THAMYRES CARDOSO DA SILVEIRA

EVOLUTIONARY HISTORY OF *Manihot carthagenensis* (EUPHORBIACEAE) AND ALLIED SPECIES IN EASTERN SOUTH AMERICA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2018

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

S587a
2018

Silveira, Thamyres Cardoso da, 1987-
Evolutionary history of *Manihot carthagenensis* (Euphorbiaceae) and allied species in eastern South America / Thamyres Cardoso da Silveira. - Viçosa, MG, 2018.
vii, 48 f. : il. (algumas color.) ; 29 cm.

Texto em inglês

.

Orientador: Luiz Orlando de Oliveira.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Microssatélites (Genética). 2. *Manihot carthagenensis* - Filogenia. 3. Filogeografia. I. Universidade Federal de Viçosa. Departamento de Bioquímica e Biologia Molecular. Programa de Pós-Graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 572.8633

THAMYRES CARDOSO DA SILVEIRA

EVOLUTIONARY HISTORY OF *Manihot carthagenensis* (EUPHORBIACEAE) AND ALLIED SPECIES IN EASTERN SOUTH AMERICA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 28 de junho de 2018.

Alberto Soares Corrêa

Eveline Teixeira Caixeta

Jéferson Nunes Fregonezi

Márcio Lacerda Lopes Martins
(Coorientador)

Luiz Orlando de Oliveira
(Orientador)

Aos meus pais, Maricélia e Gilson.

Às minhas tias, Josear e Gilda.

Dedico.

AGRADECIMENTOS

À minha mãe Maricélia por sempre acreditar, torcer e orar por mim. A painho pela confiança. A ambos pelo amor e por acreditarem sempre em meus sonhos e me apoiarem em todas as decisões.

Ás minhas tias Josear e Gilda por mostrarem que a educação é o melhor caminho a ser trilhado.

Ao professor Luiz Orlando de Oliveira pela orientação durante o mestrado e o doutorado. Muito obrigada pela paciência, por todo conhecimento compartilhado, por me fazer pensar “fora do círculo”, pela convivência respeitosa e por ser um exemplo de vida profissional.

Ao meu coorientador Márcio Lacerda Lopes Martins por sugerir que trabalhasse com o gênero *Manihot* no doutorado, pela grande ajuda e disponibilidade de sempre, por todo conhecimento transmitido, pelo incentivo, paciência e amizade.

Ao meu noivo Danilo Camêlo pelo companheirismo, amor e por sempre me incentivar a crescer profissionalmente.

Aos amigos que conheci no laboratório de Biologia Molecular e Filogeografia: Chris, Érica, Hugo, Leandro, Jefferson e Thiago. Agradeço os momentos alegres e o conhecimento que compartilhamos.

Aos amigos do Biocafé e da Botânica: Dênia, Tiago e Cristiele, muito obrigada pelo carinho e disponibilidade de ajuda, sempre que necessário.

As amigas de longa caminhada, Gabriela e Reiza. Obrigada pelo carinho e companheirismo, mesmo à distância.

À Universidade Federal de Viçosa, pela oportunidade de me capacitar profissionalmente.

Ao Programa de Pós-Graduação em Genética e Melhoramento por me proporcionar aprender com os mais qualificados professores.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa.

A todos que direta ou indiretamente me ajudaram ao longo dessa caminhada, o meu muito obrigada!

BIOGRAFIA

Thamyres Cardoso da Silveira, filha de Gilson Sampaio da Silveira e Maricélia Cardoso da Silveira, nasceu em 08 de agosto de 1987 em Ipiaú, Bahia.

Ingressou no curso de Ciências Biológicas, modalidade bacharelado em março de 2008, na Universidade Federal do Recôncavo da Bahia em Cruz das Almas, vindo a graduar-se em fevereiro de 2012.

Em março de 2012 ingressou no Programa de Pós-graduação em Botânica, área de concentração Sistemática Vegetal, da Universidade Federal de Viçosa e obteve o título de *Magister Scientiae* em fevereiro de 2014.

Em março de 2014 iniciou o Doutorado no Programa de Pós-graduação em Genética e Melhoramento, área de concentração Genética Vegetal, na Universidade Federal de Viçosa. Submeteu-se aos exames finais para aprovação da tese em 28 de junho de 2018 obtendo o título de *Doctor Scientiae* em Genética e Melhoramento.

SUMÁRIO

ABSTRACT

SILVEIRA, Thamyres Cardoso, D.Sc., Universidade Federal de Viçosa, June, 2018. **Evolutionary history of *Manihot carthagenensis* (Euphorbiaceae) and allied species in eastern South America**. Adviser: Luiz Orlando de Oliveira. Co-adviser: Márcio Lacerda Lopes Martins.

Manihot Mill. (Euphorbiaceae) is a Neotropical genus with approximately 100 species. *Manihot carthagenensis* is a very polymorphic species and associated with dry environments, mostly the caatinga and the chaco. Currently, morphological criteria associated with geographic distribution distinguish three infraspecific taxa in *M. carthagenensis*: *M. carthagenensis* subsp. *carthagenensis*, *M. carthagenensis* subsp. *glaziovii*, and *M. carthagenensis* subsp. *hahnii*. Herein, we assembled multilocus datasets with DNA sequence data obtained from four nuclear genes (sts, ch_metE, g3pdh, and nia-i3) for 34 samples of the three subspecies of *M. carthagenensis* and 14 samples from 10 closely-related species and carried out Bayesian phylogenetic analyses. We also obtained microsatellite data from 19 representative populations sampled throughout most of the known range of *M. carthagenensis* to investigate genetic structure and diversity. Our phylogenetic study suggested that *M. carthagenensis*, as presently circumscribed, does not constitute a monophyletic clade, but represents three well-differentiated lineages: *M. carthagenensis*, *M. glaziovii* and *M. hahnii*. These three lineages were supported based on morphological differences, genealogical relationships, and vegetation associations. Microsatellite data suggest that *M. carthagenensis* consists of at least three distinct gene pools partly structured according to geography. We hyphothesized that these gene pools evolved in allopatry but remained interfertile and were able to produce hybrid zones after reconnecting. Thereby the genetic admixture is of recent origin and owing to population expansion.

RESUMO

SILVEIRA, Thamyres Cardoso, D.Sc., Universidade Federal de Viçosa, junho de 2018. **História evolutiva de *Manihot carthagenensis* (Euphorbiaceae) e espécies afins no leste da América do Sul**. Orientador: Luiz Orlando de Oliveira. Coorientador: Márcio Lacerda Lopes Martins.

Manihot Mill. (Euphorbiaceae) é um gênero Neotropical com aproximadamente 100 espécies. *Manihot carthagenensis* é uma espécie muito polimórfica e associada com ambientes secos, principalmente a caatinga e o chaco. Atualmente, critérios morfológicos associados com distribuição geográfica distingue três táxons infraespecíficos em *M. carthagenensis*: *M. carthagenensis* subsp. *carthagenensis*, *M. carthagenensis* subsp. *glaziovii*, and *M. carthagenensis* subsp. *hahnii*. Reunimos conjuntos de dados multilocus com dados de sequências de DNA obtidos de quatro genes nucleares (sts, ch_metE, g3pdh, and nia-i3) para 34 amostras das três subespécies de *M. carthagenensis* e 14 amostras de 10 espécies relacionadas e realizamos análise filogenética Bayesiana. Também obtivemos dados de microssatélites para 19 populações representativas amostradas ao longo da maior parte da extensão conhecida de *M. carthagenensis*, para investigar a estrutura genética e diversidade. Nosso estudo filogenético sugeriu que *M. carthagenensis*, como atualmente circunscrita, não constitui um clado monofilético mas representa três linhagens bem diferenciadas: *M. carthagenensis*, *M. glaziovii* e *M. hahnii*. Essas três linhagens foram apoiadas com base em diferenças morfológicas, relações genealógicas e associação com a vegetação. Dados de microssatélites sugerem que *M. carthagenensis* consiste em pelo menos três *pools* gênicos distintos, parcialmente estruturados de acordo com a geografia. Hipotetizamos que esses *pools* gênicos evoluíram em alopatria mas permaneceram interférteis e foram capazes de produzir zonas híbridas após reconexão. Assim, a mistura genética generalizada que observamos é de origem recente e devido à expansão populacional.

1. INTRODUCTION

Manihot Mill. (Euphorbiaceae) comprises about 98 congeners (split across 19 sections); the geographic distribution is extensive and scattered across the Neotropics: from southern USA, to northwestern Argentina, to the Atlantic coast of eastern South America (Rogers and Appan, 1973). The most well-characterized species within *Manihot* is the domesticated cassava (*M. esculenta* Crantz ssp. *esculenta*); likely one of the most important staple food with an origin in the tropics (Alves-Pereira et al., 2018). *Manihot* is a relatively young genus; the age of the crown node of the group has been estimated at about 6.6 million years ago; the genus is believed to have experienced a recent, rapid diversification, which gave rise to many derived species (Chacon et al., 2008).

The frequent association with dry environments is one of the most intriguing ecological features of *Manihot*. The congeners are frequently associated with the Caatinga of northeastern Brazil (15 species; 11 endemic), the Cerrado of Central Brazil (38 species; 29 endemic), and the Chaco of northwestern Argentina, Paraguay and lowland Bolivia (11 species; 4 endemic) (Rogers and Appan, 1973; Duputie el al., 2011). Those dry environments constitute the "*diagonal of open formations*" (Vanzolini, 1963), which runs through eastern South America and separate the moist forests of the Amazon basin (in northern South America) from the Mata Atlântica (along the Atlantic coast of eastern Brazil).

Eastern South America is believed to have experienced intermittent episodes of climate changes; recurrently, wetter and cooler climates replaced semi-arid climates, and vice-versa (De Oliveira et al., 1999; Auler et al., 2004; Wang et al., 2004; Werneck, 2011). Those recurrent episodes of climate changes triggered vegetation shifts that likely had a profound impact on regional flora. Plant assemblages that were more adapted to the humid phases replaced those assemblages that were favored during the drier phases, and vice versa. During the last 210 kyr BP (thousands of years before the present), the region known today as the northeastern Brazil was subjected to at least 12 pulses of moister climatic regimes; recurrently, closed-canopy forests underwent retraction and the xeric shrublands of the Caatinga experienced periods of

expansion, and vice versa (Auler et al., 2004; Wang et al., 2004). Since the early Holocene, the semi-arid environment of northeastern Brazil favors the xeric shrublands of the Caatinga (Werneck, 2011). The rapid diversification of *Manihot* (Chacon et al., 2008) may have overlapped those intermittent episodes of climate changes. If so, newly derived congeners may have had several opportunities to occupy distinct niches and establish viable populations during the expansion phase of the dry environments across the "*diagonal of open formations*". During the unfavorable phases, however, local extinctions may have eliminated entire populations throughout the "*diagonal of open formations*".

Manihot carthagenensis (Jacq.) Müll. Arg. (section *Carthaginenses* Rogers & Appan) is a species that is associated with dry environments (Rogers and Appan, 1973; Allem, 2001). Currently, the species range encompasses disjunct areas, most of which are located within the "*diagonal of open formations*": the Caatinga and the Chaco (parts of Argentina, Bolivia, and Paraguay). Within Brazil, the species also occur within ecotonal zones, where the Caatinga intermingles with either the seasonal forest of the Mata Atlântica along the Brazilian coast or the Cerrado of central Brazil. The Guajira-Barranquilla xeric shrublands of northern Colombia and Venezuela constitute another disjunct range of the species (Magill et al., 2018). The current circumscription of *M. carthagenensis* encompasses three recognized sub-species, which were defined based on both morphological characters and geographic distribution: (1) *M. carthagenensis* subsp. *carthagenensis* (this subspecies occurs outside Brazil, in disjunct areas in Argentina, Paraguay, Bolivia, Colombia, and Venezuela); (2) *M. carthagenensis* subsp. *glaziovii* (Müll. Arg.) Allem (this subspecies is restricted to northeastern Brazil); and (3) *M. carthagenensis* subsp. *hahnii* Allem (a fruit-winged, narrow endemic species that is restricted to northwestern Minas Gerais state, Brazil) (Allem, 2001). The name *Manihot carthaginensis* (Jacq.) Müll. Arg. contains an epithet with an orthographic variation, thus this species name is considered invalid (Magill et al., 2018).

We hypothesize that the alternating climate changes that took place during the last 210 kyr BP in eastern South America likely shaped the population dynamics of *M. carthagenensis*. During an unfavorable cycle, groups

of interconnected populations of *M. carthagenensis* underwent range fragmentation; local extinctions wiped out many populations throughout the species range. Meanwhile, some populations retracted and survived the unstable phase within refugial areas. Under both restricted gene flow and genetic isolation within disjunct ranges, the alopatric populations accumulated novel genetic variation through genetic drift. Subsequently, distinct gene pools emerged within each of those diverging, allopatric populations. Thus far, it remains unexplored the extent of which the three current sub-species within *M. carthagenensis* would match those hypothetical, divergent gene pools.

The molecular imprints that were recorded in the genome of *M. carthagenensis* over time can be uncovered through molecular techniques (Alves-Pereira et al., 2018). Therefore, the analyses of contemporary populations of *M. carthagenensis* may disclose insights into the species' evolutionary history that remain hidden. The recency of the diversification events in *Manihot* (Chacón et al., 2008) poses enormous difficulty for gene choice as molecular markers. Most of the commonly used marker genes for either phylogeny studies (among closely-related congeners) or phylogeographic studies (among populations of closely-related species) in plants detain low levels of polymorphism across *Manihot*, particularly across *M. carthagenensis*.

There are few genus-wide molecular phylogenetic reconstructions available across *Manihot* (Chacón et al., 2008; Duputié et al., 2011); those analyses were carried out using lowly polymorphic, nuclear gene regions. The best gene choices currently available — the low copy number genes *g3pdh* and *nia-i3* — were unable to perform satisfactory during previous studies; they lacked the resolution power to resolve many parts of the phylogeny at the genus level. Moreover, the possibility of spurious relationships among paralogous may have contributed to increase the noise-to-signal ratio, particularly within a scenario in which *g3pdh* and *nia-i3* detain low levels of phylogenetic signals. Chloroplast gene regions represent no better alternative, because they too lacked sufficient variable sites to improve significantly the resolution in previous phylogenetic studies (Chacón et al., 2008; Duputié et al., 2011).

Combining novel gene regions with the existing gene choices may bring forth novel phylogenetic evidence in *Manihot*; thus, increasing the resolving power of molecular data for genus-wide phylogenetic reconstruction. Moreover,

adding different loci may contribute to a better resolution of different parts of the phylogeny (Kimball and Braun, 2014; Santos et al., 2017) and shade light into phylogeographic process among populations of closely-related species.

Orthologous genes (orthologs) arose from a common ancestral gene (Koonin, 2005; Jensen, 2001). Because they exhibit a common origin, orthologs hold the most effective and reliable molecular clues to accurately infer evolutionary relationships among species (Altenhoff et al., 2016). After a parent species go through speciation events, the degree of differentiation among orthologs across daughter species is a time-dependent process; the differentiation is dependent of the rate with which the polymorphisms occur and are retained over time (Lynch and Conery, 2000; Hanada et al., 2008; Wang, 2013). In the genome of *M. esculenta*, the cobalamin-independent methionine synthase (*metE*) gene family presents three members, which encode gene products with distinct subcellular distribution — either as cytosolic or chloroplastic forms (Rody & Oliveira, 2017). The *metE* family member that encodes the chloroplastic form of the enzyme is a single copy gene in *Manihot* (Rody & Oliveira, 2017); hereafter we will refer to this copy as the *ch_metE* gene. The stachyose synthase (*sts*) gene also is a single-copy gene in the genome of *M. esculenta* (Rody and Oliveira, unpublished data).

Herein, we assembled multilocus datasets with DNA sequence data obtained from four nuclear genes (*sts*, *ch_metE*, *g3pdh*, and *nia-i3*) for the three subspecies of *M. carthagenensis* and closely-related species and carried out Bayesian phylogenetic analyses. We also obtained microsatellite data from representative populations sampled throughout most of the known range of *M. carthagenensis* to investigate genetic structure and diversity. This study addressed the following three questions: (1) Does *M. carthagenensis* consist of a single evolutionary lineage, or is there molecular evidence to indicate divergent gene pools? (2) To which extent would the current subspecies subdivision within *M. carthagenensis* obtain support from molecular data? (3) What levels of genetic diversity and genetic structure does *M. carthagenensis* exhibit across the species range? Investigating the molecular imprints that were left in the genome of contemporary populations of *Manihot carthagenensis* and allied species may shed light into the light-known evolutionary history of the

flora associated with the dry environments, especially the Caatinga of northeastern Brazil.

2. MATERIAL AND METHODS

2.1 Taxon sampling and DNA extraction

A total of 200 samples of the *M. carthagenensis* were collected at 29 sites in Brazil (25 sites), Bolivia (3), and Colombia (1). Sample size per site varied from 1 to 15, depending upon the number of samples available (Supporting information, Table S1). Our sampling strategy covered most of the known range of the *M. carthagenensis* in Brazil. The Terrestrial Ecoregions of the World database from the World Wildlife Fund (Olson et al., 2001) and the habitat classification from the Brazilian Institute of Geography and Statistics (IBGE, 2004) defined the major vegetation type in each sampling site. Most sampling sites were located within dry environments; such as the Caatinga of northeastern Brazil and the Cerrado of central Brazil (Ecoregions NT1304 and NT0704; respectively). Samples of non-Brazilian origin came from Colombia: the Guajira-Barranquila xeric scrub (NT1308), and Bolivia: the Dry Chaco (NT0210), the Southern Andean Yungas (NT0165), and the Chiquitano dry forests (NT0212). Additionally, we added the following 14 samples of ten congeners: *M. anomala* Pohl (1 sample), *M. caerulescens* Pohl (3), *M. dichotoma* Ule (2), *M. divergens* Pohl (1), *M. esculenta* Crantz (2), *M. grahamii* Hook. (1), *M. irwinii* D. J. Rogers & Appan (1), *M. pilosa* Pohl (1), *M. quinquepartita* Huber ex D.J.Rogers & Appan (1), and *M. tripartita* (Spreng.) Müll.Arg. (1) (Supporting information, Table S2).

For each sample, leaf tissue was dried immediately using silica gel and kept at room temperature until the DNA was extracted; for herbarium specimens, leaf tissue was used directly as found in the vouchers. Total genomic DNA was extracted following a previously published protocol (Cota-Sanchez et al., 2006) with modifications (Riahi et al., 2010) and stored at -20 ^{o}C until further use.

2.2 Two single-copy nuclear genes

To gain insights into the complete sequences of the *ch_metE* and *sts* genes, information from the full genome and from the coding-DNA sequences (CDS) of *M. esculenta* were obtained from PLAZA 3.0 Dicots (http://bioinformatics.psb.ugent.be/plaza/versions/plaza_v3_dicots/). The sequence of the chloroplastic MetE protein was obtained from the *Arabidopsis thaliana* genome, available at the Protein Data Bank (PDB) under the accession number 1U1H. The sequence of the *STS* protein was obtained from the *Glycine max* genome; it was annotated as Glyma 19G40550 in Phytozome v11.0.

Using 1U1H (for the chloroplastic MetE protein) and Glyma 19G40550 (for the STS protein) as queries during tBLASTn, we searched the CDS sequences of *M. esculenta* for homologous sequences of either the MetE or STS, respectively. We trimmed the search results using an e-value cutoff of e^{-20} and filtered out sequences with less than 40% of identity. Complete nucleotide sequences of both the *sts* gene and the *ch_metE* gene of *M. esculenta*, which included introns and exons, were retrieved from the *M. esculenta* genome based on gene coordinates. To determine introns and exons regions and their boundaries in each gene, the sequences obtained from BLAST were aligned to their respective coding-DNA reference sequence using MUSCLE (Edgar, 2004) with default parameters.

The software Primer3 (Koressaar & Remm, 2007) designed primers for the independent amplification of either the *sts* gene or the *ch_metE* gene regions through the polymerase chain reaction (PCR). For primer design, the gene sequences were obtained from the genome of *M. esculenta*. The following two primer pairs were designed: MetEcp_6F (5'-TGCAGTTGACCTAGCAAATGA-3') and MetEcp_6R (5'-AGCCTCATTTGTCACCCTTG-3'), and STS_F (5'-CTCAAGATTTGGAACCTCAACA-3') and STS_R (5'-TTGATCCTTTGCTCCTTTGG-3'), for the *ch_metE* and *sts* gene regions, respectively. The software OligoAnalyser 1.0.2 (Kuulasmaa, 2002) confirmed that each of those primer pairs displayed neither potential loops nor primer-primer interactions that could decrease the efficiency of PCR amplifications. Each primer pair was designed to produce an amplicon that consisted of a

sequence of a complete intron having partial regions of exons at both flanks. This primer was designed within gene regions that would maximize the chances of finding phylogenetically-informative polymorphisms in *Manihot*.

2.3 Pcr amplification and sequencing

Independent amplifications of either ch*_metE* or *sts* gene regions were carried out in a total volume of 25 µL, with 1X buffer IVB (Phoneutria), 0.2mM of each dNTP, 0.5uM of each primer, 1.25 U of GoTaq DNA polymerase (Promega), 2 mM $MgCl_2$, 17.5 µg/mL of Bovine serum albumin (BSA) and approximately 60 ng of genomic DNA. The PCR profile for any of the gene regions was as follows: 5 min at 95ºC followed by 30 cycles of 30 s at 95ºC for template denaturation, 30 s at 60ºC for primer annealing, and 1 min at 72ºC for primer extension and final extension of 7 min at 72ºC.

In addition to the two single-copy genes — *ch_metE* and *sts*, we also amplified additional gene regions of two low-copy gene: the glyceraldehyde-3-phosphate dehydrogenase gene (*g3pdh*) and nitrate reductase gene (*nia-i3*). Those two gene regions had been useful in previous phylogenetic studies of *Manihot* (Duputié et al., 2011). The protocols, primers, and PCR conditions to amplify either *g3pdh* or *nia-i3* were performed essentially as described (Duputié et al., 2011).

All PCR products were cleaned using ExoSAP IT (USB; 3 µl of enzyme per 9 µl of reaction). Sequencing was performed by Macrogen Inc., South Korea (www.macrogen.com), using the same primers as in the PCR amplifications.

2.4 Sequence alignments and assembly of datasets

All sequences were imported into Sequencher version 4.8 (Gene Codes Corp.) for manual inspection and edition. There were sequences in which double peaks of equal height were detected, the respective site was considered as ambiguous and coded following the International Union of Pure and Applied Chemistry (IUPAC) nucleotide code for ambiguous bases. Next, multiple-sequence alignments were performed with the introduction of gaps to

compensate for the presence of insertion or deletions (indels). For phylogenetic analysis, sequences of the four nuclear loci (*ch_metE*; 669 bp; *sts*, 284 bp; *g3pdh*, 734 bp; and *nia-i3*, 612 bp) were concatenated to give rise to dataset A (N = 48; 2299 bp).

To assess the potential utility of the newly described single-copy gene regions (*ch_metE* and *sts*) for phylogenetic and phylogeographic studies, we obtained measures of nucleotide variation and compared them to those obtained for the two previously described, low-copy gene regions (*g3pdh* and *nia-i3*). Two measures of molecular diversity — number of variable sites and number of indels — were estimated directly from the sequence alignments obtained for each of the four gene regions.

The electropherograms of the ch*_metE* gene region were re-inspected visually to further characterize the intraindividual polymorphism in both the direct and reverse strands. Ambiguities owing to the presence of multiple overlapping double peaks were resolved using the haplotype subtraction technique (Clark, 1990). Subsequently, we excluded three variable sites (242, 281, and 606; having haplotype 1 as a reference sequence) because they display more than two character states across the dataset. The infinite-sites model assumes that the mutation rate within a given DNA sequence is so small that only a single mutation can occur at any given homologous position (Kimura, 1969). Those three sites violated the infinite-sites model and would introduce homoplasious relationships among haplotypes in the network. We also discarded an indel that were bordered by mononucleotide repeats (polyA; at position 190) as a source of information in subsequent statistical analyses. We coded each of the three remaining indels as a fifth character state, such that each indel was taken as a single character, irrespective of the length of the indel. The final alignment gave rise to dataset B (N = 134; 644 bp). All sequences obtained in this study were deposited in the Genbank (Supporting information, Table S3).

2.5 Phylogenetic analyses and haplotype network

The phylogenetic analysis was carried out using a Bayesian inference method, as implemented in MRBAYES v3.1.2 (Ronquist & Huelsenbeck, 2003)

using the dataset A. As outgroup, *M. grahamii* was selected based on a previous phylogenetic study of *Manihot* (Duputié et al., 2011). Before carrying out the phylogenetic analysis, the subset of sequences of each four nuclear gene regions was input separately into the software MRMODELTEST v.2.3 (Nylander, 2004). For each gene region, the Akaike Information Criterion (Akaike, 1974) indicated the best-fit models among 24 models of molecular evolution, as following: *ch_metE* and *g3pdh* (GTR+G); *nia-i3* and *sts* (HKY+G). The partitioned analysis was performed in MRBAYES using two simultaneous runs of 20 million generations each, with one cold and seven heated chains in each run. Sampling was performed once every 1000 trees with the first 250 trees discarded as burn-in samples. Average standard deviation of split frequencies at the end of each run was below 0.01. A 50% majority-rule consensus tree of the two independent runs was obtained with posterior probabilities that were equal to bipartition frequencies.

In the network analysis, gene genealogies were inferred from dataset B (information from the ch_*metE* gene region) using the median-joining network method (Bandelt et al., 1999), as implemented in NETWORK v5.0.0.3 (Fluxus Technology Ltd.) with default parameters.

2.6 Microsatellite analyses

Microsatellite genotyping were performed with the samples that belonged to the blue lineage (See Results, phylogenetic study). A total of 184 samples from 19 populations were genotyped (Supporting information, Table S1). Populations (BOL2, BOL3, CAC, ITA, JAB, MCA, MAN, and PER) detained a small sample size (N < 6) and therefore were excluded from the microsatellite analyses. The only exception was COL, a population with small sample size (n = 2) from Colombia; this population was included in the STRUCTURE analysis, as a priori identification of population is not required.

Seven nuclear microsatellite loci for *M. esculenta* that had been developed by previous studies were used: MeESSR10, MeESSE19, and MeESSR26 (Bang et al., 2011); SSRY12 and SSRY81 (Mba et al., 2001); and GA126 and GA131 (Chavarriaga-Aguirre et al., 1998) (Supporting information, Table S4). All loci were genotyped using fluorescent-labeled primers with 6-

FAM, HEX (MWG-Biotech, Ebersberg, Germany) or NED (Applied Biosystems). PCR was performed in one multiplex (MeESSR10-MeESSE19-MeESSR26) and two duplex (SSRY12-SSRY81; and GA126-GA131) systems. PCR amplification was carried out in a 12 µL reaction, with 10 ng of DNA, 1X buffer (10 mM Tris-HCl, pH 8.4, 50 mMKCl, and 1% Triton X-100), 0.2 µM of each primer, 2.5 mM of MgCl2, 0.25 mg/mL of Bovine Serum Albumin (BSA) (Invitrogen), 0.2 mM od dNTPs, and 0.5 U of Taq DNA Polymerase (PhoneutriaBiotecnologia). The amplifications were performed in a thermocycler under the following conditions: 3 min at 94ºC followed by 30 cycles of 30 s at 94ºC for template denaturation, 1 min at 58ºC for primer annealing, and 1 min at 72ºC for primer extension and final extension of 20 min at 72ºC. PCR products were separated using an ABI PRISM 3130x1genetic analyzer (Applied Biosystems). Molecular sizes in base pairs were determined using the GENESCAN-500 ROX size standard (Applied Biosystems). Alleles were scored using GeneMapper version 4.0 (Applied Biosystems).

Null allele frequencies were estimated using the MICROCHECKER, v.2.2.3 software (Van Oosterhout et al., 2004). We used the software FSTAT, v.2.9.3.2 (Goudet, 2002) to calculate the statistical significance of deviation from Hardy-Weinberg equilibrium and linkage disequilibrium between loci in all populations. The number of alleles per locus (A), number of private alleles (A_{priv}), observed heterozygosity (H_O), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}) in each population were calculated using the GDA software (Lewis & Zaykin, 2002). AMOVA was performed with ARLEQUIN v3.0 (Excoffier et al., 2005) for two hierarchical level (among and within populations) using the option F_{ST}; the parameters of molecular variance were tested using 1000 permutations. To examine genetic differentiation among populations, we calculated pairwise F_{ST} (Weir & Cockerham, 1984) among populations using the software ARLEQUIN v3.0 (Excoffier et al., 2005). The significance was confirmed via 1000 permutations of the data.

The Bayesian model-based clustering method implemented in the software STRUCTURE v.2.3.4 (Pritchard et al., 2000; Hubisz et al., 2009) inferred the number of Bayesian groups of *M. carthagenensis* using the Monte Carlo Markov Chain (MCMC) approach. We set runs with a burn-in of 250,000 steps followed by 750,000 steps, with 20 independent replications for each

number of *K*, from *K*=1 to *K*=22. To find the best *K*, we used the ΔK method (Evanno et al. 2005), as implemented in the software STRUCTURE HARVEST (Earl & VonHoldt, 2011). The data of the 20 interactions for the best *K* was converged using the software CLUMPP (Jakobsson & Rosenberg, 2007) and results were visualized using the software DISTRUCT (Rosenberg, 2004).

3. RESULTS

3.1 Two new single-copy nuclear gene markers

The gene structures of the newly described single-copy, nuclear gene regions — *ch_metE* and *sts* — provided information about the exons and flanking intronic regions (Fig. 1). Those two newly described, single-copy gene regions presented measures of molecular diversity similar to those found in *g3pdh* and *nia-i3*, the two low-copy, nuclear gene regions that have been widely used in phylogenetic studies of *Manihot* (Supporting information, Table S5). Among 12 species of *Manihot*, the number of variable sites varied from 9% (*ch_metE*) to 12% (*g3pdh*). Among populations of *M. carthagenensis*, it varied from 4% (*ch_metE*) to 8% (*nia-i3*).

3.2 Phylogenetic study

The Bayesian consensus tree depicted the phylogenetic relationships among 34 samples of the *M. carthagenensis* complex and 14 samples from 10 congeners, having *M. grahamii* as outgroup (Fig. 2). Information from the four nuclear gene regions allowed for the recovery of a tree in which most of the internal nodes showed support from posterior probability (PP) above 90%. In the consensus tree, three lineages were highlighted for reference purpose. Hereafter, each of those three lineages will be referred to as the blue, green, and red lineages, respectively. (Fig. 2).

The blue lineage contained the majority of the samples of the *M. carthagenensis* complex (*M. carthagenensis* subsp. *glaziovii* from the Caatinga of northeastern Brazil together with *M. carthagenensis* subsp. *carthagenensis* from either Colombia or Bolivia). This lineage split further into sub-clades A1

and A2. The sub-clade A1 included 22 samples from 21 sites that were collected from across a wide geographic area in Brazil, in addition to two samples from a single population (COL) in Colombia, and five samples from three populations (BOL1, BOL2, and BOL3) in Bolivia. The five samples of Bolivian origin had been identified as *M. guaranitica* Chodar & Hassl. Based on morphological characters, *M. guaranitica* had been synonymized under *M. carthagenensis* (Allem, 1979). The sister sub-clade A2 included samples from populations that were located in the extreme northern Minas Gerais State, Brazil. Based on morphology alone, the samples of sub-clade A1 were indistinguishable from the samples of the sub-clade A2.

Showing a sister relationship to the blue lineage, the green lineage comprised sequences from samples that had been identified as *M. carthagenensis* subsp. *glaziovii*. While the samples of *M. carthagenensis* subsp. *glaziovii* that belonged to the blue lineage had origin within the drier environments of the Caatinga, the samples of *M. carthagenensis* subsp. *glaziovii* that were placed within the green lineage had been collected within the more humid environments of the Atlantic rainforest (Fig. 2). The phylogenetic reconstruction placed *M. anomala* as having a sister relationship to the super-clade that comprised the blue and green lineages.

Unexpectedly, *M. carthagenensis* subsp. *hahnii* did take part in neither the blue lineage nor the green lineage. Instead, *M. carthagenensis* subsp. *hahnii* branched off and together with *M. caerulescens*, *M. tripartita* and *M. quinquepartita* formed the red lineage. None of the three congeners that clustered with *M. carthagenensis* subsp. *hahnii* in the red lineage were part of the *M. carthagenensis* complex. The red lineage was closely related to *M. dichotoma*.

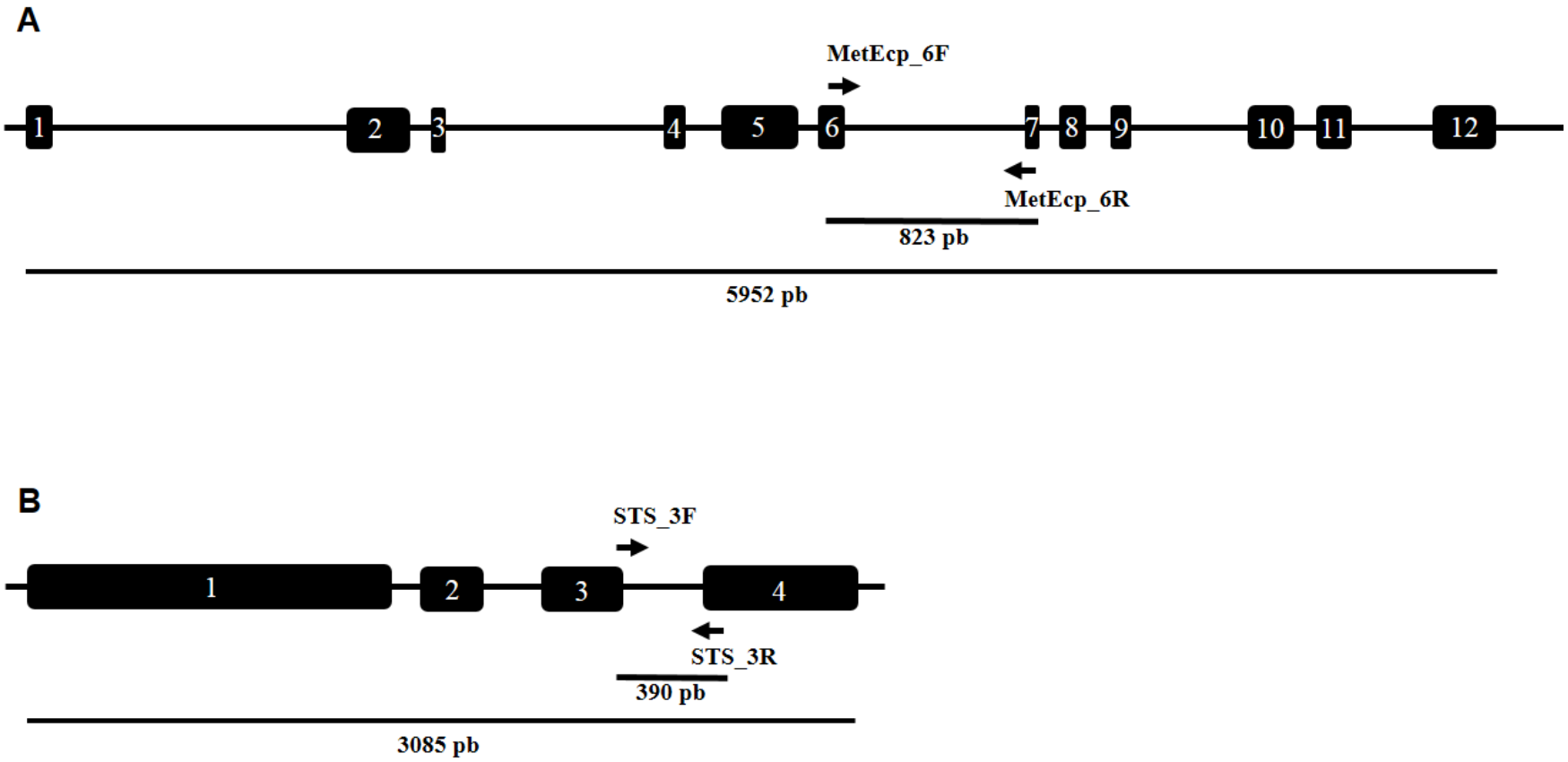

Figure 1. Schematic diagram for the two novel, single-copy genes markers in *Manihot*. (A) The cobalamin-independent methionine synthase gene (ch_*metE*). (B) The stachyose synthase gene (*sts*). Arrows indicated the approximate location and orientation of the primer pairs. The primer names are showed above each arrow. Boxes denote the coding regions (exons) and lines connecting the coding regions represent the non-coding regions (introns). The sizes of boxes and connecting lines are proportional to the number of base pairs in each sequence. The lines below each diagram inform the expected size (in base pairs) of each amplicon and the full size of each gene region.

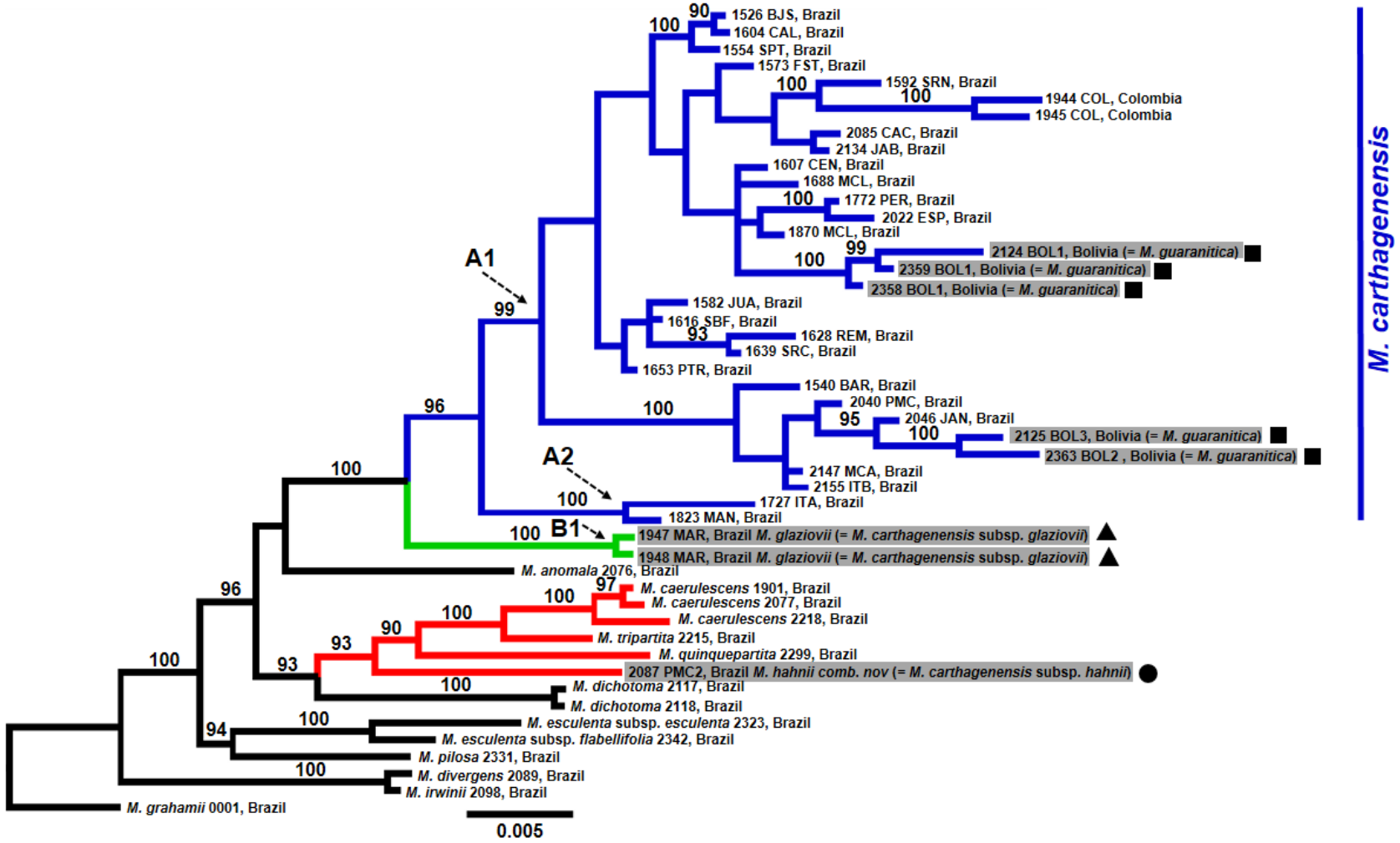

Figure 2. Bayesian phylogenetic analysis (consensus tree) showing the relationships among the *Manihot spp.,* with *M. grahamii* as outgroup. The dataset was 2299 bases long and resulted from the concatenation of four nuclear genes (*g3pdh*, *ch_metE*, *nia-i3*, and *sts*). Branch lengths are drawn to scale; nodal support values are given as posterior probabilities (%) above the branches (when ≥ 90%). Scale bar corresponds to the expected number of substitutions per site. Color in branches (blue, green, and red) according to phylogenetic lineages (see text). Color in terminals: gray highlight, samples that underwent taxonomic changes. Black squares, *M. carthagenensis* previously identified as *M. guaranitica*; black triangles, *M. glaziovii* previously identified as *M. carthagenensis* subsp. *glaziovii;* and black circle, *M. hahnii* previously identified as *M. carthagenensis* subsp. *hahnii*. Gene codes: *ch_metE*, cobalamin-independent methionine synthase; *sts*, stachyose synthase; *g3pdh*, glyceraldehyde-3-phosphate dehydrogenase; and *nia-i3*, nitrate reductase.

3.3 Haplotype network

The haplotype network recovered 26 haplotypes among the 134 sequences of the *ch_metE* gene region (Fig. 3). The analysis provided further insights into the genealogical relationships within the *M. carthagenensis* complex. Overall, the network comprised two major sets of haplotypes (haplogroups) and its topology was highly congruent with the results we had obtained with the phylogenetic analysis (Fig. 2). Therefore, we represented the haplogroups according to the color we had depicted the corresponding phylogenetic lineages.

The blue haplogroup took place around its most frequent member (haplotype 1, N = 80) and contained sequences obtained from samples that belonged to the blue lineage: *M. carthagenensis* subsp. *glaziovii* from the Caatinga of northeastern Brazil and *M. carthagenensis* subsp. *carthagenensis* from either Colombia or Bolivia. The network topology placed the green lineage immersed within the blue haplogroup. The network analysis did not distinguished *M. carthagenensis* subsp. *carthagenensis* from *M. guaranitica*; both of which were included within the blue haplogroup.

The red haplogroup comprised three rare haplotypes (20, 21, and 24) that were each recovered from a specimen of *M. carthagenensis* subsp. *hahnii* (Fig. 3, depicted in red). The distances between nearest members of distinct required a large number of mutational steps, which varied from 14 to 16.

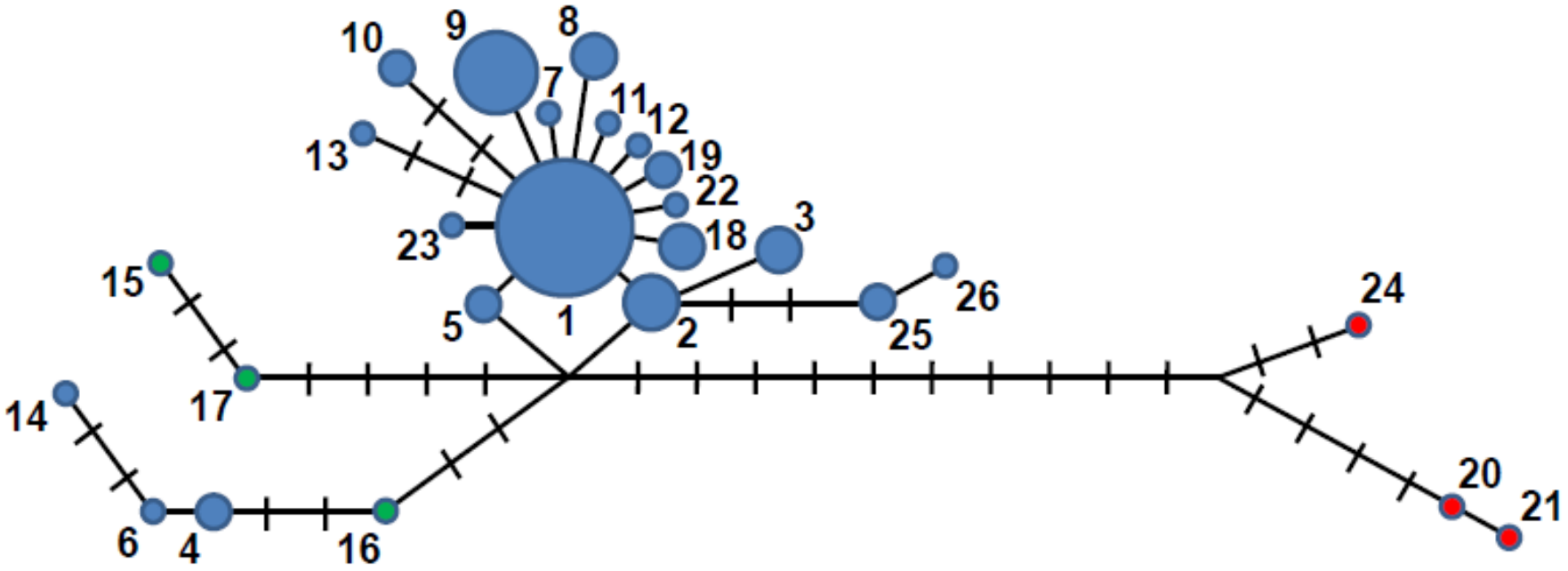

Figure 3. Median-joining network for 134 samples of *Manihot spp.*, based on sequences (644 pb) of a cobalamin-independent methionine synthase gene region (*ch_metE*). In the network, a circle represents a given haplotype (coded with numbers); size is proportional to the relative frequencies. Numbers of mutational steps are indicated with bars when there is more than one (unless indicated otherwise). Color codes: blue, *M. carthagenensis*; red, *M. hahnii* previously identified as *M. c. hahnii*; green, *M. glaziovii* previously identified as *M. c. glaziovii.*

3.4 Taxonomic treatment

Manihot carthagenensis (Jacquin) Müll.Arg., Prodr. 15(2): 1073 1866.

TYPE: COLÔMBIA, Bolivar, Cartagena, *Jacquin s/n* (Holotype P!, Isotypes photo F!, W!)

= *Manihot guaranitica* Chodat and Hassl., Bull. Herb. Boiss. ser. 2. 5 : 671. 1905.

Treelet or shrub decumbent, 1–5 m tall. Roots tuberous. Stem glabrous, smooth or with pelling periderm, cylindrical, latex white or translucent, abundant. Stipules 5–12 × 3–8 mm, deltoid, elliptical or oval, margin laciniate, generally deciduous. Leaf simple, glabrous, carthaceous to coriaceous, (3–)5(–7) lobed, margin entire or pandurate, rarely entire leaves, elliptical or oblong, central lobe 5–10 × 2–5 cm, oval, apiculate, venation camptodromous, veins glabrous, petioles 5–12 cm long, cylindrical, glabrous, green or variably purple, insertion basal, rarely peltate. Inflorescence paniculate, rarely racemose, erect to pendule, 6–20 long, glabrous, bracts linear, glabrous, 0.3–0.5 cm long, oval, margin entire or laciniate, bracteoles inconspicuous, glabrous, ca. 0.1 cm, oval, margin entire or laciniate. Staminate buds 0.5–1.2 × 0.3–0.6 cm, oval to orbicular, pedicel 5–10 mm long. Staminate flowers 1.2–1.5 × 0.6–0.8 cm, sepals 5, fused to the middle portion, greenish to purple, glabrous; stamens 10, in two whorls, ca. 1 cm long, yellowish, staminal disc yellowish. Pistillate buds1–1.2 × 0.5–0.6 cm, opposite to subopposite, in the base of inflorescence, orbicular to oval, pistillate flowers 1–1.2 × 0.5–0.7 cm, sepals 5, free, yellow-greenish to purple, glabrous, sometimes persistent, nectariferous disc yellow-purple. Capsules 1.5–2 cm diam., orbicular, surface smooth or verrucous, no ribbed, apex rounded, green. Seeds 1–1.3 × 0.5–0.7 cm, elliptical, inner face convex, brown, with dark spots; caruncle 2–3 mm long, brown, in the ventral face.

Comments: It resembles *M. glaziovii* Müll. Arg. for presenting oval or oval - shaped stipules (> 2 mm wide), leaf lobes oval,inflorescences paniculate with bracts setaceous (= 1 mm wide) and orbicular fruits and seeds with caruncula with ventral disposition. It differs, however, by the shrub habit, rare arborescent, decumbent, up to 5 m in height, by the insertion of the petiole at the base of the leaf blade and by the occurrence in arid and/or semi-arid environments, while

M. glaziovii presents arboreal habit, erect, normally above 4 m height, petiole with insertion peltate and distribution exclusively in atlantic rain forest.

Phenology: It flowers and fruits practically throughout the whole year. Under cultivation between October and June.

Distribution and habitat: It occurs in dry environments, such as cerrado and caatinga, Chaco and Chiquitano vegetation, in South America, from Colombia to Argentina, with wide distribution in the northeast and center-west of Brazil.

Conservation: The wide distribution and good representativeness of its natural populations do not show threats to this species (Bachman et al. 2011).

Material examined: BRASIL. ALAGOAS: Olho d'Água do Casado, Capelinha Farm, 01/III/2000, *D. Moura 1055 & R. A. Silva* (SP); Pão de Açúcar, Boqueirão, 09º43'44"S, 37º30'49"W, 23/II/2002, *R. P. Lyra-Lemos et al. 6180* (SP); BAHIA: Argoim, ca. 65 km N of Itaberaba, 12º34'S, 39º45'W, 07/II/1995, *A. C. Allem 4518 & V. S. Silva* (CEN); Casa Nova, BR–235, 09º25'6.6"S, 41º51'22.0"W, 10/III/2010, *M. Martins et al. 1730* (HURB); Castro Alves, road Castro Alves to Argoim, 12º41'13.9"S, 39º28'04.6"W, 202 m, 06/III/2013, *M. Martins et al. 1929* (HURB); Cotegipe, 7.2 km after Cotegipe, toward BR–242, 12º04'04.6"S, 44º18'38.8"W, 617 m, 08/III/2013, *M. Martins et al. 1938* (HURB); Dom Basílio, ca. 2 km of Lagoa do Morro Village, ca. 13km toward Livramento de Nossa Senhora, 13º34'23"S, 47º04'14"W, 02/XI/2010, *L. P. Queiroz 15223 et al.* (HUEFS); Feira de Santana, BR–116, toward Ipirá, 12º16'22.7"S, 38º59'41.8"W, 08/III/2010, *M. Martins et al. 1714* (HURB); Iaçu, BA–048, 12º50'14.7"S, 39º51'38.5"W, 18/IV/2010, *M. Martins et al. 1639* (HURB); Ipirá, BA–052, toward Bom Jesus de Lapa, 12º11'6.1"S, 39º28'39.0"W, 18/IV/2010, *M. Martins et al. 1720* (HURB); Jacobina, BR–324, Ponte do Roncador, 11º12'22.7"S, 40º25'51.1"W, 08/III/2010, *M. Martins et al. 1721* (HURB); Maniaçu, road out Itanagé, 13º47'21.4"S, 42º07'19.2"W, 20/IV/2010, *M. Martins et al. 1646* (HURB); Remanso, BR–235, 9º34'51.1"S, 42º7'36.2"W, 10/III/2010, *M. Martins et al. 1731* (HURB); Riachão das Neves, settlement after the Pedras Brancas River, BR–135, 18 km toward Barreiras/Riachão, road in the left, 6.5 km, 11º57'37.5"S, 44º58'46.4"W, 07/III/2013, *M. Martins et al. 1933* (HURB); Riacho de Santana, BR–430, toward Bom Jesus de Lapa, 13º37'34.5"S, 42º35'0.6"W, 18/IV/2010, *M. Martins et al. 1647* (HURB); Santa Teresinha, BR–493, toward Itatim, 12º42'26.2"S, 39º41'05.9"W, 18/IV/2010, *M. Martins et al.*

1632 (HURB); Sento Sé, Volta da Serra Village, 10º03'51.5"S, 42º12'51.5"W, 12/III/2010, *M. Martins et al. 1736* (HURB); Xique-Xique, BA–160, km 31, 300 m before the clover to Santo Inácio, 11º06'7.5"S, 42º44'18.3"W, 406 m, 08/III/2013, *M. Martins et al. 1939* (HURB). CEARÁ: Aiuaba, Estação Ecológica de Aiuaba, 06º36'01"S, 40º07'15"W, 27/V/2004, *J. R. Lemos 214 & P. Matias* (SPF). PERNAMBUCO: Buíque, Catimbau Valley, 12/II/2008, *D. N. Silva et al. 16* (SPF); Fernando de Noronha, Hill of Quixaba, 08/IV/1999, *A. M. Miranda 3199* (HST); Floresta, Itapemirim Farm, 14/V/2008, *J. Ferraz et al.* (SPF 193.596); Pesqueira, Saw of Gavião, 19/VI/2005, *M. Oliveira 1814* (UFP); Petrolina, 5 km after airport, toward Casa Nova, 09º22'3.1"S, 40º34'33.6"W, 09/III/2010, *M. Martins et al. 1727* (HURB); Santa Maria de Boa Vista, 04/II/1998; *W. Ribeiro* (SP 329.744); Serra Talhada, UFRPE, Saw of Macacos, road to Açude do Saco, 25/II/1999, *M. Alves et al EBNN 1545* (SP); Sobradinho, Highway Sobradinho/Sento Sé, 9º33'14.2"S, 40º57'58.7"W, 11/III/2010, *M. Martins et al. 1733* (HURB); PIAUÍ: Monsenhor Gil, ca. 14 km SW of city, along the BR–316, toward Picos, 05º42'S, 42º36'W, 01/II/1995, *A. C. Allem 4482 & V. S. Silva* (CEN); São Raimundo Nonato, São João Vermelho, 22/I/1984, *L. Emperaire* (RB 367.022); Teresina, 36 km NW of Picos, along BR–316, 06º53'S, 41º48'W, 24/I/1993, *W. W. Thomas et al. 9604* (NY). RIO GRANDE DO NORTE: Acari, Sítio Talhados, km 17, BR 427, toward Currais Novos-Acari, 06º19'53"S, 36º37'29"W, 26/II/2011, *A. A. Roque et al. 900* (UFRN); Grossos, Salina Salmar, 29/VI/2007, *A. A. Roque 240* (UFRN); João Câmara, Cauaçu Farm, 05º33'51"S, 35º55'57"W, 13/I/2011, *J. G. Jardim et al. 5918* (UFRN); Mossoró, Distrito de Alagoinha, Experimental Farm Rafael Fernandes, 02/IV/2007, *J. E. de Araújo & M. L. Silva* (HUEFS 150392); Punaú, 05º24'81"S, 35º23'24"W, 02/VI/2009, *A. C. P. Oliveira 999 & J. G. Jardim* (UFRN); São João do Sabugi, RN 118, ca. 5,5 km of city, 18/III/2011, *A. A. Roque 932* (UFRN); São Miguel do Gostoso, Novo Horizonte, 05º16'79"S, 35º44'44"W, 14/V/2007, *G. B. C. Paterno & M. I. B. Loiola 153* (UFRN); Serra Negra do Norte, Estação Ecológica do Seridó, 17/IV/1999, *R. G. V. Camacho 21* (SPF). SERGIPE: Canindé do São Francisco, Poço Verde Farm, 27/III/2000, *D. Moura 1099 & R. A. Silva* (SP). GOIÁS: Trindade, ca. 22 km SE Araripina, along BR–316, toward Trindade, 07º41'S, 40º21'W, 02/II/1995, *A. C. Allem 4484 & V. S. Silva* (CEN). MATO GROSSO DO SUL: Corumbá, Santa Tereza

Rural School, 07/XII/1976, *A. C. Allem 667* (CEN); Três Lagoas, Floresta Farm, 20/IX/1964, *J. C. Gomes Jr. 2205* (SPF). MINAS GERAIS: Arcos, Faroeste Farm, right margin of the São Miguel River, 20º15'05.7" S, 45º39'46.7" W, 31/XII/2002, *P. H. A. Melo 310 & E. A. Melo Jr* (SP); Descoberto, Reserva Biológica do Grama, 29/I/2001, *R. C. Forzza et al. 1766* (SP); Janaúba, BR–122 km 16, Lagoa Santa, APA Carste de Lagoa Santa, 30/XI/1995, *A. E. Brina & L. V. Costa* (SPF 137.951); Janauba-Porteirinha, 16º25'S, 43º33'W, 08/XI/1984, *A. C. Allem et al. 2888* (CEN); Januária, 32.3 km to NE of the city, along BR–135, 15º13'S, 44º13'W, 16/II/1995, *A. C. Allem 4582 & V. S. Silva* (CEN); Pains, Amargoso Farm, MG–439 km 16, 20º23'14"S, 45º38'59"W, 30/XII/2003, *P. H. A. Melo 283* (SP); Porteirinha, ca. 16 km NE of city, along BR–122, toward Mato Verde, 15º35'S, 43º00'W, 04/XII/1987, *A. C. Allem 3700 & W. L. Werneck* (CEN); Salto da Divisa, Santana Farm, 16º03'29.8"S, 40º01'50.6"W, 19/II/2003, *J. A. Lombardi et al. 5056* (SP). SÃO PAULO: Cardoso, former Porto Militão, 18/V/1995, *Bernacci et al. 1848* (SPF). **BOLÍVIA**. COCHABAMBA: Mizque, Julpe, 20 km S of Totora, 17º48'29"S, 65º17'34"W, 15/XII/2002, *P. Labiak et al. 2879* (SPF).

Manihot glaziovii Müll.Arg., Fl. Bras. 11(2): 446. 1874.

TYPE: Glaziou 1022: Brazil, Rio de Janeiro: near Rio de Janeiro, cultivated (F!, G)

= *Manihot carthagenensis* subsp. *glaziovii* (Müll.Arg.) Allem, Novon 11: 160, 2001. TYPE: BRASIL, Rio de Janeiro, 'environs of Rio de Janeiro, cultivated, *Glaziou 1022* (Syntype F!, G). *syn. nov.*

Tree 4–9 m tall, erect. Roots tuberous. Stem glabrous, smooth or with pelling periderm, cylindrical, latex white or translucent, abundant. Stipules 5–10 × 3–7 mm, deltoid, elliptical or oval, margin laciniate, generally deciduous. Leaf simple, glabrous, commonly coriaceous, 3(–5) lobed, rarely entire leaves, margin entire, lobes elliptical to oboval, apiculate, central lobe (5–)10–15 × (3–)5–8 cm, venation camptodromous, veins glabrous, petioles (5–)10–20 cm long, cylindrical, glabrous, green, insertion peltate. Inflorescence paniculate, rarely racemose, erect, 6–12 long, glabrous, bracts linear, glabrous, 0.3–0.5 cm long, oval, margin entire or laciniate, bracteoles inconspicuous, ca. 0.1 cm, glabrous, oval, margin entire or laciniate. Staminate buds 1–1.2 × 0.6–0.8 cm, oval to

orbicular, pedicel 5–10 mm long. Staminate flowers 1.2–1.5 × 0.6–0.8 cm, sepals 5, fused to the middle portion, greenish to purple, glabrous; stamens 10, in two whorls, ca. 1 cm long, yellowish, staminal disc yellowish to reddish. Pistillate buds1–1.2 × 0.5–0.6 cm, opposite to subopposite, in the base of inflorescence, orbicular to oval, pistillate flowers 1–1.5 × 0.5–0.8 cm, sepals 5, free, yellow-greenish to purple, glabrous, sometimes persistent, nectariferous disc yellow-purple. Capsules ca. 2 cm diam., orbicular, surface smooth or verrucous, no ribbed, apex rounded, glabrous, green. Seeds 1.2–1.5 × 0.5–0.8 cm, elliptical, inner surface convex, brown, with dark spots; caruncle 2–3 mm long, brown, in the ventral face.

Comments: For Allem (2001) the morphological similarity of *M. glaziovii* Müll. Arg. with *M. carthagenensis* Müll. Arg. (discussed in the comments on the latter) is sufficient for its recognition as subspecies, with only geographical distinction. Our data, however, support its recognition as a distinct species, sister of *M. carthagenensis*. Characteristics such as arboreal habit, erect, normally above 4 m height and insertion of petiole peptum, associated with its geographical distribution, exclusively in Atlantic rain forest, allow its distinction.

Phenology: Flowers and fruits in July and August.

Distribution and habitat: It occurs in fragments of atlantic rain forest, in altered areas near pastures, in the states of Bahia and Paraíba, Brazil.

Conservation: The Analysis of Occupancy Area (AOO) and Extent of Occurrence (EOO) from the available data indicates the Endangered status for this species (segundo Bachman et al. 2011) with classification EN A1c,B1abii,iii, 2abii.

Material examined: BRAZIL. Bahia: Maragogipe, road to Maragogipe–São Roque, 12º47'36.7"S, 38º55'51.2"W, 08/VIII/2012, *M. Martins et al. 2070* (HURB); *ibidem*, road to Maragogipe–São Roque, ground road on the left, back of the house, 12º47'40.4"S, 38º55'50.1"W, 08/VIII/2012, *M. Martins et al. M101* (HURB); *ib.*, road to Maragogipe–São Roque, 12º47'36.7"S, 38º55'51.2"W, 28/VII/2015, *M. Martins et al. M234* (HURB). PARAÍBA: Gurinhem, 9 km east of the junction of BR–230 and road to Alagoa Grande, toward João Pessoa, 07º10'S, 35º34'W, 12/VIII/1985, *A. C. Allem 3243 & W. L. Werneck* (CEN).

Manihot hahnii (Allem) M. Martins & T. Silveira, *comb. nov.*

Basionime: *Manihot carthagenensis* subsp. *hahnii* Allem. TYPE: BRAZIL. Minas Gerais: Januária, exatly 2.9 km after exit to ferryboat in the highway BR–135, to downtown, 15º29'S, 44º23'W, 470 m, 16 February 1995, *A. C. Allem 3687 & W. L. Werneck* (holotype CEN!, isotypes G, K, M, MBM, MO, NY, UB, US) (Fig. 4 and Fig. 5).

Tree 1.5–7 m alt. Roots not tuberous. Stem glabrous, smooth; branches cylindrical, latex cream, not abundant. Stipules 1–1.5 × 0.3–0.5 cm, oval to oblong, margin entire or laciniated, glabrous, decicuous. Leaf simple, glabrous, carthaceous, (6–)8(–12) lobed, with 3–7 developed lobes and 3–5 reduced lobes, between the two basal lobes, near to insertion of the petiole into the leaf lamina;developed lobes oboval, rarely elliptical, sometimes overlapping, apex acute and acuminated, margin entire, sometimes pandurated, base of lamina larger than 0.4 cm wide, central lobe (6–)8–10(–15) × (2–)3–4(–6) cm, nervation camptodromous, petioles (5–)7–15(–20) cm long, cylindrical, glabrous, greenish to purple, basal insertion to slightly peltated. Inflorescence paniculate or 1–4 racemes, terminal, erect or pendule, 6–10(–20) cm long, bracts lanceolate, apex acute, margin entire, 3–7(–10) × 1–3 mm, glabrous, bracteoles inconspicuous, 0.3 × ca. 0.1 cm, glabrous, margin entire. Staminate buds 0.3–0.8 × 0.3–0.4 cm, ovoids to pyramidal, sometimes slightly constrict in the middle portion, pedicel 5–10 mm long. Staminate flower 9–15 × 4–6 mm, sepals fused to 1/3, 5 lobes, apex acute, glabrous, yellow-greenish; stamens 10, in two cycles, 9–12 mm long, whitish; staminal disc yellowish. Pistillated buds not seen. Pistillated flowers 2, at base of inflorescence, sometimes isolated, 1–1.2 × 0.3–0.4 cm, pedicel 1–2 cm, sepals free, glabrous, yellow-greenish, apex acute, ovary globose to ovoid, slightly ribbed, glabrous, green, disc nectariferous not seen. Capsules 1.6–2(–2.2) × 1.5–1.8 cm, orbicular to ovoid, apex rounded, smooth surface to variably ribbed. Seeds 8–10 × 6–8 cm, ovoid, inner surface convex; caruncle ventral, inconspicuous.

Comments: The material type of *Manihot hahni* was originally described as *M. carthagenensis* subsp. *hahni* by Antônio Costa Allem (Allem 2001). The author, however, did not observe the presence of three to five lobes atrophied at the base of the limbus, in the region of the insertion of the petiole in the leaf. The

presence of these wolves is for the first time detected in *Manihot* species and allows the distinction of *M. hahni* from all species of the genus.

Phenology: Flowers and fruits of in February and December.

Distribution and habitat: There are records of *Manihot hahni* in the northwest region of the state of Minas Gerais, in areas of arboreal Caatinga, always on edges of forest fragments.

Conservation: The data obtained from the herbaria visited were checked on expeditions carried out in July 2013 and March 2016. Only the population of the 'old ferry' was located, however, with only one individual. In other areas, vegetation was degraded by road or urban expansion, with few remnants of the original vegetation. Thus, *M. hahni* was considered Endangered (EN) due to an Occurrence Extension of less than 5,000 km2 and an Occupancy Area of less than 500 Km2 (according to Bachman et al., 2011) in addition to occurring in a fragmented environment, to have its population decline observed and decrease in its extent of occurrence and area of occupation (ENB12a,bi, ii,2a,bi, ii).

Material examined: BRAZIL. Minas Gerais: Januária, 32, 3 km NE of the city exit, along road BR-135 toward Itacarambi, 15º13'00", 44º13'00"W, 500m, 16/II/1995, *A. C. Allem & V. S. Silva 4581* (CEN). *Ibidem.*, 8,3 km ahead of do ponto de cruzamento do rio São Francisco River, toward Januária, BR-135, 15º31'00"S, 44º24'00"W, 15/II/1995, *A. C. Allem & V. S. Silva 4571* (CEN). *Ib.*, ca. 5 km SW of Januária, along road BR–135, toward Lontra, 15º50'00"S, 44º20'00"W, 550 m, 02/XII/1987, *A. C. Allem 3687 & W. L. Werneck* (CEN). *Ib.*, exact 9,6 km ahead of São Francisco River toward Januária., along road BR-135, exact 2,9 km before first break springs of city, 15º29'00"S, 44º23'00"W, 16/II/1995, *A. C. Allem & V. S. Silva 4574* (CEN). Mirabela, exact 5 km NW of the last acess to Mirabela, along road BR–135 toward Januária, 16º13'00"S, 44º13'00"W, 820m, 17/II/1995, *A. C. Allem & V. S. Silva 4584* (CEN). Pedras de Maria da Cruz, 1 km after landing place of the old ferry, toward BR-135, 15º35'28"S, 44º23'45"W, 446m, 13/II/2016, *M. Martins et al*. *M247* (HURB).

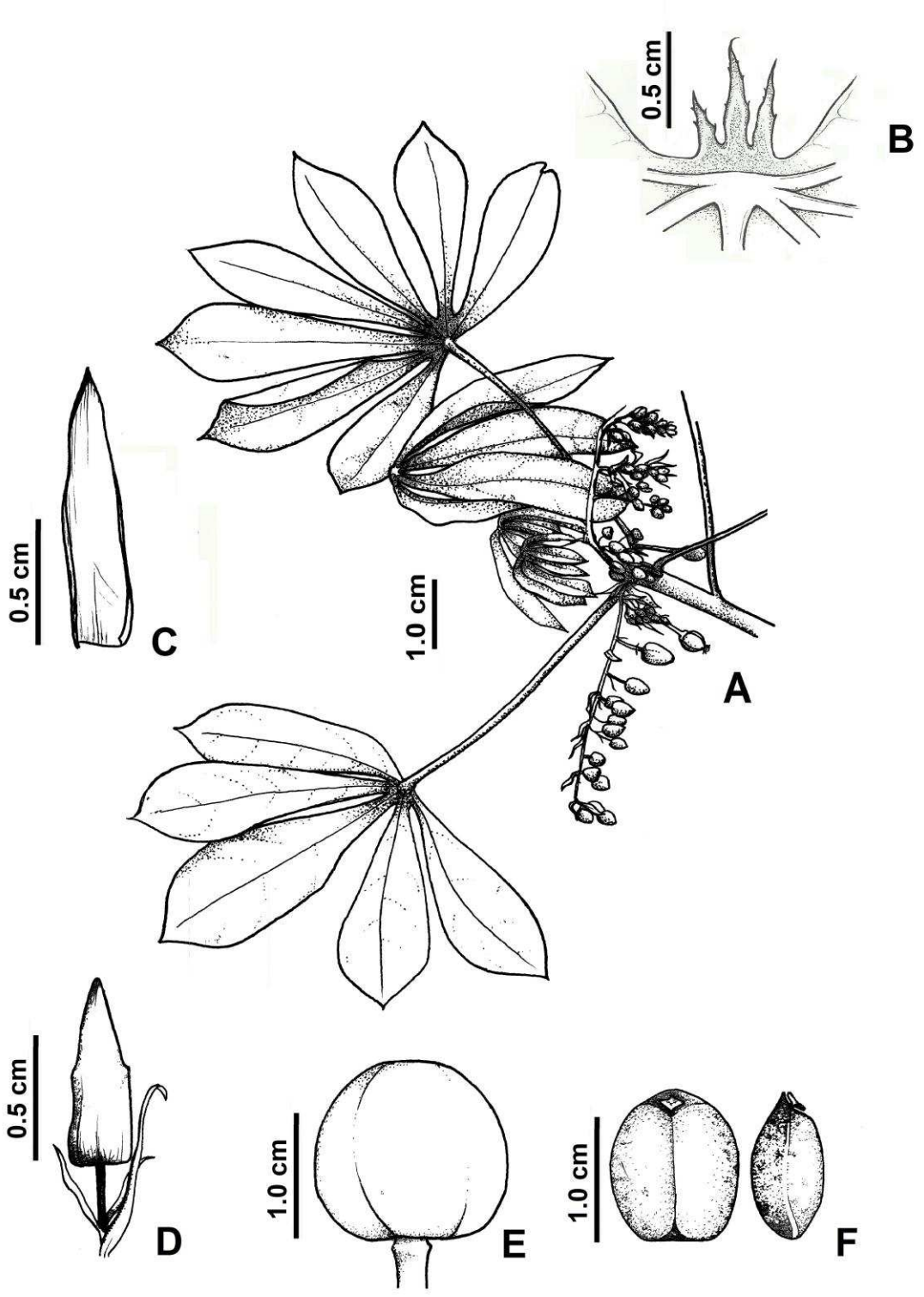

Figure 4. *Manihot hahnii*. A. Flowering branch. B. Reduced leaf lobes. C. Bract. D. Staminate bud with bract and bracteoles. E. Fruit. F. Seed, frontal and lateral view (A. C. Allem 3687 & Werneck, holotype).

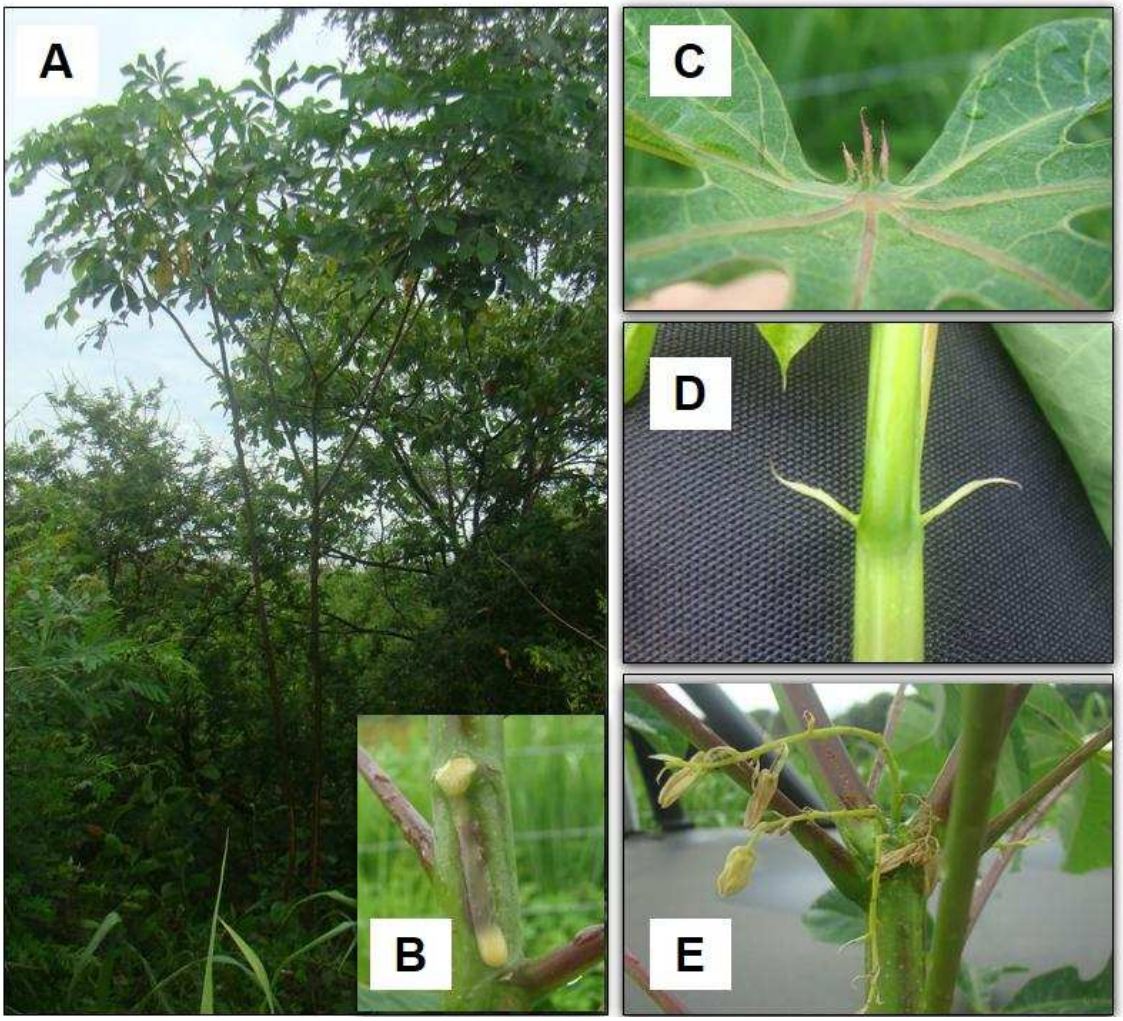

Figure 5. *Manihot hahnii*. A. Habit. B. latex; note the cream color. C. Reduced leaf lobes. D. Stipules. E. Flowering branch.

3.5 Genetic structure of *Manihot carthagenensis*

Our analysis uncovered 81 alleles across seven microsatellite loci. The number of alleles per locus ranged from three (GA-126) to 19 (SSRY81 and SSRY12). None of the seven loci exhibited null allele frequency above 0.2. (Supporting information, Table S4). Linkage disequilibrium was absent; none of the seven loci showed significant deviations from the Hardy-Weinberg equilibrium.

The overall mean H_O (0.38) was lower than H_E (0.44); (Supporting information, Table S6). Among populations, H_O ranged from 0.18 (MCL) to 0.56 (JUA); while H_E ranged from 0.25 (MCL) to 0.64 (JUA). The inbreeding coefficient (F = 0.09) ranged from –0.16 (BOL) to 0.29 (MCL). A total of 17 private alleles were identified in 13 populations: BJS, SPT, FST, JUA, CAL, CEN, SBF, SRC, ESP, JAN, and MCL (a single private allele per populations); PTR (two private alleles); and BOL (four private alleles) (Supporting information, Table S6)

Results of the AMOVA revealed that differences within populations accounted for 79.5% of the total variance. Overall mean F_{ST} (0.20) computed over microsatellite loci were high, ranging from 0.05 to 0.40. In general, pairwise F_{ST} values were higher in comparisons involving populations located at the extremes of the species range (Supporting information, Figure S1)

Results of the Bayesian model-based clustering method implemented in STRUCTURE suggested that *M. carthagenensis* consisted of three Bayesian groups (best K = 3; Fig 6A). The Bayesian groups were depicted white, gray, and black, respectively (Fig. 6B and 6C). Each of the 184 samples contained ancestry in at least two of the Bayesian groups. In nine populations, most of the samples showed the highest proportions of membership assigned to one of the groups, with a small proportion of assignment in the two remaining groups. Samples of populations COL, SRC, CEN, PMC, and MCL had the highest proportions of membership assigned to the white Bayesian group, with assignment proportion ranging from 73% (SRC) to 91% (PMC). With assignment proportion over 94%, the gray group dominated populations BJS and BAR, while the black group was prevalent in populations FST and SPT (over 90%).

The geographic placements of the three Bayesian groups do not look randomly distributed over the sampling area of *M. carthagenensis*. The white group prevails in those populations that were sampled at the westernmost part of the species range, while the gray group and the black group dominated the populations sampled in the northern and eastern parts of the range, respectively (Fig. 6C). Overall, the levels of genetic admixture among groups were high within populations. Admixture seems widespread in those populations located within the Caatinga, as for example in SRN, CAL, REM, PTR, JUA, and SBF. Admixture in those six populations resulted from a triple origin, with a mean assignment of 47% to the gray group, 37% to the black group, and 16% to the white group.

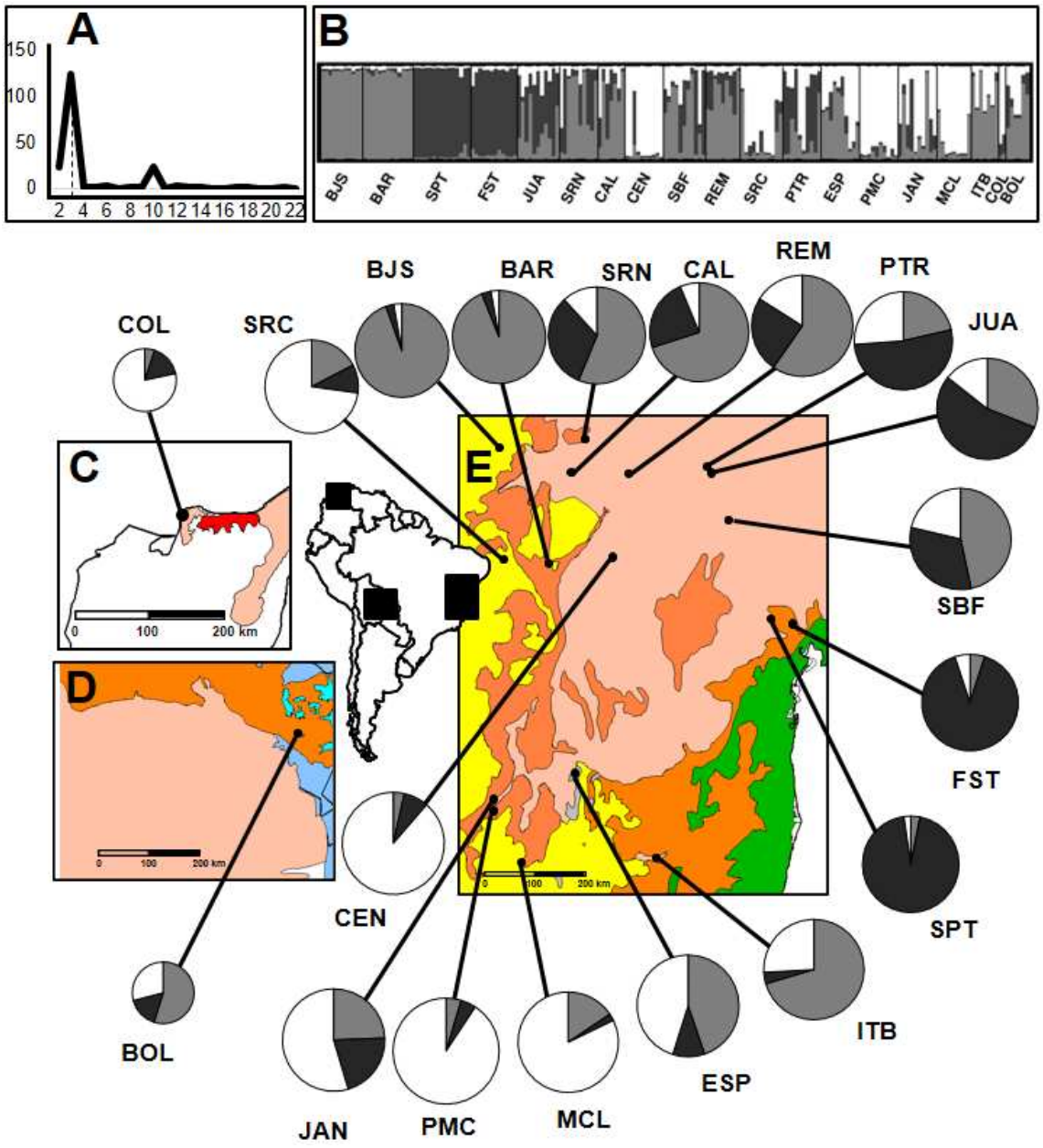

Figure 6. Clustering analyses and geographic distribution the three Bayesian groups uncovered in *M. carthagenensis*. The analyses were based on seven microsatellite markers obtained from 184 samples across 19 populations. (A) The best *K* (*K* = 3) was calculated according to the *ΔK* method (Evanno et al. 2005). (B) Plot of the clustering analysis in STRUCTURE, showing the three Bayesian groups (coded as white, gray, and black). Along the x-axis, each vertical bar represents a sample; along the y-axis, membership coefficient of a sample for a Bayesian group represents the fraction of its genome that has ancestry in that Bayesian group. (C) Geographical origin of the Colombian population (COL). (D) Geographical origin of the Bolivian population (BOL). (E) Geographical origin of the 17 Brazilians populations. Each pie diagram represents the sum of membership coefficients for all samples from that population. Small circles denote two populations (COL and BOL) with small sample sizes (N ≤ 6); large circles denote populations with large sample sizes (N > 6). (Refer to Supporting information, Table S1 for population codes).

4. DISCUSSION

4.1 Molecular evidence does not support the current infraspecific subdivision within *M. carthagenensis*

The current circumscription of *M. carthagenensis* together three subspecies, which has been defined based on both morphological characters and geographic distribution: (1) *M. carthagenensis* subsp. *carthagenensis*; (2) *M. carthagenensis* subsp. *glaziovii*, and (3) *M. carthagenensis* subsp. *hahnii* (Allem, 2001). Our phylogenetic study suggested that *M. carthagenensis*, as presently circumscribed, does not constitute a monophyletic clade. Based on morphological differences, genealogical relationships, and vegetation associations; we decided to eliminate the current subspecies subdivision within *M. carthagenensis*. We also choose to elevate two of its current subspecies to the species rank (see below).

4.2 *Manihot carthagenensis* (the blue lineage)

A certain degree of distinction is required to identify a subspecies: “The subspecies, or geographical race, is a geographically localized subdivision of the species, which differs genetically and taxonomically from other subdivisions of the species” Mayr (1942). Contrary to previous assumptions (Allem, 2001), geographic distribution seem to be an irrelevant criteria to define infraspecific taxa within *M. carthagenensis*. It was clear that the samples of *M. carthagenensis* subsp. *glaziovii* that were collected in the Caatinga exhibited neither phylogenetic nor taxonomic features that distinguished them apart from the samples of *M. carthagenensis* subsp. *carthagenensis* with either Colombian or Bolivian origin. Therefore, we recommend that *M. carthagenensis* (the blue lineage) should remain as a single taxonomic entity, with no further division at the infraspecific level.

Manihot guaranitica is a species with occurrence in Bolivia (Rogers e Appan, 1973). Previously, it has been synonymized under *M. carthagenensis* (Allem, 1979). Latter, *M. guaranitica* was taken as synonymous under the subspecies *M. carthagenensis* subsp. *carthagenensis* (Allem, 2001). Our

phylogenetic reconstruction suggested that *M. guaranitica* was more closely related to *M. carthagenensis* subsp. *glaziovii* from the Caatinga than to *M. carthagenensis* subsp. *carthagenensis* from Colombia. Thus, our result challenged the current taxonomic placement of *M. guaranitica* as a synonymous of *M. carthagenensis* subsp. *carthagenensis*. This finding adds to the mounting evidence that points to a floristic link between two drier environments in South America, the Caatinga and the Chaco an vegetations of lowland Bolivia (Pennington et al. 2006). Thus, *M. guaranitica* should remain as a synonymous under *M. carthagenensis*, as an early study had suggested (Allem, 1979).

4.3 *Manihot glaziovii* (the green lineage)

Based on morphological and ecological traits, the phylogenetic evidence (Fig. 2), and haplotype network analysis (Fig. 3); we decided to resurrect the species rank *M. glaziovii* and applied it to those populations of *M. carthagenensis* subsp. *glaziovii* that detained arboreal habit and were associated with humid environments of the moist forests of the Mata Atlântica biome (the green lineage). The sister relationship between *M. glaziovii* and *M. carthagenensis* suggested that the diversification event from a common ancestor likely took place recently during the evolutionary history of *Manihot* in Eastern South America. We hypothesize that *M. carthagenensis* became associated with drier environments, while *M. glaziovii* became associated with more humid environments. *Manihot carthagenensis* presents a habit that varies from treelets and decumbent shrubs and insertion of the petiole at the base of the leaf lamina, while *M. glaziovii* presents tree and erect habit and peltate petiole insertion. Thus, the distinct morphological traits each of those two sister species developed may be related to the adaptive responses to the distinct levels of water availability each species found in those two contrasting environments.

Previously, *M. epruinosa* Pax & K. Hoffm., *M. johannis* Pax, *M. glaziovii* Mull. Arg., and *M. pseudoglaziovii* Pax & Hoffm have been synonymized under *M. carthagenensis* subsp. *glaziovii* (Allem, 2001). The plethora of names was attributed to populations regardless of the environments in which they were found. They include populations that occur in the dry environments of the

Caatinga (*M. epruinosa*, *M, johannis*, and *M. pseudoglaziovii*) and populations that occur in moist forests of the Mata Atlântica biome (now resurrected as *M. glaziovii*). Those taxa shared many morphological similarities, mostly related to the shape of stipules and bracts, inflorescence type and fruits (Rogers and Appan, 1973). However, the arboreal habit, peltate leaves, and the occurrence in more humid environments set *M. glaziovii* apart from *M. epruinosa*, *M. johannis*, and *M. pseudoglaziovii*, which should remain under *M. carthagenensis*.

4.4 *Manihot hahnii* comb. *nov*. (the red lineage)

One of the subspecies formerly attributed to *M. carthagenensis* — *M. carthagenensis* subsp. *hahnii* — clearly belonged within neither the blue lineage nor the green lineage (Fig. 2 and Fig. 3). In addition to the molecular evidence, the presence of a regularly winged and slightly larger fruit (Allem, 2001) supported *M. hahni* M. Martins & T. Silveira comb, nov. as a taxon morphologically and genetically distinct from both *M. carthagenensis* and *M. glaziovii*. Another autapomorphic character in *M. hahnii* was the presence of three to five reduced lobes, located between the two basal lobes and near to insertion of the petiole into the leaf lamina.

Manihot hahni shared the semi-leafless stipules, the leaf with oval lobes to oboval, and the globose fruits with *M. carthagenensis*. Most likely, those shared characters arose as homoplasies; this trend has been considered of generalized occurrence within *Manihot* (Cervantes-Alcayde et al., 2015).

Manihot hahni shared with *M. quinquepartita* and *M. caerulescens* a tree habit (erect to scandens) and a leaf shape (oval to oboval lobes) (Rogers and Appan, 1973). However, there were several reproductive traits that set *M. hahni* apart from the other members within the red lineage, such as the presence of capsularis fruits in *M. hahni* and baccaceo fruits in both *M. quinquepartita* and *M. caerulescens* (Rogers e Appan, 1973). The presence of sub-shrubby to shrubby habit and racemose inflorescence with typically oval and foliace bracts (>5mm wide) in *M. tripartita* distinguished it from *M. hahni* (Rogers and Appan, 1973).

The close phylogenetic relationships among *M. hahnii*, *M. quinquepartita*, *M. tripartita*, and *M. caerulescens* (the red lineage; Fig. 2) matched the known geographic distribution of those species. They occur in ecotonal zones located towards central Brazil (Northern Minas Gerais state), where three distinct ecosystems comes together (Ab'Saber, 2003; Arruda, 2003). Across the region, the scrublands of the Caatinga find its southernmost distribution and intermingle with the savanna-like Cerrado (advancing from the West) and the Mata Atlântica dry forests (advancing from the East). In *Manihot*, previous studies suggested that the phylogenetic relationships among congeners follow geography and ecology very closely (Deputié et al. 2011); this suggestion seems to apply to the members of the red lineage.

4.5 *Manihot carthagenensis* underwent extensive admixture

The Bayesian model-based clustering analyses indicated that *M. carthagenensis* detained some levels of population structure. To a certain extent, populations with close geographic proximity tended to exhibit the largest fraction of their genome into the same Bayesian group (white, gray, or black; Fig 6C). Within a given population, however, most of the samples detained fractions of its genome with ancestry in more than one Bayesian group (multicolor bars; Fig. 6B). One likely mechanism that can account for this distribution of genetic diversity is recent genetic admixture among formerly allopatric Bayesian groups.

The likely origin of the Bayesian groups we uncovered within *M. carthagenensis* lay on populations that remained trapped within separate refugial areas during episodes of unfavorable climate shifts. Under restricted gene flow, each set of alopatric populations diverged independently through genetic drift. Novel genetic variations accumulated independently, giving rise to distinct gene assemblages within each refugial area over time; we uncovered and reported herein three of such assemblages — or Bayesian groups. Both a single episode and sequential episodes of climate shifts may be plausible causes for the origin of the Bayesian groups we uncovered witin *M. carthagenensis*.

The genomic imprints we recovered allow us to postulate that *M. carthagenensis* retracted to at least three distinct refugial areas throughout its evolutionary history. Assuming that the largest proportion of ancestry within contemporary populations of *M. carthagenensis* suggests the geographic origin of the refugial areas, we hypothesize at least three, independent refugial areas: (1) the northernmost part of the range of the Caatinga of northeastern Brazil — populations with the highest proportion of ancestry in the gray group; (2) the Central Brazil, where the Caatinga meets together with Cerrado (advancing from the west) and the Atlantic seasonal forests (advancing from the south) — populations with the highest proportion of ancestry the white group; (3) the eastern coast of Brazil, where Caatinga intermingle with both seasonal forests and humid forests of the Mata Atlântica Biome — populations with the highest proportion of ancestry in the black group.

The lack of precise information from the fossil record, the absence of a reliable molecular clock, and the recency of the diversification events prevented us from estimating the times of divergence among the Bayesian groups with confidence. However, the biogeographic evidence (De Oliveira et al. 1999; Auler et al. 2004; Wang et al. 2004; Werneck, 2011) allowed us to postulate that *M. carthagenensis* likely diversified around the late Pleistocene and the early Holocene. Recent episodes of climate changes during the Last Glacial Maximum (LGM; about 21 kyr BP) favored the expansion of the vegetation types associated with drier environments, such as the Caatinga and seasonal forests (Pennington et al. 2000). Likely, the range of *M. carthagenensis* also expanded during the LGM and formerly isolated gene pools were able to reconnect. Because populations of trapped within distinct refugial areas remained interfertile, they could interbreed after expansion; thus giving rise to the widespread admixture we detected in this study.

While distinct gene pools of *M. carthagenensis* remained interfertile and were able to produce hybrid zones after reconnecting, *M. hahnii* and M. *glaziovii* remained each genetically isolated. More likely, the vicariance events that gave rise to *M. hahnii*, and later to M. *glaziovii* promoted not only geographic isolation but also genetic isolation through time to an extent that speciation became complete.

4.6 Two new sources of molecular variation in *Manihot*

The two new regions we report herein — *ch_metE* and *sts* — performed well in our studies. The combined molecular evidence of the four nuclear gene regions — *ch_metE*, *sts*, *g3pdh*, and *nia-i3* — allowed us to uncover phylogenetic details of the evolutionary history of *Manihot* that remained hidden during previous phylogenetic reconstructions (Chacón et al. 2008; Duputié et al., 2011). As an extra bonus, those two new markers are each truly single-copy nuclear genes; thus, their orthologous relationships are validated in the *Manihot* genome. We anticipate that the two gene markers we described herein will be useful in shedding further light into the phylogenetic relationships among other congeneric lineages. Additionally, the single-copy gene *ch_metE* allowed us to uncover previously unexplored genealogical details within a set of three closely-related lineages that had remained hidden within the poorly known *M. carthagenensis*. Thus, we anticipate that the fine-scale probing of other parts of the phylogeny across *Manihot* may benefit from using the *ch_metE* gene region as a molecular marker.

5. REFERENCES

Ab'Sáber, A. N. 2003. Os domínios da natureza no Brasil: potencialidades paisagísticas. São Paulo: Ateliê Editorial,159 p.

Akaike, H. 1973. Information theory and an extension of maximum likelihood principle. Pages 267-281 in: Second International Symposium on Information Theory. B. N. Petrov and F. Csaki, eds. Akademiai Kiado, Budapest.

Allem, A. C. 1979. Notas taxonômicas e novos sinônimos em espécies de *Manihot* – V (Euphorbiaceae). Rev. Brasil. Biol. 39, 735-738.

Allem, A. C. 2001. Three new infraspecific taxa of *Manihot* (Euphorbiaceae) from the Brazilian Neotropics. Novon. 11 (2), 157-165.

Altenhoff, A. M., Boeckmann, B., Capella-Gutierrez, S., Dalquen, D. A., DeLucca, T., Forslund, K., Huerta-Cepas, J., Linard, B., Pereira, C., Pryszcz, L. P., et al. 2016. Standardised benchmarking in the quest for orthologs. Nat Methods. 13 (5), 425-430.

Alves-Pereira, A., Clement, C. R., Picanço-Rodrigues, D., Veasey, E. A., Dequigiovanni, Ramos, S. L. F., Pinheiro, J. B., Zucchi, M. I. 2018. Patterns of nuclear and chloroplast genetic diversity and structure of manioc along major Brazilian Amazonian rivers. Annals of Botany. 121, 625-639.

Arruda, M. B. 2003. Estudo de representatividade ecológica com base na biogeografia de biomas e ecoregiões continentais do Brasil. O caso do bioma Cerrado. Tese de Doutorado, Universidade de Brasília, Brasília.

Auler, A. S., Wang. X., Edwards, R. L., Cheng, H., Cristalli, P. S., Smart, P. L., Richards, D. A. 2004. Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. J. Quat. Sci. 19, 693-701.

Bachman, S., Moat, J., Hill, A. W., Torre, J., Scott, B. 2011. Supporting red list threat assessments with GeoCAT: Geospatial conservation assessment tool. ZooKeys. 150, 117-126.

Bandelt, H. J., Forster, P., Rohl, A. Median-joing. 1999. Median-Joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37-48.

Bang, T. C., Raji, A. A., Ingelbrecht, I. L. A. 2011. A multiplex microsatellite marker kit for diversity assessment of large cassava (*Manihot esculenta* Crantz) germoplasm collections. Plant. Mol. Biol. Reporter. 29 (3), 655-662.

Cervantes-Alcayde, M., Olson, M. E., Olsen, K. M., Eguiarte, L. E. 2015. Apparent similarity, underlying homoplasy: Morphology and molecular

phylogeny of the North American clade of *Manihot*. Am. J. Bot. 102 (4), 520-532.

Chacón, J., Madrinán, S., Debouck, D. Rodriguéz, F., Tohme, J. 2008. Phylogenetics patterns in the genus *Manihot* (Euphorbiaceae) inferred from analyses of nuclear and chloroplast DNA regions. Mol. Phylogenet. Evol. 49, 260-267.

Chavarriaga-Aguierre, P., Maya, M. M., Bonierbale, M. W., Kresovich, S., Fregene, M. A., Tohme, J., Kochert, G. 1998. Microsatellites in cassava (*Manihot esculenta* Crantz): discovery, inheritance and variability. Theor. Appl. Genet. 97, 493-501.

Clark, A. G. 1990. Inference of haplotypes from PCR-amplified samples of diploid populations. Mol. Biol. Evol. 7, 11-122.

Cota-Sánchez, J. H., Remarchuk, K., Ubayasena, K. 2006. Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. Plant Mol. Biol. Report. 24 (2), 161-167.

De Oliveira, P. E., Barreto, A. M. F., Suguio, K. 1999. Late pleistocene/holocene climatic and vegetational history of the Brazilian caatinga: the fossil dunes of the middle São Francisco River. Palaeogeogr. Palaeoclimatol. Palaeoecol. 152, 319-337.

Duputié, A., Salick, J., Mckey, D. 2011. Evolutionary biogeography of Manihot (Euphorbiaceae), a rapidly radiation Neotropical genus restricted to dry environments. J. Geogr. 38 (6), 1033-1043.

Earl, D. A., vonHoldt, B. M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Res. 4 (2), 359-361.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high through-put. Nucleic. Acids. Res. 32, 1792-1797.

Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14 (8), 2611-2620.

Excoffier, L., Laval, G., Schneider, S. 2005. Arlequin: An integrated software package for population genetics data analysis, version 3.0. Evol. Bioinform. Online. 1, 47-50.

Goudet, J. 2002. FSTAT, version 2.9.3.2. A program to estimate and test gene diversities and fixation índices.

Hanada, K., Zou, C., Lehti-Shiu, M. D., Shinozaki, K., Shiu, S. H. 2008. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant. Physiol. 148, 993-1003.

Hubisz, M. J., Falush, D., Stephens, M., Pritchard, J. K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9, 1322-1332.

IBGE, 2004. Mapa de Vegetação do Brasil. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

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

Kimball, R. T., Braun, E. L. 2014. Does more sequence data improve estimates of galliform phylogeny? Analyses of a rapid radiation using a complete data matrix. PeerJ. 2, e361.

Kimura, M. 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics. 61, 893-903.

Koonin, E. V. 2005. Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. 39, 309-38.

Koressaar, T., Remm, M. 2007. Enhancements and modifications of primer design program Primer 3. Bioinformatics. 23 (10), 1289-1291.

Kuulasmaa. T. 2002. Oligo Analyser 1.0.2. Distributed by the author. Kuopio.

Lewis, P. O. Zaykin, D. 2002. Genetic Data Analysis (GDA): computer program for the analysis of allelic data – version 1.0.

Lynch, M., Conery, J. 2000. The evolutionary fate and consequences of duplicate genes. Science. 290, 1151-1154.

Magill, B., Solomon, J., Stimmel, H. 2016. Tropicos specimen data. Missouri Botanical Garden. Occurrence dataset https://doi.org/10.15468/hja69f acessed via GBIF.org on 2018-06-08.

Mayr, E. 1942. Systematics and the origin of species. New York: Columbia University Press, 334p.

Mba, R. E. C., Stephenson, P. Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J. Fregene, M. 2001. Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. Theor. Appl. Genet. 102, 21-31.

Nylander, J. A. A. 2004. MrModeltest v2. Computer program distributed by author. Evolutionary Biology Ceentre, Uppsala University.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Louks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., Kassem, K. R. 2001. Terrestrial ecoregions of the world: a new map of life on earth. Bio Science 51, 933–938.

Pennington, R. T., Prado, D. E., Pendry, C. A., 2000. Neotropical seasonally dry forests and quaternary vegetation changes. J. Biogeogr. 27, 261–273.

Pennington, R. T., Ratter, J. A., Lewis, G. P., 2006. Neotropical savannas and seasonally dry forests: Diversity. Biogeography and Conservation. CRC Press, Boca Raton, pp. 1–30.

Pritchard, J. K., Stephens, M., Donnelly, P. 2000. Inference of population structure using multilocus genotype data, version 2.3.4. *Genetics* 155, 945 – 959.

Riahi, M., Zarre, S., Maassoumi, A. A., Attar, F., Kazempour Osaloo, S. 2010. An inexpensive and rapid method for extracting papilionoid genomic DNA from herbarium specimens. Genet. Mol. Res. 9, 1334-1342.

Rody, H. V. S., Oliveira, L. O. 2018. Evolutionary history of the cobalamin-independent methionine synthase gene family across the land plants. Mol. Phylogenet. Evol. 120, 33-42.

Rogers, D. J., Appan, S. G. 1973. *Manihot* Manihotoides (Euphorbiaceae). Flora Neotropica, 13, 1-272.

Ronquist, F., Huelsenbeck, J. P. 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure, version 1.1. Mol. Ecol. Notes 4, 137– 138.

Santos, L., Alves, A., Alves, R. 2017. Evaluating multi-locus phylogenies for species boundaries determination in the genus *Diaporthe*. PeerJ. 5, e3120.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., Shipley, P. 2004. MICRO-CHECKER: soft ware for identifying and correcting genotyping errors in microsatellite data, version 2.2. Mol. Ecol. Notes. 4, 535–538.

Vanzolini, P. E. 1963. Problemas faunisticos do cerrado. *In* M. G. Ferri [ed.], Simposio sobre o Cerrado, Sao Paulo, 1963, 307-320. Universidade de São Paulo, Sao Paulo, Brazil.

Wang, G., Eltahir, E. A. B., Foley, J. A., Pollard, D., Levis, S. 2004. Decadal variability of rainfall in the Sahel: results from the coupled GENESIS-IBIS atmosphere-biosphere model. Climate Dynamics. 22, 625-637.

Wang, Y. 2013. Locally duplicated ohnologs evolve faster than nonlocally duplicated ohnologs in *Arabidopsis* and rice. Genome Biol. Evol. 5, 362–9.

Weir, B. S., Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38, 1358-1370.

Werneck, F. 2011. The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. Quat. Sci. Rev. 30, 1630-1648.

6. SUPPORTING INFORMATION

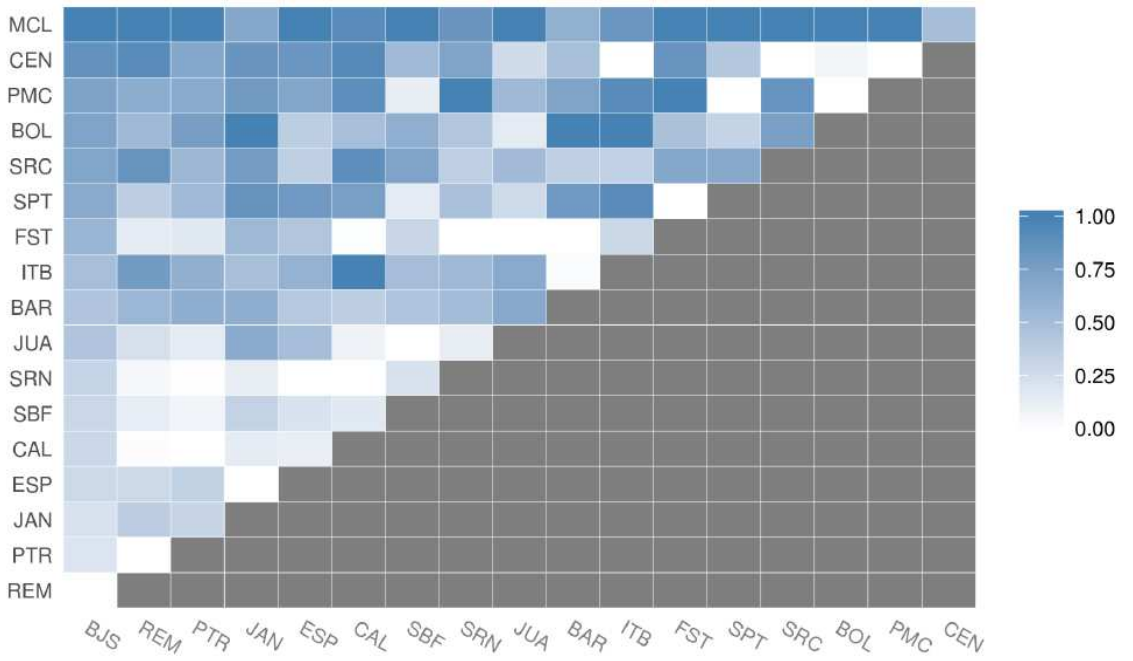

Figure S1. Heatmap plot of pairwise F_{ST} values estimated with microsatellite data from seven loci for 19 populations of *Manihot carthagenensis*. Only populations with N ≥ 6 are included. Refer to (Supporting information, Table S1 for populations size).

Table S1. Study populations of *M. carthagenensis, M. glaziovii e M. hahnii*, with locality, geographic coordinates, samples sizes (*N*, number of samples used in phylogenetic analyses; *N'*, number of samples used in SSR analyses; *N"*, number of samples used in haplotype network analyses), accession numbers, and voucher information.

Locality (population code)	Coordinates (lat/long)	N/N'/N"	Accession numbers	Voucher information	Former classification	Present classification
Barra, BA, Brazil (BAR)	S11°08'49.1"/W43°24'09.2"	1/13/5	From 1535 to 1547	Martins et al. M221	*Manihot carthagenensis* subsp. *glaziovii* *	*Manihot carthagenensis*
Bom Jesus, BA, Brazil (BJS)	S09°01'51.6"/W44°18'20.4"	1/11/6	From 1524 to 1534	Martins et al. M210	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Calculé, BA, Brazil (CAC)	S14°22'01.3"/W42°14'26.5"	1/0/0	2085	Martins et al. M236	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Campo Alegre de Lourdes, BA, Brazil (CAL)	S09°28'34.9"/W42°59'19.0"	1/7/4	From 1599 to 1604 and 1647	Martins et al. M204	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Central, BA, Brazil (CEN)	S11°03'08.9"/W42°13'34.7"	1/10/6	From 1605 to 1614	Martins et al. M222	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Feira de Santana, BA, Brazil (FST)	S12°16'38.2"/W38°59'45.4"	1/12/5	From 1563 to 1569 and 1571 to 1576	Martins et al. M225	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Itacarambi, MG, Brazil (PER)	S15°04'44.3"/W44°05'14.3"	1/0/0	1772	T. Silveira et al. 102	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Itaobim, MG, Brazil (ITB)	S16°34'08.8"/W41°27'27.3"	1/7/4	From 2153 to 2159	Martins et al. M265	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Janaúba, MG, Brazil (JAB)	S15°47'01.2"/W43°16'51.3"	1/0/4	2134 and from 2139 to 2141	Martins et al. M259	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Januária, MG, Brazil (JAN)	S15°29'34.3"/W44°23'44.7"	1/10/3	From 2041 to 2050	Martins et al. M250	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Juazeiro, BA, Brazil (JUA)	S09°28'44.5"/W40°28'32.1"	1/11/4	From 1577 to 1587	Martins et al. M199	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Magdalena, Colombia (COL)	N11°14'31"/W74°12'19"	2/2/2	1944 and 1945	CTH 415-5T and CTH 417-2D (CIAT)	*M. carthagenensis* subsp. *carthagenensis* *	*M. carthagenensis*
Itacarambi, MG, Brazil (ITA)	S15°15'82.6"/W44°13'70.4"	1/0/2	1727 and 1729	T. Silveira et al. 101	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Manga, BA, Brazil (MAN)	S14°54'85.1"/W43°59'09.3"	1/0/3	1822, 1823and 1824	T. Silveira et al. 103	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Matias Cardoso, MG, Brazil (MCA)	S14°52'32.3"/W43°55'32.3"	1/0/4	2143 and from 2146 to 2151	Martins et al. M257	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Espinosa, MG, Brazil (ESP)	S15°00'39.7"/W42°55'55.6"	1/10/5	From 2021 to 2030	Martins et al. M238	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Montes Claros, MG, Brazil (MCL)	S16°40'14.5"/W43°53'29.3"	2/9/4	1688, 1870 and from 2054 to 2062	Martins et al. M244	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Pedras de Maria da Cruz, MG, Brazil (PMC)	S15°41'52.0"/W44°23'08.0"	1/10/5	From 2031 to 2040	Martins et al. M247	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Petrolina, PE, Brazil (PTR)	S09°22'03.1"/W40°34'34.5"	1/10/5	From 1648 to 1657	Martins et al. M013	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Remanso, BA, Brazil (REM)	S09°29'54.4"/W41°56'40.2"	1/9/6	From 1627 to 1635	Martins et al. M201	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Santa Rita de Cássia, BA, Brazil (SRC)	S11°05'06.8"/W44°13'41.5"	1/11/6	From 1636 to 1646	Martins et al. M219	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*

Table S1. Continuation

São Raimundo Nonato, PI, Brazil (SRN)	S08°52'44.5"/W42°44'34.0"	1/10/5	From 1589 to 1598	M. Martins et al. M206	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Senhor do Bonfim, BA, Brazil (SBF)	S10°20'43.1"/W40°10'05.4"	1/11/5	From 1615 to 1625	Martins et al. M198	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Serra Preta, BA, Brazil (SPT)	S12°10'35.6"/W39°23'55.8"	1/15/5	From 1549 to 1588	Martins et al. M224	*M. carthagenensis* subsp. *glaziovii* *	*M. carthagenensis*
Santa Cruz, Cordillera, Bolivia (BOL)	S20°26'45"/W63°37'41"	1/6/4	2124, 2126, 2127 2358, 2359 and 2362	M. Mendoza et al. 5152	*Manihot guaranitica* **	*M. carthagenensis*
Santa Cruz, Andrez Ibañez, Bolivia (BOL2)	S18°10'00"/W63°30'56"	1/0/0	2363	M. Mendoza 2376	*Manihot guaranitica* **	*M. carthagenensis*
Santa Cruz, German Busch, Bolívia (BOL3)	S18°39'08"/W58°29'27"	1/0/0	2125	M. Mendoza et al. 4582	*Manihot guaranitica***	*M. carthagenensis*
Maragojipe, BA, Brazil (MAR)	S12°47'36.7"/W38°55'51.2"	2/0/0	1947 and 1948	Martins et al M100	*M. carthagenensis* subsp. *glaziovii* *	*M. glaziovii*
Pedras de Maria da Cruz, MG, Brazil (PMC2)	S15°35'27.8"/W44°23'47.1"	1/0/0	2087	Martins et al. M248	*M. carthagenensis* subsp. *hahnii* *	*M. hahnii*

Table S2. Study congeners of *M. carthagenensis, M. glaziovii e M. hahnii*, with accession numbers, locality, geographic coordinates, and voucher information.

Accession numbers	Taxon	Locality	Coordinates (lat/long)	Voucher information
2076	*M. anomala*	Itacarambi, MG, Brazil	S15°10'38.0"/W44°12'51.0"	M. Martins et al. M251
1901	*M. caerulescens*	Joaquim Felício, MG, Brazil	S17°41'01.9"/W44°05'34.9"	T. Silveira et al. 101
2077	*M. caerulescens*	São João das Missões, BA, Brazil	S14°51'04.9"/W44°04'07.1"	M. Martins et al. M256
2218	*M. caerulescens*	Cocos, BA, Brazil	S14°10'43.8"/W44°34'17.8"	M. Martins 2111
2117	*M. dichotoma*	Manuel Vitorino, BA, Brazil	S14°08'00"/W40°16'00"	A. C. Allem & V. S. Silva 4520
2118	*M. dichotoma*	Cruz das Almas, BA, Brasil	S12°40'12"/W39°06'60.84"	M. Martins et al. 1422
2323	*M. esculenta* subsp. *esculenta*	Viçosa, MG, Brazil	S20°45'14.040"/W42°52'54.840"	T. Silveira, 100
2342	*M. esculenta* subsp. *flabellifolia*	Cruz das Almas, BA, Brasil	S12°40'12"/W39°06'60.84"	Martins et al. 1413
0001	*M. grahamii*	Curitibanos, SC, Brazil	S27°16'58.08"/W50°35'30.84"	V. dos Santos 109
2098	*M. irwinii*	Cocalzinho de Goiás, Go	S15°49'08"/W48°43'21"	M. Mendoza 4948
2331	*M. pilosa*	Caparaó, MG, Brazil	S20°20'00"/W41°49'00"	L. B. Bianchetti et al. 1255
2299	*M. quinquepartita*	Itacarambi, MG, Brazil	S15°08'17.2"/W44°14'54.8"	M. Martins et al. M253
2215	*M. tripartita*	Mirabela, MG, Brazil	S16°14'31.2"/W44°10'33.3"	M. Martins et al. M245
2089	*M. violacea* subsp. *divergens*	Cocalzinho de Goiás, Go	S15°45'56"/W48°43'29"	M. Mendoza 4903

Table S3. Taxa, Acession number, and GenBank accession numbers for specimens included in phylogenetic analysis.

Taxon	Acession number	GenBank acession number			
		g3pdh	ch_metE	nia-i3	sts
Manihot carthagenensis	1526	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1540	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1554	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1573	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1582	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1592	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1604	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1607	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1616	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1628	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1639	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1653	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1688	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1727	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1772	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1823	In Progress	In Progress	In Progress	In Progress
-Manihot carthagenensis	1870	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1944	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	1945	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2022	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2040	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2046	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2085	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2124	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2125	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2134	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2147	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2155	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2358	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2359	In Progress	In Progress	In Progress	In Progress
Manihot carthagenensis	2363	In Progress	In Progress	In Progress	In Progress
Manihot caerulescens	1901	In Progress	In Progress	In Progress	In Progress
Manihot caerulescens	2077	In Progress	In Progress	In Progress	In Progress
Manihot caerulescens	2218	In Progress	In Progress	In Progress	In Progress
Manihot dichotoma	2117	In Progress	In Progress	In Progress	In Progress
Manihot dichotoma	2118	In Progress	In Progress	In Progress	In Progress
Manihot divergens	2089	In Progress	In Progress	In Progress	In Progress
Manihot esculenta subsp. *esculenta*	2323	In Progress	In Progress	In Progress	In Progress
Manihot esculenta subsp. *flabellifolia*	2342	In Progress	In Progress	In Progress	In Progress
Manihot hahnii	2087	In Progress	In Progress	In Progress	In Progress
Manihot grahamii	0001	In Progress	In Progress	In Progress	In Progress
Manihot irwinii	2098	In Progress	In Progress	In Progress	In Progress
Manihot pilosa	2331	In Progress	In Progress	In Progress	In Progress

Table S3. Continuation

Manihot quinquepartita	2299	In Progress	In Progress	In Progress	In Progress
Manihot tripartita	2215	In Progress	In Progress	In Progress	In Progress

Table S4. Description of the seven microsatellite loci used in the present study. Codes: $N_{A,}$ number of alleles per locus; bp, base pairs.

Locus	Primer Sequence (5'-3')	Array	Fluorescence	N_A	Allele Size (bp)	Null Allele Mean Frequency	Reference
GA-126	F: AGTGGAAATAAGCCATGTGATG R: CCCATAATTGATGCCAGGTT	$(GA)_{126}$	NED	3	205-209	-0.0102	Chavarriaga-Aguirre et al. (1998)
GA-131	F: TTCCAGAAAGACTTCCGTTCA R: CTCAACTACTGCACTGCACTC	$(GA)_{131}$	6-FAM	10	82-100	-0.0447	Chavarriaga-Aguirre et al. (1998)
MeESSR10	F: TTTTTCCCTGCTGGCTTAGA R: TTGCAGCACCCATACTGAAG	$(AT)_7$	6-FAM	7	129-143	0.0014	Bang et al. (2011)
MeESSR19	F: TTCTCGTCGGCTCCTTTCTA R: CCCCACTTGATCTGCCTTTA	$(AT)_{10}$	HEX	7	183-237	0.0430	Bang et al. (2011)
MeESSR26	F: CGGAAATGACGAAAGAAAGG R: AATTCCAATTCCACCCACAC	$(CT)_{10}$	NED	16	224-266	0.0597	Bang et al. (2011)
SSRY12	F: AACTGTCAAACCATTCTACTTGC R: GCCAGCAAGGTTTGCTACAT	$(CA)_{19}$	NED	19	206-298	-0.0286	Mba et al. (2001)
SSRY81	F: GGCGATTTCATGTCATGCTT R: TGATTTTCTGCGTGATGAGC	$(GA)_{22}$	FAM	19	173-213	0.0575	Mba et al. (2001)

Table S5. Summary of nucleotide variation in aligned sequences to two datasets to four nuclear genes.

Characteristics	Taxonomic group	*G3pdh*	*ch_metE*	*nia-i3*	*sts*
Fragment size	*Manihot* spp.	734	669	612	284
	M. carthagenensis	697	628	520	283
Number of sequences	*Manihot* spp.	46	45	42	44
	M. carthagenensis	31	30	30	28
Variable sites (%)	*Manihot* spp.	89 (12)	58 (9)	66 (11)	30 (11)
	M. carthagenensis	32 (5)	24 (4)	41 (8)	17 (6)
Number of indels	*Manihot* spp.	13	10	15	2
	M. carthagenensis	7	1	12	0

Table S6. Within-population genetic diversity in *M. carthagenensis* across seven microsatellite loci*. Codes: Codes: *n*, number of samples evaluated; *A/locus*, mean number of alleles per locus; A_{priv}, number of private alleles; H_E, expected heterozygosity; H_O, observed heterozygosity; and *F*, inbreeding coefficient.

Populations	*n*	A_{priv}	*A/locus*	H_E	H_O	*F*
Barra, BA, Brazil (BAR)	13	0	3.9	0.46	0.43	0.07
Bom Jesus, BA, Brazil (BJS)	11	1	3	0.39	0.35	0.09
Campo Alegre de Lourdes, BA, Brazil (CAL)	7	1	3.9	0.53	0.45	0.16
Central, BA, Brazil (CEN)	10	1	3.4	0.55	0.46	0.17
Espinosa, MG, Brazil (ESP)	10	1	3.6	0.51	0.41	0.21
Feira de Santana, BA, Brazil (FST)	12	1	3.6	0.47	0.44	0.07
Itaobim, MG, Brazil (ITB)	7	0	2.6	0.41	0.42	-0.03
Januária, MG, Brazil (JAN)	10	1	3.4	0.47	0.37	0.22
Juazeiro, BA, Brazil (JUA)	11	1	4.9	0.64	0.56	0.13
Montes Claros, MG, Brazil (MCL)	9	1	1.7	0.25	0.18	0.29
Pedras de Maria da Cruz, MG, Brazil (PMC)	10	0	3	0.50	0.49	0.02
Petrolina, PE, Brazil (PTR)	10	2	4.9	0.60	0.47	0.22
Remanso, BA, Brazil (REM)	9	0	4	0.48	0.44	0.09
Santa Rita de Cássia, BA, Brazil (SRC)	11	1	3.6	0.46	0.42	0.08
São Raimundo Nonato, PI, Brazil (SRN)	10	0	3.9	0.55	0.47	0.14
Senhor do Bonfim, BA, Brazil (SBF)	11	1	4.4	0.59	0.46	0.23
Serra Preta, Ba, Brazil (SPT)	15	1	3.1	0.45	0.34	0.25
Santa Cruz, Cordillera, Bolivia (BOL)	6	4	3.5	0.46	0.52	-0.16
Overall sum	182	17				
Overall mean	10.11	0.94	3.33	0.44	0.38	0.09