DEPARTMENT OF BIOLOGICAL AND
ENVIRONMENTAL SCIENCES

SPECIES DELIMITATION IN *GENIPA* (RUBIACEAE) USING INTEGRATIVE TAXONOMY

Rhonda Ridley

Degree project for Master of Science (120 hec) with a major in Biodiversity & Systematics

BIO 707, MSc Nordic Academy of Biodiversity and Systematic Studies (60 hec)

Second cycle

Semester/year Autumn 2020 – Spring 2022

Supervisors: Dr. Christine D. Bacon & Dr. Claes Persson

Examiner: Prof. Bengt Oxelman, Department of Biological and Environmental Sciences

Photo credit front-page: Claes Persson

Table of Contents

Abstract

The genus *Genipa* L. is a widespread, lowland, Neotropical lineage of trees in the coffee family, Rubiaceae. There is long-standing disagreement on the number of species that should be recognised in the genus. Here, I use an integrative taxonomy approach encompassing genomic, morphological and distribution data to resolve the classification of *Genipa*. A comprehensive species phylogeny was produced under the multi-species coalescent model, using a high-resolution dataset from target sequence capture data. Results from a 245 loci dataset strongly supports *Genipa spruceana* Steyerm., often synonymised with *Genipa americana* L., as a distinct monophyletic species. Similarly, the monophyly of *Genipa infundibuliformis* Zappi & Semir is also strongly supported. The species delimitation of *Genipa spruceana* is corroborated by morphological data and *Genipa infundibuliformis* is corroborated by morphological and distribution data. The phylogeny also shows that the widespread species *G. americana* has three distinct well-supported clades within it. These are interpreted as three independently evolving lineages. However, following an integrative taxonomy approach no new classification is recommended at this point, until reliable determination can be made with evidence other than genomic data. Additionally, the importance of leaf indumentum as a diagnostic character in *Genipa* was investigated and scanning electron micrograph images of leaf trichomes are presented.

Keywords: Angiosperms 353, SEM, SECAPR, STACEY, Multi-species Coalescent, Maximum Likelihood, Bayesian inference

Introduction

Genipa L. is a widespread Neotropical tree genus in the coffee family Rubiaceae, tribe Gardenieae DC. The most well-known species is *Genipa americana* L. This species is of economic and cultural significance. It has many uses, for example, the fruit is eaten or made into beverages and it is used as a natural blue food colourant. It is important to several indigenous groups who extract an ink from the unripe fruit which is used as body paint (Steyermark, 1972). This practice has been commercialised and today it is marketed as a henna alternative – so called jagua tattoos. It is also important for its medicinal uses and its timber. The tree is cultivated around Amazonian villages (Milliken et al., 1992), and it has also been proposed as a potential shade tolerant tree crop by the United Nations Conference on Trade and Development (Profound, 2005). The large number of common names demonstrates the ethnobotanic value of Genipa. Despite its seeming ubiquity and importance, the systematics of this genus is not well resolved. Previous classifications are based on morphological data, phylogenetic studies have been restricted to one or two loci and only for *G. americana.* This study investigates species delimitation in the genus and infers the phylogeny using an integrative taxonomy approach that combines phylogenomic and traditional taxonomic data. The necessity of a stable taxonomic framework is magnified given recent increases in deforestation in lowland tropical habitats where *Genipa* is distributed and the implications for biodiversity conservation when species concepts are ill-formed (Cavers et al, 2013, Frankham et al., 2012, Mace 2004; Ruhsam et al 2016).

Taxonomic history

The original concept of *Genipa* was not well defined and may explain the high level of synonymy in the genus. Early descriptions of the genus are scant and lack specific morphological detail. Linnaeus described the genus in Genera Plantarum (1754) and the species, *G. Americana,* in the tenth edition of Systemae Naturae (1759). As with most of his tropical plant descriptions he was not especially familiar with the genus, and his description was based on illustrations and the work of Tournefort (1700). He cited two illustrations, one by Plumier of a specimen from either Haiti or Martinique dated between 1687 – 1689 (published posthumously by Burman, 1757); the other, is a drawing by Marcgrave of a specimen from Brazil dated 1648. The latter was selected by Howard, in the Flora of the Lesser Antilles, Leeward and Windward Islands (1989) as the lectotype of *G. americana*. Both illustrations show inaccuracies. The illustration by Plumier has flowers that differ from *Genipa* most notably in the morphology of the stamens. In the drawing by Marcgrave the leaves are alternate whereas *Genipa* has opposite leaves and the leaf venation also differs from *Genipa* in that it does not show brochidodromous venation, were the secondary veins link together in loops at the leaf margin. Despite these inadequacies the latter specimen was selected as the type by Steyermark (1972), as it was preferred to have an illustration representing a specimen from Brazil. Steyermark (1972) noted that Urban (1920) describes the type location as Haiti or Martinique after the illustration by Plumier. The sub-optimal type specimen is likely to have contributed to the taxonomic confusion surrounding this genus. *Genipa* has been through several taxonomic expansions and contractions over the years. According to the International Plant Names Index (IPNI) 76 specific names exist in the genus and it has a further five infraspecific names. Previous circumscriptions of *Genipa* for example by Baillon (1880) and Drake (1898) were much larger and encompassed a pantropical distribution. Over the last twenty years *Genipa* was shown to be paraphyletic and has been gradually modified (for example, Persson, 1996; Persson, 2000a, 2000b, 2003; Rakotonas and Davis, 2006) and it is now a much reduced solely Neotropical genus. Previous *Genipa* species have been found to be congeneric with a diversity of Rubiaceae genera including: *Agouticarpa* C.H.Perss.*, Aidia* Lour.*, Alibertia* A.Rich. ex DC., *Benkara* Adans., *Bertiera* Aubl., *Burchellia* R. Br., *Casasia* A.Rich., *Catunaregam* Wolf, *Ceriscoides* (Hook.f.) Tirveng., *Duroia* L.f., *Gardenia* J.Ellis, *Glossostipula* Lorence, *Hyperacanthus* E.Mey. ex Bridson, *Randia* Houst. ex L., *Rosenbergiodendron* Fagerl., *Rothmannia* Thunb., *Sphinctanthus* Benth. and *Tocoyena* Aubl.

Recent systematic work and current taxonomic status

Existing treatments and floras of Central and South American countries (Bernal et al., 2019; Burger, & Taylor, 1993; Delprete & Cortes, 2012; Gomes, M. 2020; Mendoza et al., 2004; Steyermark & Persson, 2004; Woodson et al., 1980 and Zappi et al., 1995) recognise a different number of species and infraspecific taxa without consensus, summarised in Table 1. Kew's The World Checklist of Vascular Plants (2022) and International Plant Names Index ((IPNI), 2022) list three valid species: *Genipa americana*, *G. infundibuliformis* Zappi & Semir and *G. spruceana* Steyerm. *G. spruceana* was first described by Steyermark in The Botany of Guyana Highlands (1972), which also contains a detailed description of *G. americana*. The most recently described species is *G. infundibuliformis* by Zappi et al.,1995. The Missouri Botanical Garden database, Tropicos.org (2022) accepts three species, *G. americana, G. chapelieri (A. Rich.) Drake* and *G. infundibuliformis.* In Tropicos.org *G. spruceana* is treated as a synonym of *G. americana.* The other major global botanical taxonomic databases, the Leipzig Catalogue of Vascular Plants (Freiberg et al., 2020) and World Flora Online (WFO, 2022) list four species in *Genipa: G. americana, G. infundibuliformis, G. spruceana* and *G. chapelieri. Genipa chapelieri* is a somewhat puzzling Madagascan species (Bridson and Robbrecht, 1985). It is synonymous with *G. talangninia* (DC.) Drake, now moved to *Hyperacanthus talangninia (*DC.*)* Rakotonas. & A.P. Davis in the Aidia clade (sensu Mouly et al., 2014) and is excluded from this study. Zappi et al., (1995). The online Flora do Brasil (2020) also treat *G. spruceana* as conspecific with *G. americana*, whereas other treatments recognise it as a separate species (Bernal et al., 2019, Mendoza et al., 2004, and Steyermark and Persson, 2004). The entry in the Checklist of the Plants of the Guiana Shield (Funk et al., 2007) is *G. spruceana = G. americana?* indicating that it is a species of unknown certainty.

As the different botanical works summarised in Table 1 indicate there is disagreement within the botanical community. Some view *G. americana* as a single highly phenotypically variable species (for example, Burger and Taylor, 1993; Gomes 2020; Zappi et al., 1995) while others view the phenotypic variation to be of taxonomic merit. For example, in Pittier's 'Century of Trees of Panama' (1931), he has a detailed description of *G. caruto* Kunth, the hairy-leaved genipa which he recognises as a distinct species from *G. americana* the smooth-leaved genipa (Pittier, 1931). The former view, is the more commonly adopted approach among botanists today. *G. caruto* is now demoted to *G. americana* var. *caruto* Kunth (K. Schum) or not recognized at all. Zappi et al., (1995) state that "the indumentum of *G. americana* are quite variable" and as the character is not discontinuous it should not be used to determine taxa. However indumentum can be important diagnostic characters for species determination in plants (Payne, 1978), and they have been treated as taxonomically informative in *Genipa* and feature in all keys to the genus (Bernal et al., 2019; Berry et al., 2004; Mendoza et al., 2004; Steyermark & Persson, 2004; Woodson et al., 1980 and Zappi et al., 1995).

This tendency of lumping is applied to the infraspecific taxa in *Genipa*. IPNI lists five infraspecific names *G. americana* var. *caruto, G. americana* f. *grandifolia* Chodat & Hassl., *G. americana* f. *jorgensenii* Steyerm., *G. americana f. parvifolia Chodat & Hassl and G. americana var. riobranquensis Kuhlm.* Many of the botanical works listed in Table 1 do not recognize these infraspecific taxa (Burger and Taylor, 1993; Gomes, 2020; Woodson et al., 1980; Zappi et al., 1995).

Table 1. Summary of *Genipa* species recognised in different works

Work	*G. americana*	*G. americana var. caruto*	*G. caruto*	*G. infundibuliformis*	*G. spruceana*
PoWO	✔			✔	✔
Tropicos	✔			✔	
WFO	✔			✔	✔
LCVP	✔			✔	✔
Kew Bul Zappi	✔			✔	
Flora Panama	✔			Na	Na
Flora Guatemala			✔	Na	Na
Guyana Highlands	✔	✔		Na	✔
Costaricensis	✔			Na	Na
Venezuelan Guyana	✔	✔		Na	✔
Bolivia	✔	✔		Na	
Central French Guiana				Na	✔
Rubiaceae de Colombia	✔	✔		Na	✔
Plants & lichens of Colombia	✔			Na	✔
Mato Grosso	✔			Na	✔
Online Flora do Brasil	✔			✔	

Na denotes that it is outside the known distribution of the species.

PoWO: Plants of the World Online, Royal Botanic Gardens, Kew; Tropicos:Tropicos.org, Missouri Botanical Garden; WFO: World Flora Online; LCVP: Leipzig Catalogue of Vascular Plants; Kew Bul Zappi: Kew Bulletin, 50(4), 761–771; Flora Panama: Flora of Panama. Part IX. Family 179. Rubiaceae--Part 1. Annals of the Missouri Botanical Garden, 67(1), 1–256; Flora Guatemala: Flora of Guatemala (Steyermark, 1950); Guyana Highlands: The Botany of the Guyana Highlands; Costaricensis: Flora Costaricensis Family #202 Rubiaceae. Fieldiana, 33; Venezuelan Guyana: Flora of the Venezuelan Guayana: Poaceae – Rubiaceae, Berry, P. E., & Missouri Botanical Garden (Eds.); Bolivia: Guia de Arboles de Bolivia; Central French Guiana: Guide to the Vascular Plants of Central French Guiana; Rubiaceae de Colombia: Rubiaceae de Colombia. Guia ilustrada de generos; Plants & lichens of Colombia: Catalogue of the Plants and Lichens of Colombia; Mato Grosso: A synopsis of the Rubiaceae of the states of Mato Grosso and Mato Grosso do Sul; Online Flora do Brasil:.http://floradobrasil.jbrj.gov.br/reflora/floradobrasil

An integrated taxonomic approach

A large body of literature exists covering the species concept debate, the prevailing popular species concepts such as the biological species concept (Mayr, 2000) are best applied to sexually reproductive species were allopatric speciation predominates. For botanists, a different species concept is required, one that accommodates hybridisation, clonal reproduction, apomixis and polyploidy. One widely applicable theory is 'the unified species concept', which is based on separately evolving metapopulation lineages (De Queiroz, 2007). It can easily be applied as it is a catch-all that separates species conceptualization – the ontological theory, from species delimitation – the epistemological practice. This species concept works well with genomic data and phylogenies can readily be produced that reveal distinct evolutionary lineages. We have in a sense solved the species concept problem, by shifting it to a species delimitation problem.

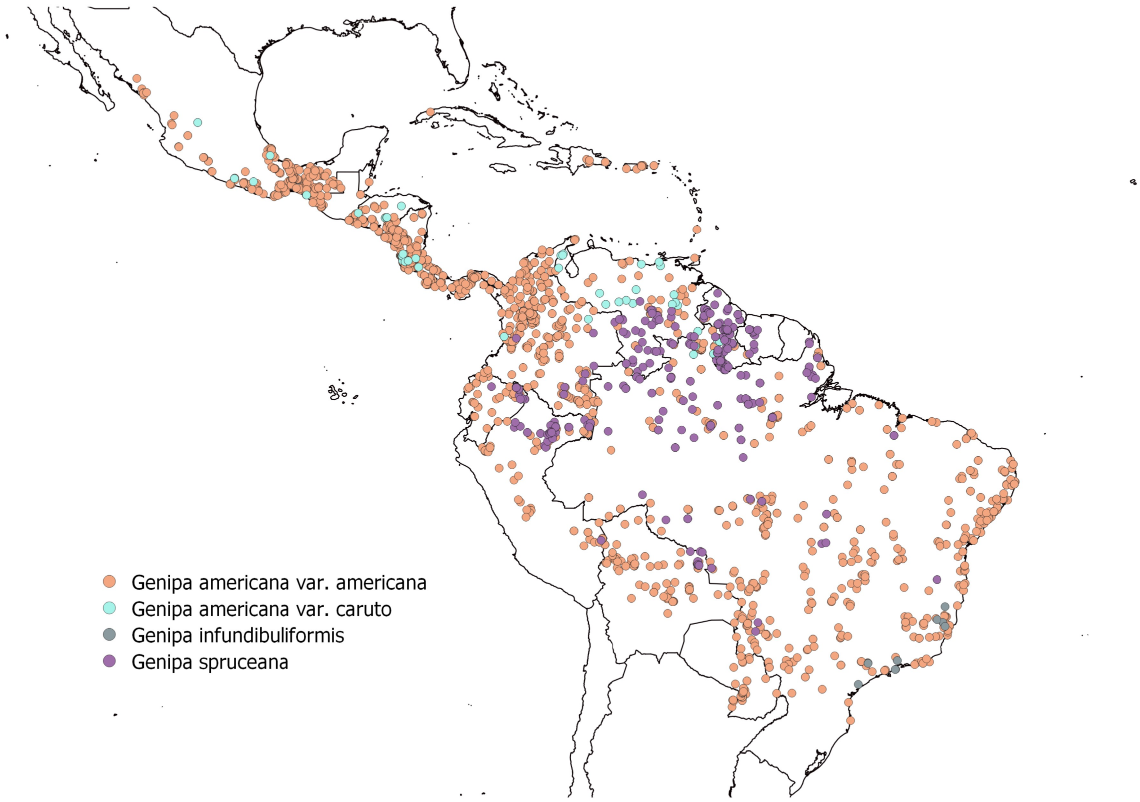
If you consider that there are four aspects to taxonomy: delineation, classification, identification and naming (Dayrat, 2005), using sequence data and increasingly available genomic data to delimit taxa alone is problematic as it precludes identification on a practical level as molecular facilities are unavailable to many taxonomic users. If a taxa cannot be identified then its utility is limited. The use of genomic data as the only identifying feature for angiosperms is therefore not a desirable model and should only be considered for truly cryptic taxa, as defined by Struck et al., (2018). Furthermore, current phylogenomic methods do not readily distinguish between population structure and species (Carstens et al., 2013; Sukumaran and Knowles, 2017). This can result in taxonomic inflation whereby previously identified infraspecific taxa or new clades are erroneously recognised as new species (Issac et al., 2004; Sukumaran & Knowles, 2017). In order to avoid conflating species limits with population structure, I use multiple lines of evidence, in an integrated approach (Carstens, 2013; Denham et al., 2018; Fernández et al., 2017; Karbstein, 2020; Wortley and Scotland, 2006). Independently evolving lineages will be identified with genomic data using two different phylogeny inference methods: i) a heuristic two-step approach were gene trees are created first independently and then combined to create a species tree and ii) using Bayesian inference were gene trees and the species tree are co-estimated. Datasets with different numbers of loci will also be tested. This data will be cross-referenced with morphological and distribution data to determine species boundaries in the genus. This approach will decrease the likelihood of conflating population structure with species boundaries (Carstens et al., 2013; Sukumaran and Knowles, 2017). Morphological data will serve not only as a proxy for gene flow between populations but also to ensure that the species limits are diagnostic. The lack of statistical rigour in morphological studies of plants is problematic due to the likelihood of introducing subjectivity or bias into species delimitation. In order to avoid this I shall employ a number of analytic statistical bioinformatic tools to ensure morphological species determinations are demonstrably data driven.

Target sequence capture

Target sequence capture is a genomic method that balances cost, data scale and computational requirements (Jones and Good, 2016). Using massive parallel sequencing technology, specific loci are sequenced on a large scale of hundreds or sometimes thousands of loci. This method is widely used for phylogenetic inference. It is suitable for DNA of limited quality that is more fragmented such as herbarium specimens or degraded silica dried plant material. A major benefit of target sequence capture is the existence of pre-designed bait or probe kits that target known regions of the genome. One such kit is the Angiosperms 353 bait kit which targets 353 single-copy protein-coding genes and works across all angiosperms (Johnson et al., 2019). Target sequence capture data will be used to produce a phylogeny of the genus with heuristic and Bayesian inference methods to determine consistency in species delimitation. The multi-species coalescent model (Rannala and Yang, 2003; Degnan and Rosenberg 2009; Edwards 2009) is applied here for phylogeny construction. This applies mathematical and probabilistic theory to explain the evolution of alleles and accounts for the incongruence between gene trees and species trees as a result of incomplete lineage sorting. Incomplete lineage sorting is a genetic phenomenon through which two alleles of the same gene fail to coalesce within the species.

Summary of geographic distribution

Genipa is widely distributed from Mexico and the Caribbean to Argentina (Figure 1). *Genipa americana* var. *americana* has the widest distribution and is present in 31 countries. *Genipa americana* var. *caruto* has a northerly distribution, it is restricted to Central America, Colombia and Venezuela. *Genipa spruceana* which is sometimes sunk within *G. americana,* has a distribution that overlaps with *G. americana, G. spruceana* is the most common species in the Guiana shield region. *Genipa infundibuliformis* has the most narrow distribution; it is only found in the Atlantic Forest on the south east coast of Brazil. In parts of the Guiana Shield: Venezuela (Amazonas, Bolivar), Guyana and Surinam, *G. americana* var. *americana, G. americana* var. *caruto* and *G. spruceana* are sympatric. Given the differences in the number of species recognised in *Genipa* (Table 1) the distribution map for the genus in Figure 1 reflects an approximate distribution as we do not know how the determination of each record was reached. Given the known differences in taxonomic classification in the genus it is likely that *G. spruceana, G. americana* var. *caruto* and *G. infundibuliformis* are under-recorded and have been recorded as *G. americana*.



Genipa is found in a variety of tropical and subtropical habitats Cerrado Savanna (sensu lato), Gallery Forest, (Zappi et al., 1995) Seasonally Inundated Forest (Igapó), Terra Firme Forest, Inundated Forest (Várzea), Seasonally Deciduous Forest, Seasonal Evergreen Forest, Seasonally Semi-deciduous Forest, Ombrophyllous Forest (Tropical Rain Forest), Coastal Forest (Restinga) and anthropic areas (Gomes, 2020). Additionally Pittier (1931) states that it is found in deciduous tree clumps in savanna. It is found from sea level to 900 m (Burger and Taylor, 1993). The ecology of *G. spruceana* differs from *G. americana* as it is a riparian species (pers. com. Claes Persson). *G. infundibuliformis* is known from humid Atlantic Forest in South East Brazil (Zappi et al., 1995). Many of these habitats are vulnerable to deforestation primarily through conversion to agricultural land, especially the Atlantic Forest were around 85% of the original area has been deforested (Ribeiro et al., 2009). The IUCN threat status has not been calculated for *Genipa*. The future status has been projected to be endangered (*G. americana*) and vulnerable (*G. spruceana*) (Steege et al., 2015). However it is not possible to accurately calculate conservation status until the infra-generic taxonomy is stable.

Aim

Here I implement an integrated taxonomy approach leveraging phylogenomics, morphology, and distribution data. This will enable species delimitation in the genus and specifically test if:

1. *G. spruceana* is a separate species to *G. americana;*
2. *G. caruto* is a separate species to *G. americana;*
3. leaf trichomes are diagnostic in the genus.

I test these hypotheses using target sequence capture data to infer relationships within the genus and create a phylogeny based on the multi-species coalescent model. Finally, species delimitations will be made by comparing independently evolving lineages elucidated from phylogenomic data to morphological and distribution data.

Material and methods

Taxon Sampling

Twenty-eight *Genipa* samples representing all putative species in the genus were available for the phylogenomic analyses. Sample numbers for the phylogenomic analyses are as follows: 12 *G. americana var. americana,* 7 *G. americana var. caruto*; 2 *G. infundibuliformis* and 7 *G. spruceana.* Taxon sampling for the morphological analyses also covered all study taxa, specimen numbers were as follows: 102 *Genipa americana* var. *americana*; 57 *Genipa americana* var. *caruto*; 1 *Genipa infundibuliformis* and 62 *Genipa spruceana*. A total of 13 type specimens were included in the morphological study (two holotypes of *G. americana* var. *caruto* f. jorgensenii and *G. spruceana* var. *ramosa*; four isotypes of *G. barbata, G. americana* var. *caruto* f. *grandifolia, G. americana var. caruto f. parvifolia* and *G. spruceana* var. *ramosa*; two holotypes or isotypes designation not determined of *G. codonocalyx* and *G. spruceana*; four paratypes and one syntype of *G. spruceana*). The geographic coverage of specimens was representative of the distribution of the genus and samples from nine countries were in the genomic analysis. The list of specimens examined is given in Appendix 1 and a key to the genus is given in Appendix 2.

Phylogenomic analyses

Leaf tissue samples collected in the field and dried in silica gel were homogenised using a Tissuelyser II (Qiagen, Venlo, Netherlands). Total genomic DNA was extracted using the NucleoSpin Plant II Kit (Macherey-Nagel, Düren, Germany) or DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The protocol followed manufacturers instructions apart from the cell lysis time which was increased to overnight to maximise DNA yield. DNA quality was assessed using a Nanodrop 2000 spectrophotometer and quantified using the Qubit 2.0. The Nanodrop 2000 and Qubit 2.0 results were used to determine samples that needed concentration by vacuum centrifugation. Gel electrophoresis was also carried out to assess

DNA fragment size. Multiple extraction rounds were pooled as necessary when initial DNA quantity was low, in order to meet the minimum concentration requirements 8ng/µl. DNA samples were sent to Rapid Genomics, Florida, USA for target capture library preparation and sequencing. The DNA was mechanically sheared to a size of 200 – 500 base pairs. Illumina libraries were constructed and barcode adapters for the Illumina Sequencing platform were ligated to the libraries then PCR-amplified using standard cycling protocols. Samples were pooled into 16 barcoded libraries with equimolar amounts to a total of 500 ng for hybridization. Target enrichment was performed using the “Angiosperms 353” probe set (Johnson et al., 2019) targeting 353 orthologous genes. After enrichment, samples were re-amplified for an additional 6–12 PCR cycles and sequenced using an Illumina NovaSeq 6000 with paired-end 250 bp reads.

Bioinformatic processing of target sequence capture data

The target sequence capture raw read data was processed using the bioinformatic pipeline SECAPR 2.2.5 (Andermann et al., 2018). The bioinformatic pipeline was run on the Sigma2 High-Performance Computing cluster at NTNU, Norway. Raw sequence data was quality checked using FastQC (Andrews, 2010) and MultiQC (Ewels et al., 2016) to gain an overview of sequence quality and determine cleaning parameters. Illumina adapters were removed and cleaning of sequences was carried out using FastP 0.23 (Chen et al., 2018). FastP default settings implemented in SECAPR were: i) the read was cut if the Phred score between adapter and read was below 20; ii) maximum percent of low-quality nucleotides allowed 40%, reads were discarded if they had a higher percentage of low quality nucleotides; iii) size of sliding window for quality trimming 5 nucleotides; iv) reads below complexity threshold of 10 removed; v) trim poly repeats at end of read of length 7; vi) low complexity filtering was enabled and vii) length filtering was disabled. The quality of cleaned reads was checked, using FastQC, MultiQC and the plotting function in SECAPR.

De novo contig assembly was performed on cleaned reads using Spades 3.15.2 (Bankevich et al., 2012). Spades is based on a de Bruin graph building algorithm that searches all reads looking for overlapping sequences that it combines into contig sequences using kmer values 21, 33, 55, 77, 99 and 127. The minimum contig length set was 200, contigs under this threshold were discarded. Target loci were selected from the contigs using Blastn (Camacho et al., 2009), minimum coverage and minimum identity was 80. Loci with multiple contig matches were discarded as they may represent paralogous sequences. The *Gardenia philastrei* Pierre ex Pit., Davis, A.P. 4055 (K) sequence from the Royal Botanic Gardens Kew PAFTOL project (Baker et al., 2022) was the reference sequence for Blastn. A multiple species alignment was created from the contig data using MAFFT 7.490 (Katoh et al., 2019) with default settings in SECAPR. Referenced-based mapping was performed using a consensus sequence for each alignment from the de novo assembly step to create a reference library. The minimum coverage parameter was set at four reads. Multiple sequence alignments were performed for each locus using MAFFT 7.490 (Katoh et al., 2019) with default settings in SECAPR.

Phylogenetic analysis

Three different phylogenies were inferred. Two using ASTRAL-III (Zhang et al., 2019) with different datasets. The input data used by ASTRAL-III was maximum likelihood gene trees generated using IQ-TREE 2 (Minh et al., 2020), with ModelFinder (Kalyaanamoorthy et al., 2017) and 1000 bootstrap replicates using UFBoot2 (Hoang, 2018). The two ASTRAL-III datasets were: i) the multiple sequence alignment from the de novo contig assembly, using a dataset of 36 gene trees with four or more samples ii) the multiple sequence alignment from the reference assembly using a dataset of 245 gene trees. The trees were visualised using

Figtree v.1.4.3 (Rambaut, 2017). *Tocoyena pittieri* (Standl.) Standl. a closely related member of Rubiaceae was used as the outgroup to root the species trees.

The third species phylogeny was produced using Bayesian inference, created with Species Tree And Classification Estimation, Yarely (STACEY. Jones, 2017) for BEAST2 (Bouckaert et al., 2014) on the CIPRES Science Gateway web portal (Miller et al., 2012). This method simultaneously estimates gene trees and species trees using a birth-death collapse model. The input data was a subset of six loci from the de novo contig assembly dataset. The subset selection was the first loci in the de novo assembly dataset (5, 9, 20, 43, 55 and 62), with the exception of locus 59, it was excluded from the analysis as it only had seven out of 29 samples. The xml input was generated in BEAUTi 2.6., Java 1.8.0_212 (Bouckaert et al., 2019). The samples were not preassigned to species and no partitions were selected. The following parameters and priors were selected: species tree model collapse height: 1e-5; clock model: each locus were set as relative to each other; bdcGrowthRate: lognormal (M=5, S=2); collapseWeight:beta (alpha=2, beta=2); population prior log normal (M=-7, S=2); relativeDeathRate: beta (alpha=1, beta=1). The MCMC was run for 100 million generations and Tracer Version v1.7.1 (Rambaut et al., 2018) was used to explore convergence of parameters. The tree was generated using TreeAnnotator 2.6.3 (Drummond and Rambaut, 2007), after discarding 10% as burn-in, then visualised using Figtree v.1.4.3 (Rambaut, 2017).

Morphology analysis

The morphological analysis was undertaken at Herbarium GB. Pressed herbarium specimens were studied from the following herbaria: GB, MO, NY, U, abbreviations follow Thiers (2020 continuously updated). A pilot study of 74 features was undertaken to determine the morphological traits that may be informative in *Genipa*. Final trait selection was based on the pilot study information combined with characters deemed to be of diagnostic importance in previous studies of *Genipa* (Bernal et al., 2019; Berry et al., 2004; Mendoza et al., 2004; Steyermark & Persson, 2004; Woodson et al., 1980 and Zappi et al., 1995). The terms character and trait are used in the sense of that described by Nixon & Wheeler (1990) whereby characters are qualitative variables of which only one state is found in all comparable individuals within a species. In contrast, traits are qualitative variables of which more than one state occurs within a species. Those traits that are uninformative or difficult to measure in herbarium specimens or absent in the majority of specimens were excluded from the morphological analysis. A total of 16 traits were measured comprising seven continuous traits and nine categorical traits. The list of morphological traits and abbreviations is given in Appendix 3. To study the morphology a light microscope was used and measurements were taken with a 30 cm ruler with the exception of leaf trichomes, were scanning electron microscopy (SEM) was used to investigate trichome morphology.

SEM Study

A study of trichomes on the abaxial side of leaves was carried out using SEM at the Centre for Cellular Imaging (CCI) Sahlgrenska Academy, University of Gothenburg with the assistance of CCI electron microscopy staff. Due to the dehydrated state of the herbarium specimens, minimal sample preparation was required. Leaf samples (approximately 0.5 cm in diameter) were taken from specimens of *G. americana var. americana*, *G. americana var. caruto* and *G. spruceana* and mounted on a stub. *G. infundibuliformis* was not included in the SEM study as it is an undercollected species a sample was not available, however this does not impact the results of the study as presence or absence of trichomes are not important for species determination in *G. infundibuliformis*. The adaxial side of each leaf was adhered tightly to the carbon surface of the stub using Pelco conductive silver paint. Samples were then sputter coated with 5 nm thickness of gold using a Quorum Q150T sputter coater machine. SEM micrographs were obtained using a scanning electron microscope. A comparison of trichomes between different taxa was undertaken.

Statistical Analyses

All statistical analyses of the morphological data were undertaken using R (R Core Team, 2020). Descriptive statistics such as the mean, median and standard deviation (SD) were calculated for each continuous morphological trait. Each continuous observation was examined for distribution frequency in each taxon. K-means clustering was used to explore the patterns of variation in the data and identify morphogroups based on the continuous morphological dataset to detect if they follow putative species boundaries. The K-means clustering used the base R package and was plotted using ggplot2 (Wickham, 2016). The character states for categorical traits were coded as 0, 1 for binary traits and 0, 1, 2, 3 for multistate traits. The term NA is used for both absent traits and traits that are not applicable, however they are not interchangeable as the latter instance of NA is treated as a character state, for example, trichome colour coded as NA in glabrous plants is coded as character state 0. Univariate statistical analysis was performed on the nine categorical morphological traits and a contingency table was produced using Arsenal 3.6.3 in R to show the proportions of each character state per taxon.

Results

Phylogenomic Analyses

The mean number of raw reads for the samples was 1,126,098, the maximum was 2,183,270 and the minimum was 535,602. After cleaning the mean raw reads remaining was 1,108,523. The maximum percentage reduction after cleaning was a reduction of 4.48% and the minimum was a reduction of 0.57%. The mean number of target loci present in each contig extracted from de novo assembly was 198, 28 loci were present in all samples, 36 loci were in four or more samples. The reference assembly resulted in recovery of more loci for more samples, 245 loci were recovered, 240 contained all 29 samples (28 *Genipa* and one *Tocoyena pittieri* outgroup) and five loci had missing samples. A graphical representation of the loci recovered for each sample is shown for both types of assembly in Figure 2.

Phylogenies were inferred using ASTRAL-III which produces a species tree using quartet scores. The resulting phylogeny for the 36 loci multiple sequence alignment from the de novo assembly (Appendix 4) and the multiple sequence alignment from the reference-based phylogeny containing all 245 loci (Figure 3) show a similar topology with the exception of the placement of *G. americana* sample G_am6 from Peru. The same clades were formed in both phylogenies, *G. infundibuliformis* and *G. spruceana* are well supported as independently evolving metapopulation lineages, they received maximum local posterior probability support in ASTRAL-III. Within *G. americana* there are three subclades that are strongly supported: clade A which contains eight samples, clade B is comprised of three Bolivian and one Colombian sample and clade C is comprised of six *G. americana* var. *caruto* samples, one *G. americana* var. *caruto* from Bolivia is in clade B. In the de novo assembly phylogeny (Appendix 4) G_am6 was separate from the other *G. americana* samples and not within the three clades. In the reference based phylogeny this sample is placed within *G. americana* clade B. The Astral-III branch support levels were higher in the larger reference assembly dataset, the species and clades within *G. americana* received full support.

The phylogeny produced using STACEY also shows that *G. infundibuliformis and G. spruceana* are monophyletic and are independently evolving lineages. The same three clades are present within *G. americana* A, B and C. However, the Peruvian sample G_am6 is placed within *G. americana* clade A whereas in the ASTRAL-III tree it is in clade B. *Genipa americana* clades B and C received maximum posterior probability scores in STACEY and 0.98 for clade A. The node bars shown on the tree are the height posterior density which represents the 95% central posterior distribution of species tree split times, from this we can see that the *G. americana* clades split relatively long ago.

Morphological analysis

The results from the morphological analysis of the 16 traits are shown in full in Appendix 5. The morphological analyses shows no taxonomic differentiation was detected for vegetative traits and that fruit and indumentum are the key traits that are informative in Genipa. A summary of the statistical analysis undertaken, for seven continuous and nine categorical traits follows. Figure 5 shows the density plot for longest leaf length, it shows the degree of overlap between the taxa. A similar pattern was detected for leaf width (measured at the widest point) and leaf distance from the widest point to the leaf tip. For the other continuous traits the distribution frequency data indicate some correlation between fruit traits (FrL: fruit length; FrW: fruit width, FrDis: fruit distance from widest point to tip, FrNo: fruit number, based on number of pedicels per peduncle, actually maximum number of potential fruit) and taxon. *G. americana* var. *caruto* generally have larger fruit in small numbers and *G. spruceana* have smaller fruit in larger numbers (Figure 6 and Figure 7). These fruit traits show overlap between taxa, for example, *G. americana* var. *caruto* falls within the *G. americana* var. *americana* range. Association between sampling month and fruit characteristics was investigated. The scatter plot in Figure 8 shows that *G. spruceana* fruit (which are smaller) are mostly collected between December and April, only one *G. spruceana* fruit was collected between May and August; Jansen-Jacobs 3881 U from Guyana, this fruit is 55 mm x 50mm, almost double the median length. However, further investigation is required due to the low fruit sample numbers n= 15 for *G. spruceana*. There is a caveat in the interpretation of results from fruit size measurements, in that it is difficult to determine the fruit growth stage or it the fruit has reached maturity from herbarium specimens.

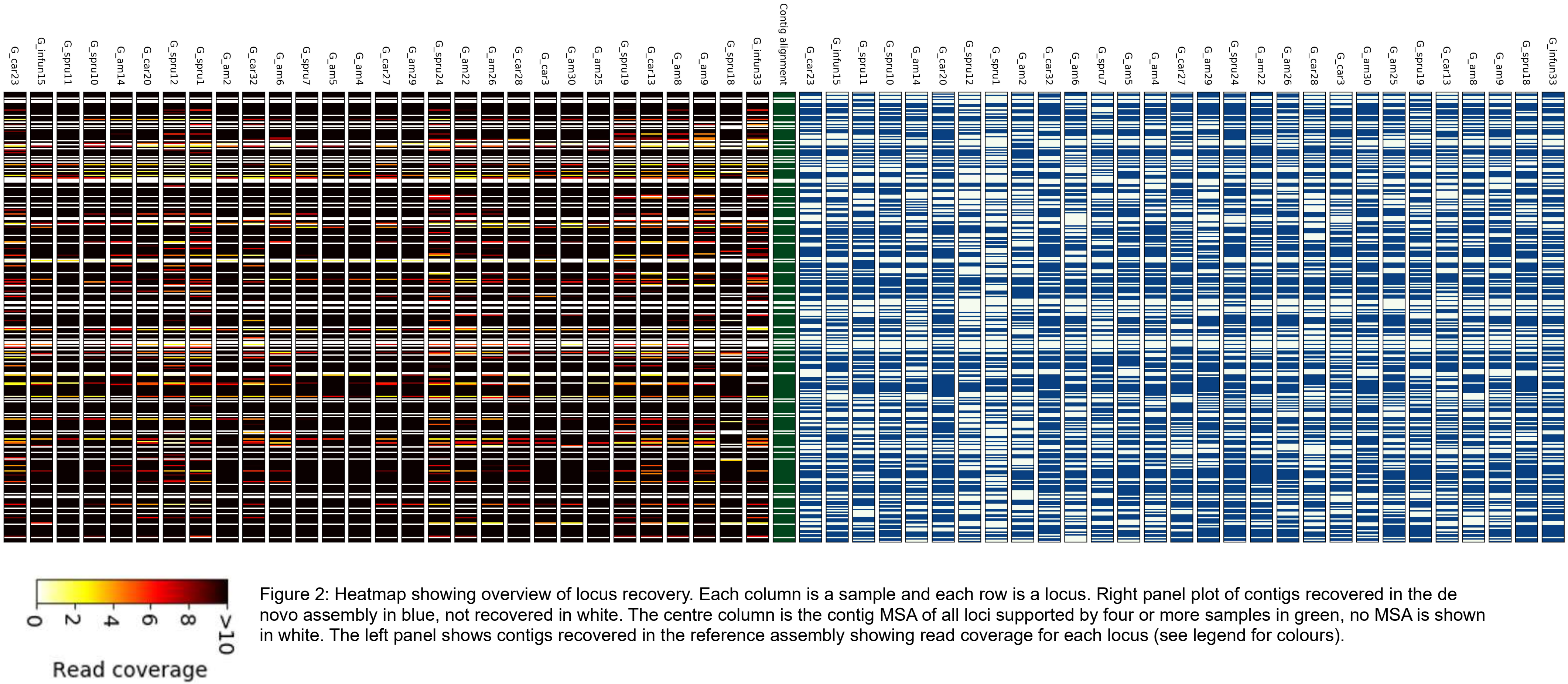


Figure 2: Heatmap showing overview of locus recovery. Each column is a sample and each row is a locus. Right panel plot of contigs recovered in the de novo assembly in blue, not recovered in white. The centre column is the contig MSA of all loci supported by four or more samples in green, no MSA is shown in white. The left panel shows contigs recovered in the reference assembly showing read coverage for each locus (see legend for colours).

15

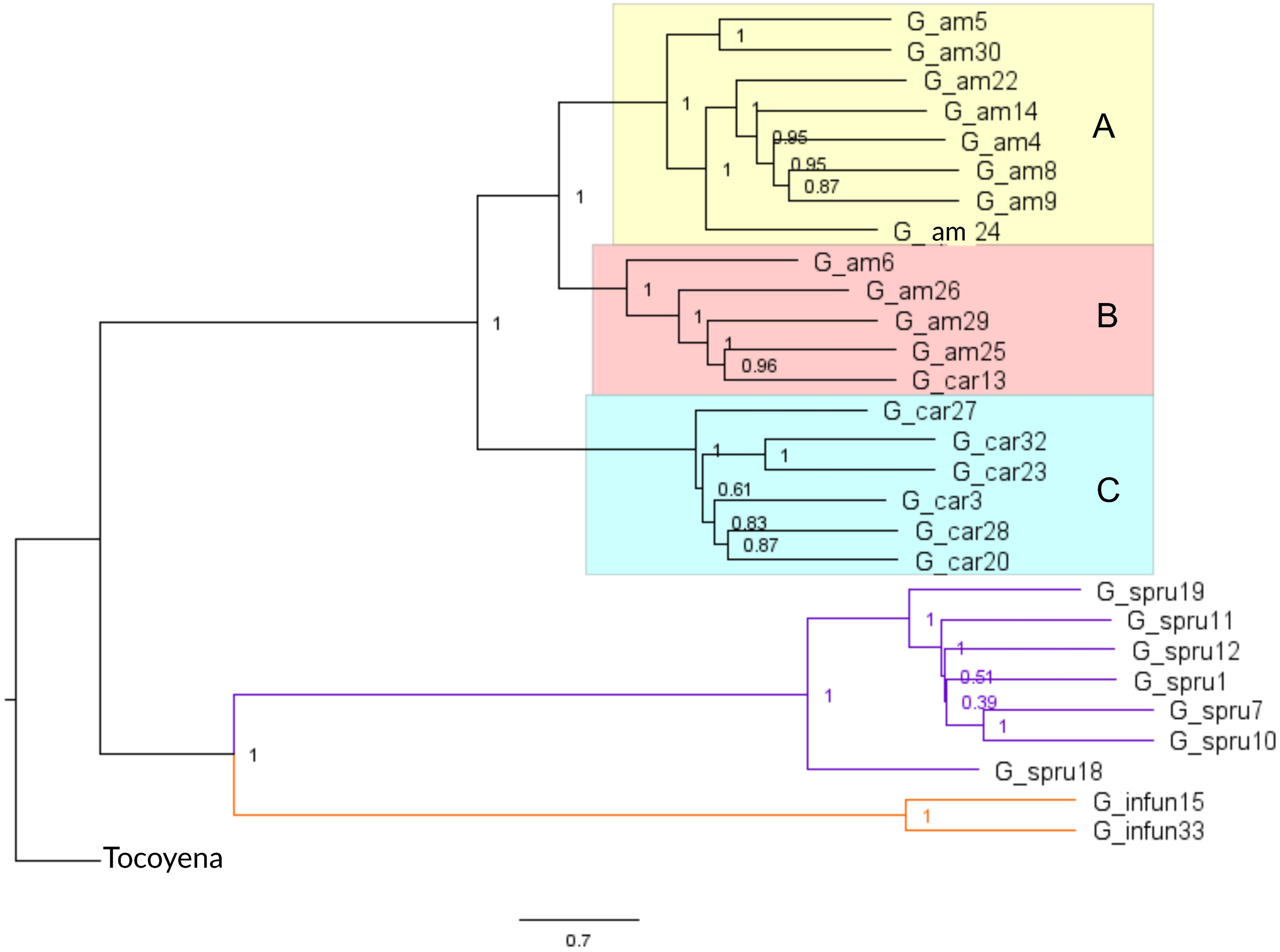


Figure 3: Cladogram produced using ASTRAL-III, of 28 *Genipa* samples, based on 245 nuclear loci, tree rooted on *Tocoyena pittieri,* with ASTRAL local posterior branch support shown. Three separate clades are shown in *G. americana*: clade A; clade B and clade C –this clade is *G. americana var. Caruto*. The scale bar is for internal branches, terminal branches have undefined branch length.

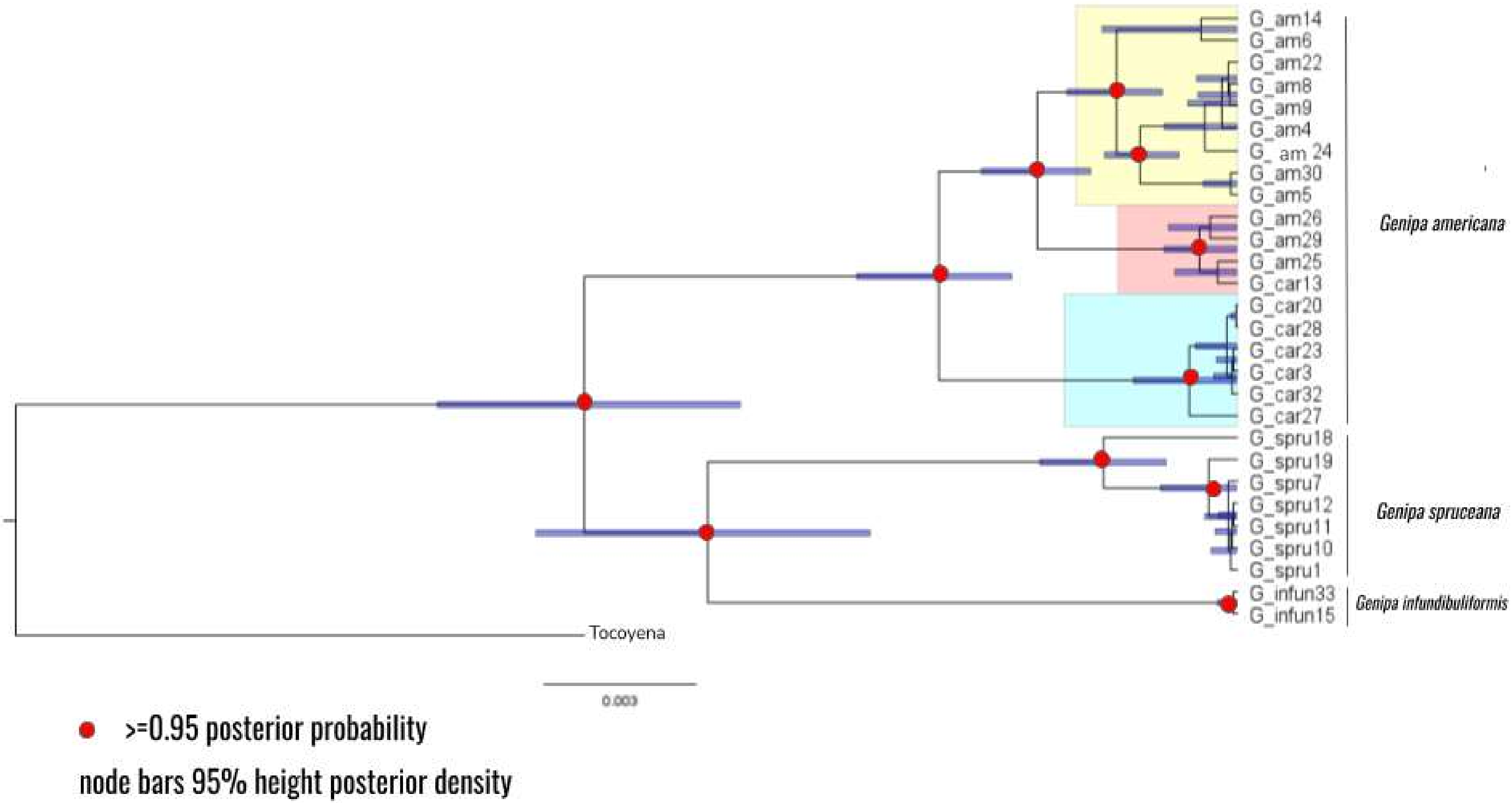
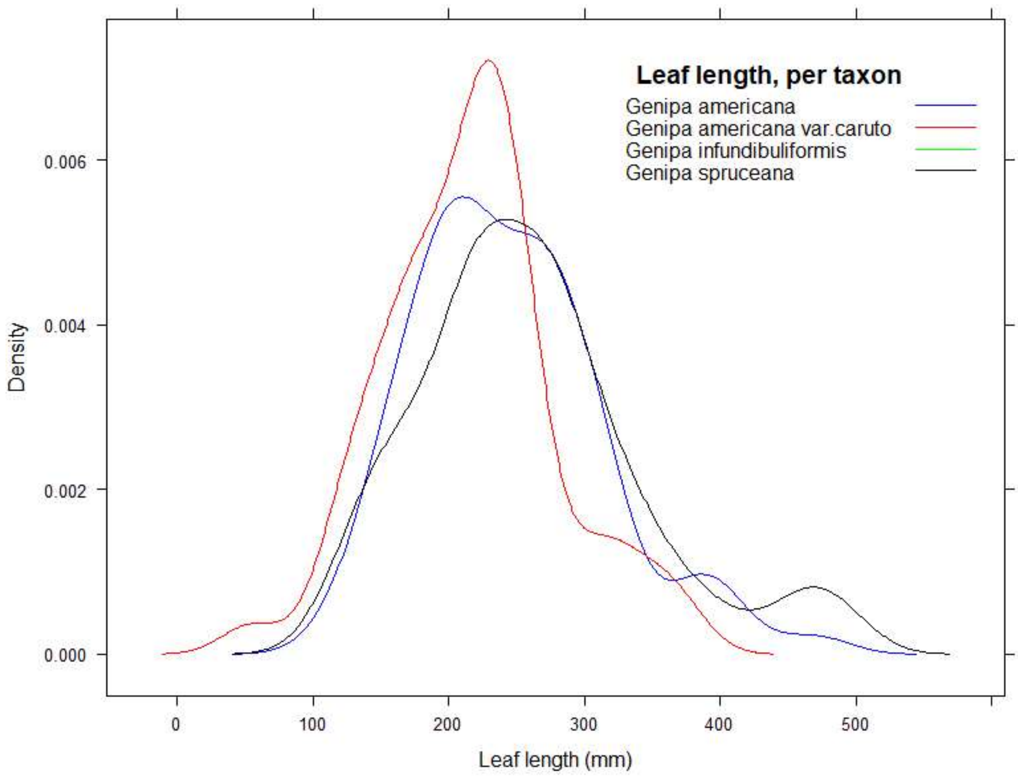


Figure 4: Phylogeny from STACEY analysis from five locus dataset, node values with red dot show posterior probabilities >0.95, node bars show 95% height posterior density.



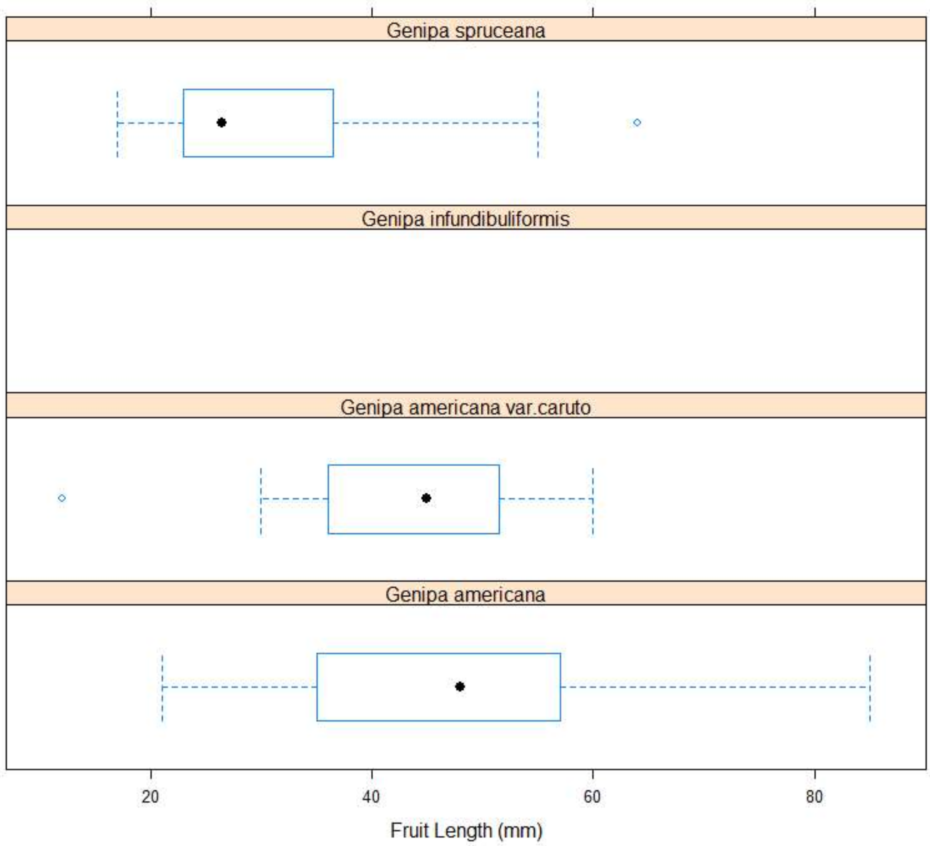
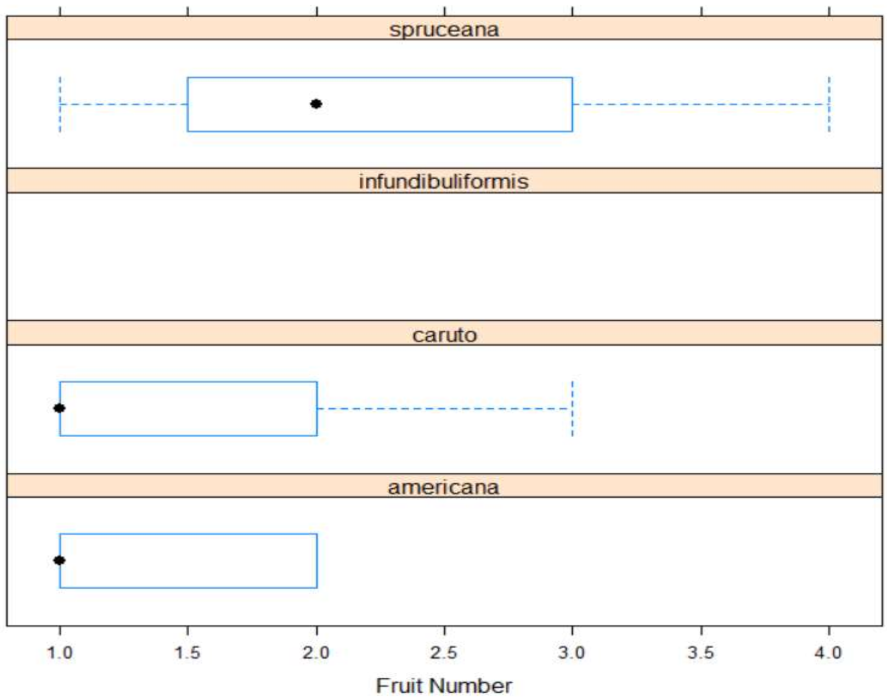


Figure 5 & Figure 6: Density plot of leaf length for individuals in each *Genipa* taxon. Box plot of fruit length for each *Genipa* taxon.



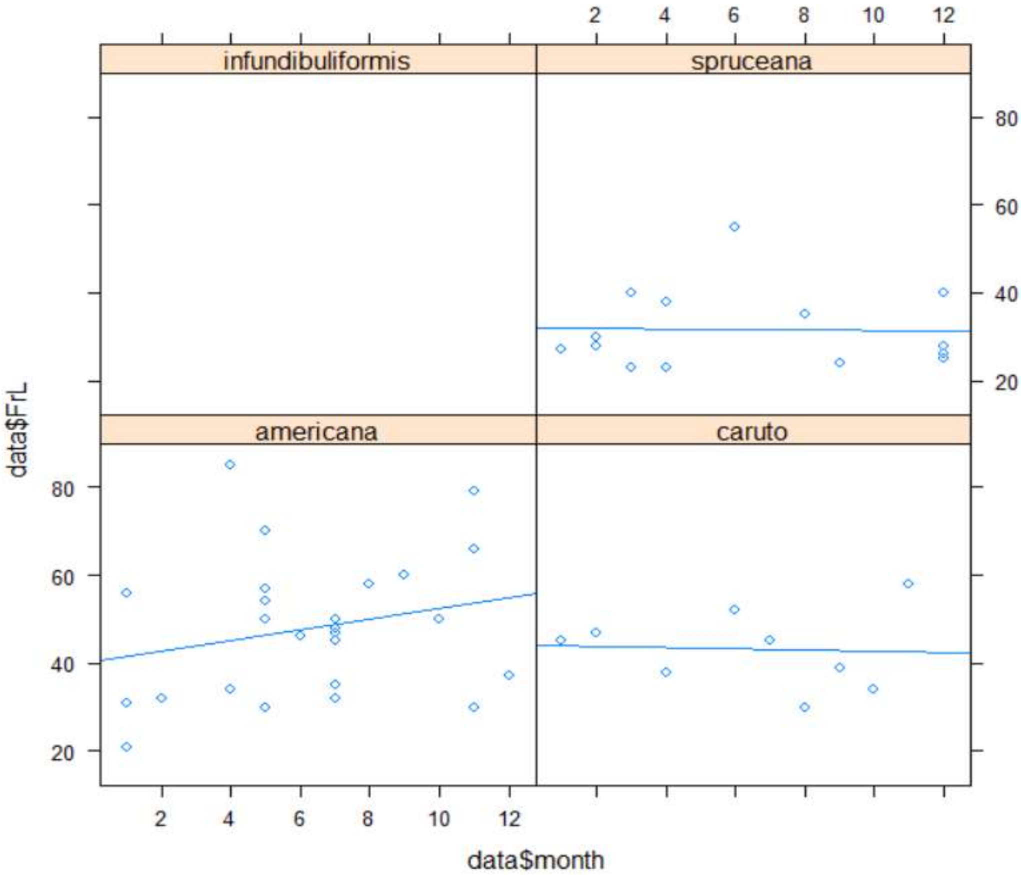


Figure 7 & Figure 8: Box plot of fruit number for individuals in each *Genipa* taxon. Scatter plot testing correlation between fruit length and month of collection.

K-means clustering was performed using the seven continuous traits (Figure 9). *Genipa* does not fruit and flower at the same time therefore for biological reasons the data matrix contains gaps or missing data. After removing any rows where values were not recorded, 44 observations remained. Data were removed rather than imputed due to the overall sparse matrix which could make imputation unreliable and bias results. Clustering was performed with a range of 3–6 clusters as determined by the scree plot and number of expected taxa. Resultant clusters were a mix of taxa, some hinting towards taxon grouping but overall inconclusive. This is likely attributed to the lack of correlation of individual traits to taxon (i.e. no association with leaf measurements and some overlap in fruit traits) and that there is no interaction between traits that is taxon specific.

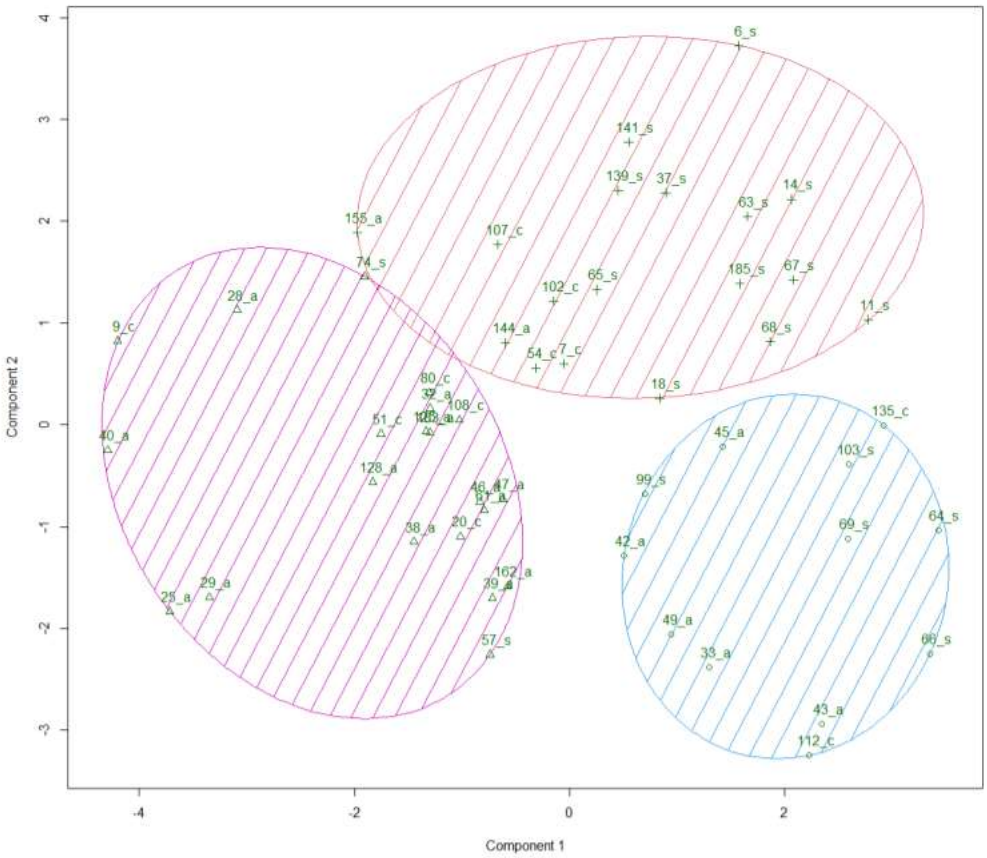


Figure 9: K-means clustering, K=3, points are labelled as follows, numbers are randomly assigned, c = *Genipa americana* var. c*aruto*, s= *Genipa spruceana,* a=*Genipa americana.*

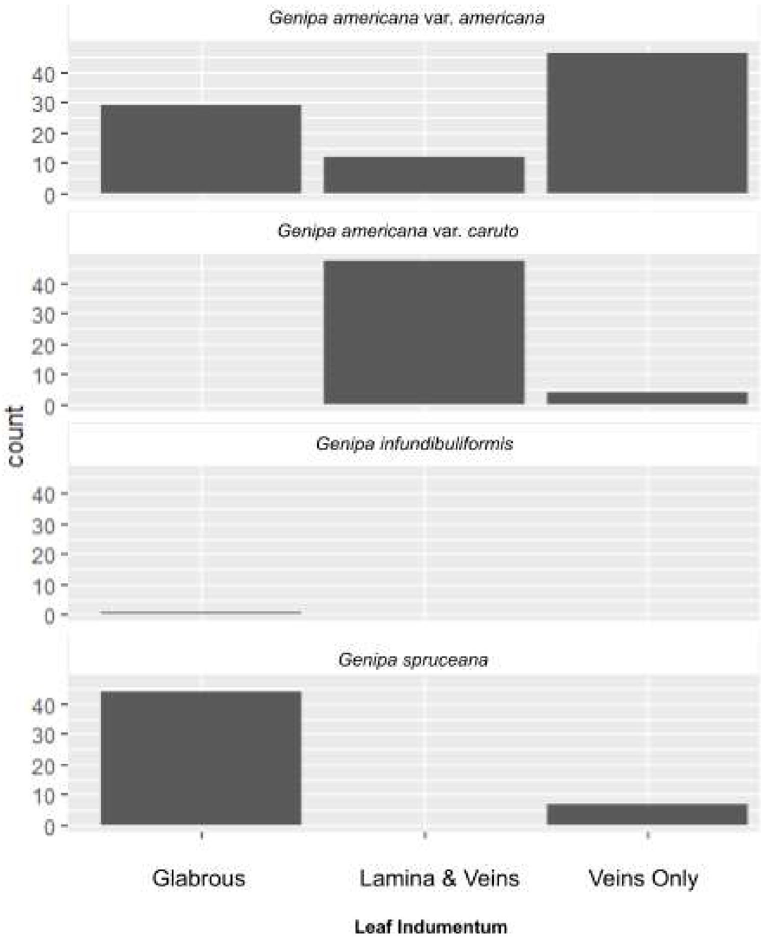


Figure 10: Leaf indumentum histogram showing the count of categories by by taxon.

A contingency table of main categorical traits showing the proportion of each character state observed per taxon is shown in Appendix 6. The categorical data show some association with taxon in the calyx interior indumentum and abaxial leaf indumentum traits (Figure 10). *Genipa spruceana* have a distinctive calyx interior, the interior calyx walls are glabrous with a small ring of minute hairs at the base of the calyx whereas other *Genipa* taxa have a sericeous or tomentose calyx interior. Abaxial leaf indumentum in *Genipa americana* var. *americana* are primarily restricted to the veins only, the primary and secondary veins have trichomes that do not cover the entire surface. *Genipa americana* var. *caruto* have a densely tomentose indumentum that covers the surface of leaf lamina and veins. *Genipa spruceana* is glabrous or sometimes nearly glabrous with sparse short indumentum on the primary vein rarely on the secondaries. Figure 10 demonstrates that there is variability in indumentum type in each taxa.

The SEM study shows that the trichomes in *Genipa* are simple, elongate, unbranched, they vary in length within a sample and vary in density between samples. Previous SEM studies have been carried out for *G. americana,* these disagree on whether *Genipa* has glandular trichomes (Vasconcelos et al., 2017) or non-gladular trichomes (Erbano and Duarte 2010). This study found no evidence of glandular trichomes in line with Erbano and Duarte (2010). The differences in indumentum between taxa are shown in the SEM micrographs Figure 11, it shows the three main abaxial leaf indumentum types found in *Genipa:* 1) glabrous leaf lamina (A) mostly short trichomes, restricted to veins (B) and detail of trichome structure (C). 2) densely tomentose trichomes, contorted or straight and longer than 1 on leaf veins (A) and lamina (B) and detail of trichome structure (C). 3) glabrous or nearly glabrous lamina (A) and veins (B). The indumentum in *G. americana* var. *americana* are variable in trichome length and density, the majority have type 1 indumentum, rarely, *G. americana* var. *americana* specimens have type 2 indumentum typical of *G. americana* var. *caruto* and sometimes they are glabrous or nearly glabrous typical of *G. spruceana* type 3 indumentum. The densely tomentose type 2 indumentum is present on all *G. americana* clade B.

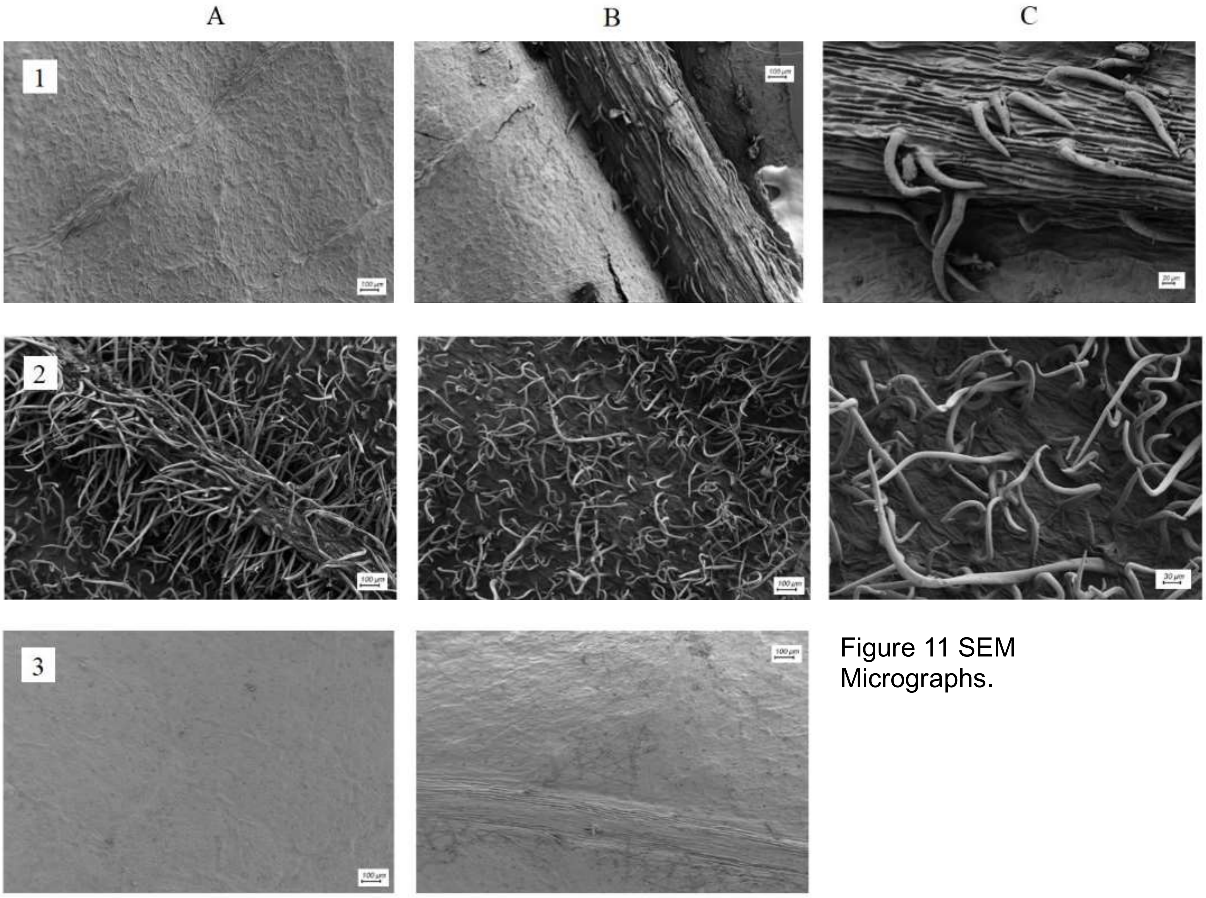


Figure 11 SEM Micrographs.

1. *Genipa americana* var. *americana* C. Bonifaz, 2535, GB
    - A) *Genipa americana* var. *americana* leaf lamina, mag. = 148 X
    - B) *Genipa americana* var. *americana* leaf vein, mag. = 148 X
    - C) *Genipa americana* var. *americana* trichome detail, mag. = 560 X
2. *Genipa americana* var. *caruto* L.G. Gomez 23043, GB
    - A) *Genipa americana* var. *caruto* leaf vein, mag. = 148 X
    - B) *Genipa americana* var. *caruto* leaf lamina, mag. = 148 X
    - C) *Genipa americana* var. *caruto* trichome detail, mag. = 437 X
3. Genipa spruceana, C. Persson,1959, GB
    - A) *Genipa spruceana* leaf lamina collector number 1959, mag. = 148 X
    - B) *Genipa spruceana* leaf vein collector number 1959, mag. = 148 X

Geographic distribution

The distribution of the molecular samples (Figure 12) and the location data for the morphological samples (Appendix 1) show that *G. americana var. caruto* has a northerly distribution in Central America and northern South America (Genipa americana Clade C on Figure 12). Samples of different taxa that were collected within 50 km of each other: G_car27 and G_spru18; G_spru11 and G_am9 clade A; G_spru11 G_am14 clade A 45591 m. Samples of different taxa that were collected within 100 km of each other: G_am30 clade A and G_am26 clade B; G_am8 clade A and G_spru1; G_am24 clade A and G_car28. This shows that the phylogeny is not the result of the sample locations, these samples could potentially interbreed but they are independently evolving lineages. The samples of *G. americana* Clade A (G_am4) and *G. americana* Clade B (G_car13) that have the typical densely tomentose indumentum of *G. americana* Clade C are shown.

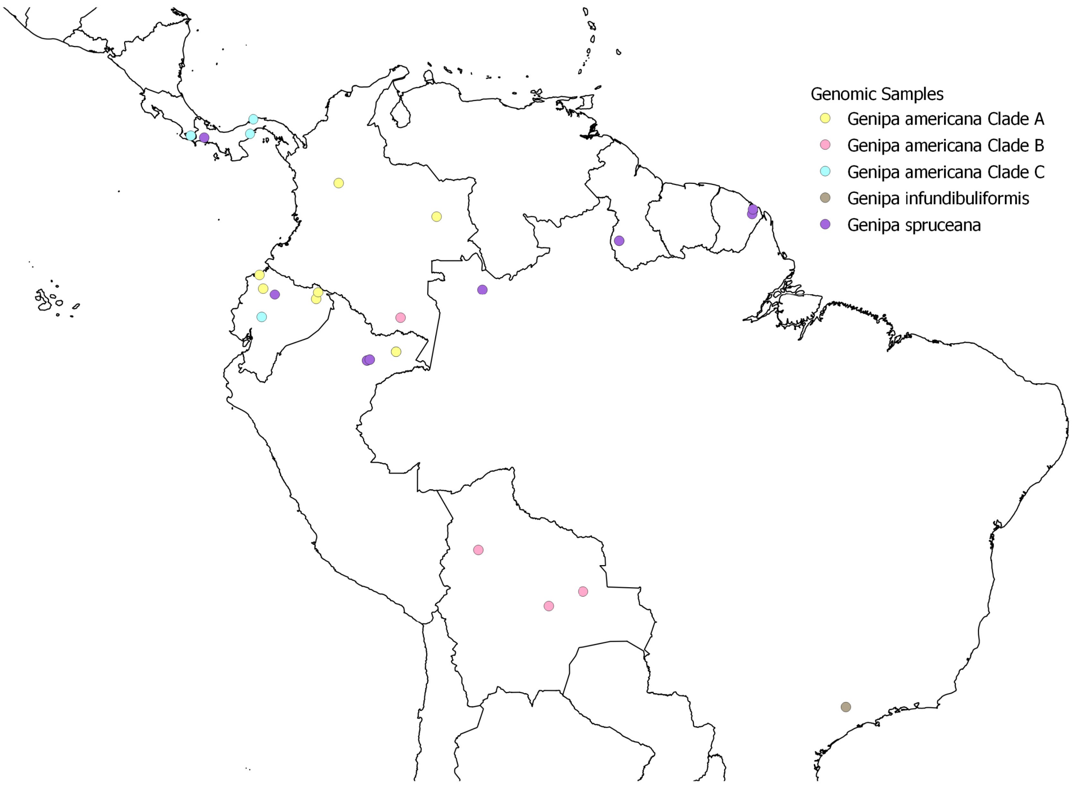


Figure 12 Location of Genomic Samples

Discussion

I produced a well resolved phylogeny from Angiosperms 353 target capture data using two coalescent methods, that are consistent in the clades returned. The data support the monophyly of *G. americana, G. infundibuliformis* and *G. spruceana.* This brings the total number of species in the genus G*enipa* to three*.* This study shows for the first time using genomic evidence, support for the recognition of *G. infundibuliformis* and *G. spruceana* as separate species from *G. americana*. *Genipa spruceana* is not universally recognised for example the recent Online Flora do Brasil (2020) treats *G. spruceana* as a synonym of *G.*

americana. This has considerable conservation implications, it falsely inflates the distribution and abundance of *G. americana* while *G. spruceana* goes unrecorded (Bickford et al., 2007; Delić et al., 2017). It echoes the conservation dilemmas presented by splitting or lumping taxa (Garnett and Christidis, 2017). The widespread species *G. americana* can be split into three genomically well supported clades, A, B and C. This pattern is understudied but it is likely to be a frequent occurrence in widespread species in the Neotropics (Antonelli et al., 2018). One of these clades, *G. americana* Clade C corresponds to *G. caruto,* today mostly recognised as a variety, commonly known as the hairy genipap. The integrative approach used genomic, morphology and distribution data as lines of evidence. As the common name suggests, all of the samples in clade C have densely tomentose abaxial leaf indumentum. However, indumentum are not diagnostically reliable in *Genipa,* one sample in clade A (G_am4, B. Stahl 5849, GB) and one sample in clade B (G_car13, C. Persson 342, GB) also have densely tomentose abaxial leaf indumentum. Indumentum are commonly considered to be a defence response or a means to control transpiration, they may reflect adaptations to humidity, wind, temperature or herbivores and pathogens (Ehleringer & Mooney, 1978; Gruner et al., 2005), the role in *Genipa* is unknown. No other distinguishing characters have been identified for *G. americana* clade C in this study. The distribution data shows clade C are only found in Central America or north of South America. The study includes good infraspecific sampling, sample density is especially high in the north of South America where *G. americana* clade A and C and *G. spruceana* are sympatric. This shows that the samples form independent clades even when there is distributional overlap, as in these contact zones potential interbreeding could occur. The other two clades in *G. americana* cannot be distinguished morphologically or distributionally. Based on the current level of study it is not accurate to describe them as cryptic species as so many lines of diagnostic evidence remain to be tested. Therefore in an attempt to increase taxonomic stability and not add to the already lengthy list of synonyms in this genus, no taxonomic changes are recommended in the genus until further evidence is acquired to decide if the clades in *G. americana* warrant species status or if an infraspecific rank is more appropriate. Described below is the support or lack of support of each line of evidence for each taxa.

Genipa infundibuliformis

Genipa infundibuliformis is fully supported as a separate species in all phylogenies. It has a discrete distribution as it grows only in Atlantic Forest habitat in south-east Brazil. It can be readily determined by its morphology, namely the long corolla tube, reflexed petal lobes and lobed juvenile leaves which are all distinct characters only found in this species of *Genipa*.

Genipa spruceana

Genipa spruceana is fully supported as a separate species in all phylogenies. Its distribution is not discrete and it is sympatric with *G. americana* var. *americana* and *G. americana var. caruto.* There is overlap in morphology between *G. americana* var. *americana* and *G. spruceana* but reliable determination of this species can be made by using a suite of morphological traits where several corroborating characters are used to ensure accurate determination in this species. Morphological traits characteristic of *G. spruceana* include glabrous or near glabrous abaxial leaf surface, glabrous calyx interior with a minute ring of hairs at the base and small fruit c. 25 mm in length in multiples of 2 or 3 per peduncle. It is also noted that there is anecdotal evidence of *G. spruceana* characters that are useful for determination that are not included in this study, these are: shiny adaxial leaf surface (this seems to deteriorate in herbarium specimens), riparian habitat and shrubby habit.

Genipa americana s.l.
There are three clades A, B and C present in *G. americana* s.l. that are well supported in the phylogenies and represent independently evolving lineages. There is a degree of correlation in distribution pattern, though geography does not perfectly correspond to each clade. *Genipa americana* var. *caruto* clade C is restricted to Central America and northern South America. Clade B contains three Bolivian samples and a Peruvian sample and a sample from the state of Amazonas, in South East Colombia. Its distribution overlaps with *G.* americana clade A. There is some taxonomic structure within the morphological data in that *Genipa americana* var. *caruto* clade C have a densely tomentose abaxial leaf surface. However this feature is not synapomorphic for this clade as one sample in clade A and one sample in clade B have densely tomentose abaxial leaf lamina. Therefore, *G. americana* var. *americana* clade A and clade B are both polymorphic in leaf indumentum. Most commonly they have trichomes only on the abaxial leaf veins, however, some individuals of *G. americana* var. *americana* have glabrous abaxial leaf surfaces. This indicates that indumentum are evolutionarily labile in *Genipa.* The three clades in *G.americana* cannot be distinguished morphologically based on the current study. Given that distributional boundaries are not diagnostic and that they lack morphological differentiation no taxonomic designation should be attributed to the clades until they can be reliably determined by means other than genomic data. This study exemplifies the importance of restraint in making taxonomic changes until multiple line of supporting evidence supports species delimitation. Referring back to the initial hypotheses *G. spruceana* is a separate species to *G. americana;* *G. caruto* is should not be recognised as a separate species to *G. americana* and leaf trichomes are not diagnostic in the genus.

Recommendations for further study
In order to establish reliable diagnostic characters corresponding to each clade in *G. americana* it is recommended that further morphological and ecological study is carried out. This may elucidate diagnostic features for these clades. While recognising that speciation is not always accompanied by morphological differentiation and that cryptic plant taxa have long been recognised, it is not possible to make a decision on whether the clades within *G. americana* represent cryptic taxa with the extent of the present morphological study. Instead, it is recommended to expand the number of morphological traits with a specific focus on traits that differ within *G. americana*. Some examples of potential traits include: leaf venation type, petiole length and indumentum density of other structures, such as corolla tube or petals. A study is recommended to determine if there are any ecological preferences or habitat characteristics that are taxa specific or if there is an ecological cause for the differences in abaxial leaf indumentum in *Genipa.* A species distribution model that includes climate, soil and hydrology data may elucidate distinctions in the *G. americana* clades. A glasshouse experiment growing *Genipa* under different hydrological regimes and measuring leaf indumentum is an additional approach that could be studied to link leaf indumentum density with hydrological conditions. It is noted that the distribution of *Genipa* is potentially the result of human cultivation and as a result it may lack strong adaptation to ecological conditions.

Once the above avenues are investigated a decision can be made on the taxonomic rank applicable (if any) to the clades within *G. americana*. If *G. americana* is not split into separate taxa then epitypification would be beneficial given the deficiencies in the existing type specimen.

Conclusion

To date, there was a lack of consensus from morphological data on how many species are in the genus *Genipa.* No previous molecular study has been carried out to test the various taxonomic viewpoints. By applying the multi-species coalescence model to detect independently evolving lineages in *Genipa* I show support for three species and evidence of infraspecific genomic structure within *G. americana* s.l. The ultimate goal is to produce a stable systematic framework for the genus *Genipa* based on an integrative taxonomy approach reliant on the corroboration of taxonomic status from independent lines of data and this study is a first step in this process.

Acknowledgements

Thank you to my supervisors Christine Bacon and Claes Persson and my examiner Bengt Oxelman for their help and support throughout the work. Many members of the University of Gothenburg Department of Biological and Environmental Sciences helped me in this work I would particularly like to thank Adrian Hill, Maria Fernández Torres, Tobias Andermann and Weston Testo for their assistance and advice. Thanks to the staff at Rapid Genomics and the Centre for Cellular Imaging Sahlgrenska Academy, University of Gothenburg, for assistance with genomic sequencing and SEM imaging, respectively. Finally I would like to commend my family especially my wife, Nina for supporting me and not being too outwardly exasperated at my fourth (and probably not final) time at University.

References

Andermann, T., Cano, Á., Zizka, A., Bacon, C., & Antonelli, A. (2018). SECAPR—a bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. *PeerJ*, **6**, e5175.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Retrieved April 19, 2022, from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Antonelli, A., Ariza, M., Albert, J., Edwards, S. V. (2018). Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ*, **6**, e5644.

Baillon H. 1880. Rubiaceae. Historie des Plantes 7: 257–501.

Baker, W. J., Bailey, P., Barber, V., Forest, F. (2022). A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology*, **71**(2), 301–319.

Bankevich, A., Nurk, S., Antipov, D., Pevzner, P. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology*, **19**, 455–77.

Bernal, R., SR Gradstein & M. Celis (eds.). (2019). Catalog of plants and lichens of Colombia. Retrieved from http://catalogoplantasdecolombia.unal.edu.co

Bickford, D., Lohman, D., Sodhi, N., Ingram, K. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148–55.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, **15**(4), e1006650.

Bridson, D., & Robbrecht, E. (1985). Validation of the African Genus Hyperacanthus E. Mey. (Rubiaceae Tribe Gardenieae). *Kew Bulletin*, **40**(2), 273.

Burger, W., & Taylor, C. W. (n.d.). Flora Costaricensis Family #202 Rubiaceae. *Fieldiana*, **33**.

Camacho, C., Coulouris, G., Avagyan, V., Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, **10**(1), 421.

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, **22**(17), 4369–4383.

Cavers, S., Telford, A., Arenal Cruz, F., Vendramin, G. G. (2013). Cryptic species and phylogeographical structure in the tree *Cedrela odorata* L. throughout the Neotropics. *Journal of Biogeography*, **40**(4), 732–746.

Chamberlain, S., Oldoni, D., Barve, V., rOpenSci (https://ropensci.org/). (2021). rgbif: Interface to the Global "Biodiversity" Information Facility API (Version 3.5.0). Retrieved from https://CRAN.R-project.org/package=rgbif

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**(17), i884–i890.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**(3), 407–417.

Delprete, P. G., & Jardim, J. G. (2012). Systematics, taxonomy and floristics of Brazilian Rubiaceae: an overview about the current status and future challenges. *Rodriguésia*, **63**(1), 101–128.

Denham, S. S., Brignone, N. F., Johnson, L. A., & Pozner, R. E. (2019). Using integrative taxonomy and multispecies coalescent models for phylogeny reconstruction and species delimitation within the "Nastanthus–Gamocarpha" clade (Calyceraceae). *Molecular Phylogenetics and Evolution*, **130**, 211–226.

Drake del Castillo, M. E. (1897). Plantes nouvelles de Madagascar. *Bull. Mens. Soc. Linn.*, **2**(1).

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

Edgar, R. C. (2021). *MUSCLE v5 enables improved estimates of phylogenetic tree confidence by ensemble bootstrapping* (preprint), Bioinformatics. Retrieved from http://biorxiv.org/lookup/doi/10.1101/2021.06.20.449169

Ehleringer, J. R., & Mooney, H. A. (1978). Leaf hairs: Effects on physiological activity and adaptive value to a desert shrub. *Oecologia*, **37**(2), 183–200.

Erbano, M., & Duarte, M. (2010). Leaf and stem morpho-anatomy of Genipa americana L., Rubiaceae. *Revista Brasileira de Farmacognosia*, **20**, 825–832.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**(19), 3047–3048.

Fernández, M., Ezcurra, C., & Calviño, C. I. (2017). Species limits and morphometric and environmental variation within the South Andean and Patagonian *Mulinum spinosum* species-group (Apiaceae-Azorelloideae). *Systematics and Biodiversity*, **15**(5), 489–505.

Frankham, R., Ballou, J. D., Dudash, M. R., Ryder, O. A. (2012). Implications of different species concepts for conserving biodiversity. *Biological Conservation*, **153**, 25–31.

Freiberg, M., Winter, M., Gentile, A., Zizka, A. N., Muellner-Riehl, Weigelt, A., & Wirth C. LCVP, The Leipzig catalogue of vascular plants, a new taxonomic reference list for all known vascular plants. Sci Data 7, 416 (2020). https://doi.org/10.1038/s41597-020-00702-z

Funk, V., Hollowell, T., Berry, P., Kellofff, C., & Alexander, S. N. (2007). Checklist of the Plants of the Guiana Shield (Venezuela: Amazonas, Bolivar, Delta Amacuro; Guyana, Surinam, French Guiana). *Smithsonian Institution: Washington, DC. Contributions from the United States National Herbarium*, **55**, 1–584.

Garnett, S. T., & Christidis, L. (2017). Taxonomy anarchy hampers conservation. *Nature*, **546**(7656), 25–27.

Gomes, M. 2020. Genipa in Flora do Brasil 2020. Jardim Botânico do Rio de Janeiro. Available at: <http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB14044>. Accessed on: 05 Jan. 2022.

Gruner, D., Taylor, A., & Forkner. (2005). The effects of foliar pubescence and nutrient enrichment on arthropod communities of Metrosideros polymorpha (Myrtaceae). Retrieved June 7, 2022, from https://drum.lib.umd.edu/handle/1903/7106

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, **35**(2), 518–522.

Isaac, N. J. B., Mallet, J., & Mace, G. M. (2004). Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution*, **19**(9), 464–469.

Johnson, M. G., Pokorny, L., Dodsworth, S., Wickett, N. J. (2019). A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. *Systematic Biology*, **68**(4), 594–606.

Jones, G. (2017). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, **74**(1–2), 447–467.

Jones, M. R., & Good, J. M. (2016). Targeted capture in evolutionary and ecological genomics. *Molecular Ecology*, **25**(1), 185–202.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**(6), 587–589.

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *TAXON*, **69**(6), 1191–1220.

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, **20**(4), 1160–1166.

Linné, C. von,. (1754). *Genera plantarum : eorumque characteres naturales secundum numerum, figuram, situm, et proportionem omnium fructificationis partium*.

Linné, C. von. (1759). *Systemae Naturae*.

Mace, G. M. (2004). The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**(1444), 711–719.

Maldonado, C., Molina, C. I., Zizka, A., Antonelli, A. (2015). Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? *Global Ecology and Biogeography*, **24**(8), 973–984.

Mayr, E. (2000). The Biological Species Concept. In Q. D. Wheeler & R. Meier (Eds.), *Species Concepts and Phylogenetic Theory: A Debate* (pp. 17–29). Columbia University Press. http://www.jstor.org/stable/10.7312/whee10142.6

Mendoza, Humberto & Ramírez Padilla, Bernardo & Jiménez, Luis. (2004). *Rubiacea de Colombia. Guía ilustrada de géneros.*

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2012). The CIPRES science gateway: enabling high-impact science for phylogenetics researchers with limited resources. In *Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the eXtreme to the campus and beyond*, Chicago, Illinois, USA: Association for Computing Machinery, pp. 1–8.

Milliken, W., R P Miller, P S Pollard. (n.d.). Ethnobotany of the Waimiri Atroari Indians of Brazil.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, **37**(5), 1530–1534.

Mouly, A., Kainulainen, K., Persson, C., Bremer, B. (2014). Phylogenetic structure and clade circumscriptions in the Gardenieae complex (Rubiaceae). *Taxon*, **63**, 801–818.

Nixon, K. C., & Wheeler, Q. D. (1990). An Amplification of the Phylogenetic Species Concept. *Cladistics*, **6**(3), 211–223.

Payne, W. W. (1978). A Glossary of Plant Hair Terminology. *Brittonia*, **30**(2), 239.

Persson, C. (1996). Phylogeny of the Gardenieae (Rubiaceae). *Botanical Journal of the Linnean Society*, **121**(2), 91–109.

Persson, C. (2000a). Phylogeny of Gardenieae (Rubiaceae) based on chloroplast DNA sequences from the rps16 intron and trnL(UAA)-F(GAA) intergenic spacer. *Nordic Journal of Botany*, **20**(3), 257–270.

Persson, C. (2000b). Phylogeny of the Neotropical Alibertia Group (Rubiaceae), with Emphasis on the Genus Alibertia, Inferred from ITS and 5S Ribosomal DNA Sequences. *American Journal of Botany*, **87**, 1018–28.

Persson, C. (2003). Agouticarpa, a New Neotropical Genus of Tribe Gardenieae (Rubiaceae). *Brittonia*, **55**(2), 176–201.

Pittier, H., & Mell, C. D. (1931). *A century of trees of Panama*.

Plumier, C., & Burman, J. (1755). *Plantarum americanarum fasciculus primus[-decimus] :continens plantas, quas olim Carolus Plumierius, botanicorum princeps detexit, eruitque, atque in insulis Antillis ipse depinxit*, Amstelaedami :Lugd. Batav.,: Sumtibus Auctoris, Prostant in Horto Medico, atqueApud Viduam & Filium S. Schouten;Apud G. Potuliet & T. Haak.

Profund. (2005). *Market Brief in the European Union for selected natural ingredients derived from native species Genipa americana Jagua, huito*, The United Nations Conference on Trade and Development (UNCTAD).

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rakotonasolo, F., & Davis, A. (2006). Six species of Madagascan *Genipa* transferred to *Hyperacanthus* (Rubiaceae–Gardenieae) and new data on general morphology, placentation and ovary structure in *Hyperacanthus*. *Taxon*, **55**(2), 387–396.

Rambaut, A. (2017). FigTree-version 1.4. 3, a graphical viewer of phylogenetic trees. *Computer Program Distributed by the Author, Website: Http://Tree. Bio. Ed. Ac. Uk/Software/Figtree*.

Rambaut A, Drummond A.J., Xie D., Baele G. and Suchard M.A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology. syy032. doi:10.1093/sysbio/syy032

Rambaut A, & Suchard M., Xie D., Drummond A. (2014). Tracer v1. 6.*, Available at URL http://Beast. Bio. Ed. Ac. Uk/Tracer.*

Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. (2009). The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, **142**(6), 1141–1153.

Ruhsam, M., Clark, A., Finger, A., Hollingsworth, P. M. (2016). Hidden in plain view: Cryptic diversity in the emblematic *Araucaria* of New Caledonia. *American Journal of Botany*, **103**(5), 888–898.

Steege, ter H., Pitman, N. C. A., Killeen, T. J., Gamarra, L. V. (2015). Estimating the global conservation status of more than 15,000 Amazonian tree species. *Science Advances*, **1**(10), e1500936.

Steyermark, J. A. (1972). *The Botany of the Guyana Highland. Memoirs of the New York Botanical Garden.*, Vol. —Part IX, New York Botanical Garden Press.

Steyermark J. A. and C. Persson. (2004). *Flora of the Venezuelan Guayana: Poaceae - Rubiaceae*, St. Louis, Mo: Missouri Bot. Garden Press.

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, **114**(7), 1607–1612.

The World Checklist of Vascular Plants (2022). World Checklist of Vascular Plants, version 2.0. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://wcvp.science.kew.org/ Retrieved 27 July 2022.

Thiers, B. (2020). *Index herbariorum*. New York, New York Botanical Garden. http://sciweb. nybg. org/science2/IndexHerbariorum. asp.(Accessed Jan, 2022).

Tournefort, J. P. (1700) *Institutiones rei herbariæ.* Typographia Regia.

Tropicos.org. Missouri Botanical Garden. 27 Jul 2022 https://tropicos.org

Urban, I. (1920). *Reportorium specierum novarum regni vegetabilis: Plumiers leben*.

Vasconcelos, A. L., Santos, A. V., Padilha, R. J. R., Alves, L. C., & Randau, K. P. (2017). Anatomical characterization of ultra-structures, biominerals and histolocalization of metabolites in leaves of Genipa americana. *Revista Brasileira de Farmacognosia*, **27**(5), 541–548.

Werker E. 2000. Trichome diversity and development. In: Hallahan D.L., Gray J.C., eds. Advances in botanical research incorporating advances in plant pathology. Vol. 31: Plant trichomes. San Diego, CA: Academic Press, 1–35.

WFO (2022): World Flora Online. Published on the Internet; http://www.worldfloraonline.org. Accessed on: 27 Jul 2022

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Retrieved from https://ggplot2.tidyverse.org

Woodson, R. E., Schery, R. W., & Dwyer, J. D. (1980). Flora of Panama. Part IX. Family 179. Rubiaceae--Part 1. *Annals of the Missouri Botanical Garden*, **67**(1), 1–256.

Wortley, A. H., & Scotland, R. W. (2006). The Effect of Combining Molecular and Morphological Data in Published Phylogenetic Analyses. *Systematic Biology*, **55**(4), 677–685.

Zappi, D. C., Semir, J., & Pierozzi, N. I. (1995). Genipa infundibuliformis sp. nov. and Notes on Genipa americana (Rubiaceae). *Kew Bulletin*, **50**(4), 761–771.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, **19**(6), 153.

Zizka, A., Silvestro, D., Andermann, T., Antonelli, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, **10**(5), 744–751.

Supporting Information

Appendix 1
List of Specimens Examined

Species	Collector name	Collector number	Herbarium code	Catalogue number	Country	State	Municipality	Lat	Long
Genipa americana	C. Persson	612	GB	NA	Peru	Loreto	Maynas	03° 48' S	73° 25' W
Genipa americana	C. Persson	1865	GB	NA	Ecuador	Pichincha	Pedro Vicente Maldonado	00° 18' 44.7" N	79° 13' 11" W
Genipa americana	C. Persson	1866	GB	NA	Ecuador	Pichincha Canton Quininde	NA	00° 18' 44.7" N	79° 13' 11.0" W
Genipa americana	C. Persson	342	GB	182548	Bolivia	Santa Cruz	Nuflo de Chavez	16° 52' S	61° 51' W
Genipa americana	J. H. E. Rova	2372	GB	187596	Panama	Colon	Portobello	09° 30' N	79° 41' W
Genipa americana	C. Persson	306a	GB	182518	Bolivia	Santa Cruz	Ichilo	17° 32' S	63° 40' W
Genipa americana	C. Persson	306b	GB	182517	Bolivia	Santa Cruz	Ichilo	17° 32' S	63° 40' W
Genipa americana	C. Persson	2143	GB	171708	Colombia	Amazonas	Leticia	04° 14' S	69° 58' W
Genipa americana	H. H. Bartlett	16465	GB	171989	Panama	Panama	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	C. Persson	231	GB	182568	Bolivia	Beni	Ballivian	14° 23' S	67° 28' W
Genipa americana	F. Alzate	225	GB	187599	Colombia	Antioquia	San Luis	05° 59' 23" N	74° 58' 33" W
Genipa americana	S. Flores	141	GB	NA	Peru	Cajamarca	San Ignacio	05° 19' 16" S	078° 41' 5" W
Genipa americana	E. Rodriguez	1695	GB	NA	Peru	NA	NA	S 23° 23' 59"	W 57° 25' 56"
Genipa americana	R. Vasquez	24304	GB	NA	Peru	Amazonas	Imaza	05° 03' 24" S	78° 20' 17" W
Genipa americana	R. Vasquez	11263	GB	NA	Peru	NA	NA	S 23° 23' 59"	W 57° 25' 56"
Genipa americana	A. Araujo-M	1633	GB	NA	Bolivia	La Paz	Abel Iturralde	14° 14' 00" S	68° 05' 00" W
Genipa americana	G.A. Parada	1336	GB	NA	Bolivia	Santa Cruz	Sara	17° 06' 41" S	63° 33' 57" W
Genipa americana	G.A. Parada	3465	GB	NA	Bolivia	Santa Cruz	Vallegrande	18° 44' 39" S	063° 37' 08" W
Genipa americana	N. Paniagua	5820	GB	NA	Bolivia	La Paz	Franz Tamayo	14° 36' 01" S	068° 41' 20" W
Genipa americana	S. Ortiz	30	GB	NA	Bolivia	Santo Cruz	Andrez Ibanez	17° 47' 58" S	63° 11' 6" W
Genipa americana	M. Nee	44922	GB	NA	Bolivia	Santo Cruz	Ichilo	17° 40' 20" S	63° 28' 35" W
Genipa americana	B.L. Stannard	529	GB	185413	Venezuela	Territorio Federal Amazonas	Negro	0° 50' N	66° 10' W
Genipa americana	A.H. Gentry	71518	GB	170956	Costa Rica	Guanacaste	NA	10° 30' N	85° 10' W
Genipa americana	L. Carrillo	152	GB	NA	Ecuador	NA	NA	N 19° 0' 0"	W 70° 40' 0"
Genipa americana	C. E. Ceron	20286	GB	NA	Ecuador	NA	NA	N 19° 0' 0"	W 70° 40' 0"
Genipa americana	P.A. Silverstone-Sopkin	5366	GB	NA	Colombia	Valle	Toro	4° 37' 37" S	76° 04' 47" W
Genipa americana	L. G. Gomez	23043	GB	10654	Costa Rica	Guanacaste	Filadelfia	10° 26' 40" N	85° 33' 5" W
Genipa americana	C. Lero	133	GB	NA	Bolivia	Beni	Ballivian y Yacuma	14° 30' S	66° 37' W
Genipa americana	E. Rwero	209	GB	NA	Bolivia	Beni	Ballivian y Yacuma	14° 30' S	66° 37' W
Genipa americana	X. Cornejo	5123	GB	NA	Ecuador	NA	NA	N 19° 0' 0"	W 70° 40' 0"
Genipa americana	C. Bonifaz	2535	GB	NA	Ecuador	NA	NA	N 19° 0' 0"	W 70° 40' 0"
Genipa americana	A. Araujo-M	1852	GB	NA	Bolivia	NA	NA	14° 19 ' 30" S	68° 33' 57" W
Genipa americana	C. Vazquez Yanes	701	GB	NA	Mexico	Veracruz	NA	N 23° 0' 0"	W 102° 0' 0"
Genipa americana	Cooper	80	NY	NA	Panama	NA	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	W.G. D'Arcy	10394	NY	NA	Nicaragua	Leon	NA	N 23° 0' 0"	W 102° 0' 0"
Genipa americana	T. B. Croat	5805	NY	NA	Panama	Barro Colorado Island	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	T. B. Croat	7949	NY	NA	Panama	Barro Colorado Island	Canal Zone	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	E. L. Ekman	8389	NY	NA	Haiti	Plaine Centrale	NA	N 19° 10' 0"	W 72° 0' 0"
Genipa americana	Pere Duss	9916	NY	NA	Martinique	NA	NA	N 15° 24' 0"	W 86° 40' 0"
Genipa americana	R. Cardene	1439	NY	Na	Bolivia	Lake Rogagua	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa americana	Herbert H. Smith	2646	NY	Na	Colombia	Santa Marta	NA	N 3° 45' 0"	W 76° 30' 0"
Genipa americana	H. Garcia-Barriga	13845	NY	Na	Colombia	Amazonas	NA	N 5° 0' 0"	W 59° 0' 0"
Genipa americana	J. Cuatrecasas	26586	NY	Na	Colombia	Cordillera Central	NA	N 3° 30' 0"	W 73° 0' 0"
Genipa americana	Earl L. Core	1649	NY	Na	Colombia	NA	Valle del Cauca	N 3° 45' 0"	W 76° 30' 0"
Genipa americana	Leopoldina de Assis	3776	NY	Na	Brasil	Sau Paulo	NA	NA	NA
Genipa americana	Emilio Goeldi	4855	NY	Na	Brasil	Para	Anajas	NA	NA
Genipa americana	B. A. Krukoff	1552	NY	Na	Brasil	Matto Grosso	Machado River Region	NA	NA
Genipa americana	J. M. Pires	11.306	NY	Na	Brasil	Para	Belem	NA	NA
Genipa americana	R. Froes	11720	NY	Na	Brasil	Maranhao	NA	NA	NA
Genipa americana	M. Claussen	683	NY	Na	Brasil	Minas Geraes	na	NA	NA
Genipa americana	N. A. Ross	2997	NY	Na	Brasil	Maranhao	NA	NA	NA
Genipa americana	F. Souza Santos	770	NY	Na	Brasil	Bahia	Santa Cruz Cabralia	16° 23' S	39° 8' W

Species	Collector name	Collector number	Herbarium code	Catalogue number	Country	State	Municipality	Lat	Long
Genipa americana	G. Hatschbach	46937	GB	Na	Brasil	Povocao	Linhares	NA	NA
Genipa americana	R. C. Gill	55	NY	Na	Ecuador	Orientes	Napo-Pastaza	S 1° 15' 0"	W 78° 15' 0"
Genipa americana	Elbert L. Little Jr.	21187	NY	Na	Ecuador	Esmeraldas	NA	S 1° 15' 0"	W 78° 15' 0"
Genipa americana	SEF	8954	NY	Na	Ecuador	Napo	Anangu	00°31' S	76° 23' W
Genipa americana	D. C. Daly	5220	NY	Na	Colombia	Antioquia	Taraza	08°35' N	75° 25' W
Genipa americana	J. Zarucchi	4960	NY	Na	Colombia	Antioquia	Arboletes	08°46' N	76° 28' W
Genipa americana	R. Callejas	2450	NY	Na	Colombia	Antioquia	de Taraza	N 5° 0' 0"	W 59° 0' 0"
Genipa americana	G. T. Prance	59509	NY	Na	Brasil	Serra do Caiapo	Moncao	NA	NA
Genipa americana	G. Hatschbach	30432	NY	Na	Brasil	Mato Grosso	Miranda	NA	NA
Genipa americana	G. Hatschbach	38926	NY	Na	Brasil	Goias	Ipameri	NA	NA
Genipa americana	McDowell T.	2552	U	68829 01.05.97	Guyana	Cuyuni-Mazaruni	Along Mazaruni River	05°53' N	60° 37' W
Genipa americana	D.B. Fanshawe	4750	U	1700 23.APR.1945	Guyana	NA	NA	N 16° 15' 0"	W 61° 35' 0"
Genipa americana	I. Boldingh	3030	U	10753	Surinam	NA	NA	N 18° 7' 44"	W 66° 28' 38"
Genipa americana	Bro. Alain H. Liogier	12405	NY	NA	Dominican Republic	Puerto Plata Province	NA	NA	NA
Genipa americana	M. Mejia	10113	NY	NA	Dominican Republic	El Seibo	NA	18° 57' N to 18° 58'	68° 46' W to 68° 47' W
Genipa americana	Dr. Alain H. Liogier	28024	NY	NA	Puerto Rico	Fajardo-Altamira	NA	NA	NA
Genipa americana	T. Zanoni	31126	NY	NA	Dominican Republic	Llano Costero	Peravia	18° 16' N	70° 19' W
Genipa americana	Elbert L. Little Jr.	13653	NY	NA	Puerto Rico	Villalba	NA	N 18° 7' 44"	W 66° 28' 38"
Genipa americana	R. A. Howard	10754	NY	NA	Grenada	NA	Isle of Ronde	NA	NA
Genipa americana	H. H. Smith	632	NY	NA	St Vincents	na	na	na	na
Genipa americana	L. E. Gregory	49	NY	NA	Puerto Rico	NA	NA	na	na
Genipa americana	J. A. Shafer	3507	NY	NA	Porto Rico	Sierra de Naguabo	NA	NA	NA
Genipa americana	Pamela Beard	1212	NY	NA	The Grenadines	NA	NA	na	na
Genipa americana	G.A. Parada	270	GB	NA	Bolivia	Santa Cruz	Vallegrande	18° 44' 39" N	63° 37' 08" W
Genipa americana	H. Pittier	12.085	NY	NA	Costa Rica	NA	NA	NA	NA
Genipa americana	Blanchet	na	NY	NA	Brasil	Bahia	NA	NA	NA
Genipa americana	Delprete, P. G.	11917	GB	NA	French Guiana	Montsinery	Riviere Tonnegrande	4° 49' 02" N	52° 32' 49" W
Genipa americana	R. Vasquez	36537	GB	NA	Peru	Pasco Oxapampa	Palcazu	10° 09' 49" S	075° 09' 52" W
Genipa americana	J. Campos	3844	GB	NA	Peru	Cajamarca	San Ignacio	05°00' 11" S	78° 58' 30" W
Genipa americana	J. Campos	4135	GB	NA	Peru	Cajamarca	San Ignacio	05° 15' 00" S	78° 45' 00" W
Genipa americana	L. Valenzuela	2266	GB	NA	Peru	Madre de Dios Tambopata	Las Piedras	12° 33' 26" S	069° 03' 00" W
Genipa americana	R. Vasquez	20925	GB	NA	Peru	Cajamarca	San Ignacio	05° 22' S	78° 30' W
Genipa americana	C. Hardy	248	GB	NA	Panama	NA	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	C. Hardy	249	GB	NA	Panama	NA	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana	B. Stergios	9040	GB	NA	Venezuela	NA	NA	N 8° 0' 0"	W 66° 0' 0"
Genipa americana	I. Boldingh	3830	U	NA	Surinam	NA	NA	N 18° 7' 44"	W 66° 28' 38"
Genipa americana	M. Mejia	6617	NY	NA	Dominican Republic	Santo Domingo	Samana	19° 16' N	69° 20' W
Genipa americana	Boon	1046	U	2376 30.MRT.1922	NA	NA	NA	NA	NA
Genipa americana	BAFOG	1111	U	3681 B 4.SEP.1957	NA	NA	NA	NA	NA
Genipa americana	BAFOG	1111	U	4B B 30. Nov. 1954	NA	NA	NA	NA	NA
Genipa americana	C. Aulestia	894	GB	NA	Ecuador	NA	NA	NA	NA
Genipa americana	M. Mejia	6617	NY	NA	Dominican Republic	Santo Domingo	Samana	19° 16' N	69° 20' W
Genipa americana	J.L. Clark	4945	GB	NA	Ecuador	Esmeraldas	Quininde	00° 20' 50" N	079° 28' 35" W
Genipa americana	E. Zardini	32718	GB	NA	Paraguay	Central	Esteros de Ypoa	25° 40' S	57° 30' W
Genipa americana	NA	37587	NA	171910	NA	NA	NA	NA	NA

Species	Collector name	Collector number	Herbarium code	Catalogue number	Country	State	Municipality	Lat	Long
Genipa americana var.caruto	D. Santamaria	S-959 B	GB	NA	Costa Rica	Puntarenas	NA	08° 59' 57"	83° 13' 58"
Genipa americana var.caruto	D. Santamaria	S-959 A	GB	NA	Costa Rica	Puntarenas	NA	08° 59' 57" N	83° 13' 58" W
Genipa americana var.caruto	B. Stčhl	5849	GB	NA	Ecuador	Los Rios	Cerro Samama near Puerta N	01° 39' S	79° 20' W
Genipa americana var.caruto	J. H. E. Rova	2402	GB	NA	Panama	Chiriqui	San Lorenzo	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	M. J. Jansen-Jacobs	3680	NY	NA	Guyana	NA	Rupununi	02° 49' N	59° 09' W
Genipa americana var.caruto	C. Persson	1976	GB	171362	French Guiana	NA	Ile de Cayenne	4° 54' N	52° 16' W
Genipa americana var.caruto	M.J. Jansen-Jacobs	4031	GB	179917	Guyana	Rupununi	Dadanawa	02° 49' N	059° 31' W
Genipa americana var.caruto	Dr. E. Hassler	12403	GB	NA	Paraguay	Central	Ypacaray	25° 24' 28" S	57° 17' 20" W
Genipa americana var.caruto	R. Rojas	24850	GB	NA	Peru	Amazonas	Bagua Province	04° 52' 00" S	78° 19' 01" W
Genipa americana var.caruto	T. Plowman	11280	GB	NA	Peru	Huanuco	Leoncio Prado Tingo Maria	9° 17' 43" S	75° 59' 44" W
Genipa americana var.caruto	Reinaldo Aguilar	1608	GB	NA	Costa Rica	Puenta Arenas	Canton Golfito	08° 27' 00" N	83°29' 30" W
Genipa americana var.caruto	J. Gonzalez	418	GB	NA	Costa Rica	Puenta Arenas	Canton de Puenta Arenas	90° 39' 51" N	85° 03' 92" W
Genipa americana var.caruto	Marko Lewis	NA	GB	171910	Bolivia	Santa Cruz	Nuflo de Chavez	14° 45' S	62° 45' W
Genipa americana var.caruto	L. Rea	344	GB	NA	Bolivia	La Paz	Abel Iturralde	14° 27' S	67° 46' W
Genipa americana var.caruto	Alfonso Jimenez M.	1007	NY	NA	Costa Rica	Guanacaste	NA	N 9° 40' 0"	W 84° 0' 0"
Genipa americana var.caruto	Paul H. Allen	3657	NY	NA	Panama	Chiriqui	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	M. Nee	10147	NY	NA	Panama	Veraguas	NA	N 9° 0' 0"	W 80° 0' 0"
Genipa americana var.caruto	T. G. Yuncker	8155	BM	NA	Honduras	Yoro	NA	N 15° 24' 0"	W 86° 40' 0"
Genipa americana var.caruto	J. Dwyer	na	BM	NA	Panama	Los Santos	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	Carlos Renson	290	NY	NA	El Salvador	San Salvador	NA	N 13° 50' 0"	W 88° 55' 0"
Genipa americana var.caruto	E. Matuda	5418	NY	NA	Mexico	Chiapas	NA	N 14° 40' 0"	W 61° 0' 0"
Genipa americana var.caruto	Angela Laguna	442	NY	NA	Nicaragua	Esteli	NA	N 23° 0' 0"	W 102° 0' 0"
Genipa americana var.caruto	D. Philcox	7488	NY	NA	Trinidad	NA	NA	N 4° 0' 0"	W 56° 0' 0"
Genipa americana var.caruto	Alfonso Jimenez M.	3211	NY	NA	Costa Rica	San Jose	NA	N 9° 40' 0"	W 84° 0' 0"
Genipa americana var.caruto	N.L. Britton	6315	NY	NA	Cuba	Pinar Del Rio	NA	N 22° 25' 0"	W 83° 50' 0"
Genipa americana var.caruto	Bro. Lemente	7374	NY	NA	Cuba	Oriente	NA	N 22° 25' 0"	W 83° 50' 0"
Genipa americana var.caruto	Antonio Molina R.	1065	BM	NA	Honduras	Modrazan	NA	N 15° 24' 0"	W 86° 40' 0"
Genipa americana var.caruto	J. N. Rovirosa	480	NY	NA	Mexico	NA	NA	N 14° 40' 0"	W 61° 0' 0"
Genipa americana var.caruto	G. Martinez-Calderon	1416	NY	NA	Mexico	Oaxaca	NA	N 23° 0' 0"	W 102° 0' 0"
Genipa americana var.caruto	S.F. Glassman	2183	NY	NA	Honduras	Morazan	NA	N 15° 24' 0"	W 86° 40' 0"
Genipa americana var.caruto	R.E. Woodson Jr.	1384	NY	NA	Panama	Panama	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	H Pittier	3952	NY	NA	Panama	Canal Zone	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	R. J. Seibert	409	NY	NA	Panama	Canal Zone	NA	N 12° 26' 16"	W 86° 52' 40"
Genipa americana var.caruto	R. J. Taylor	4419	NY	NA	Costa Rica	Guanacaste	NA	N 9° 40' 0"	W 84° 0' 0"
Genipa americana var.caruto	Bernardi	20626	NY	Na	Paraguay	Concepcion	Inter Toldo	N 9° 0' 0"	W 80° 0' 0"
Genipa americana var.caruto	Javier Fernandez Casas	8539	NY	Na	Bolivia	Pando, Manuripi	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa americana var.caruto	M. Nee	37848	NY	Na	Bolivia	Santa Cruz	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa americana var.caruto	Jose Steinbach	7234	NY	Na	Bolivia	Santa Cruz	Sara	S 17° 47' 10"	W 63° 10' 52"
Genipa americana var.caruto	R. Goodland	773	NY	Na	British Guiana	Rapununi Northern Savanna	NA	N 5° 0' 0"	W 59° 0' 0"
Genipa americana var.caruto	W. L. Balee	800	NY	Na	Brasil	Maranhao	NA	NA	NA
Genipa americana var.caruto	L. V. Ferreira	86	NY	Na	Brasil	Amazonas	Manaus	03°15' S	60° 00' W
Genipa americana var.caruto	Francis Drouet	2644	NY	Na	Brasil	Ceara	Maracanaú <Maranguape>	NA	NA
Genipa americana var.caruto	Hopkins M.J.G.	577	NY	Na	Brasil	T. F. Roraima	NA	03°22' N	61° 20' W
Genipa americana var.caruto	P. Kamb	2029	NY	NA	Belize (British Hondu	NA	NA	NA	NA
Genipa americana var.caruto	T. G. Yuncker	8646	NY	NA	Honduras	Yoro	Coyoles	N 15° 24' 0"	W 86° 40' 0"
Genipa americana var.caruto	Robert Merrill King	1896	NY	NA	Mexico	Oaxca	Zanatepec	NA	NA
Genipa americana var.caruto	H. H. Smith	1793	NY	NA	St Vincents	na	na	NA	NA
Genipa americana var.caruto	Julian A. Steyermark	51455	NY	NA	Guatemala	Huehuetenango	NA	NA	NA
Genipa americana var.caruto	Ducke	1385	NY	Na	Brasil	Amazonas	Boa Vista	S 17° 47' 10"	W 63° 10' 52"
Genipa americana var.caruto	Dr E. Hassler	7851	NY	NA	Paraguay	Centralis	NA	NA	NA
Genipa americana var.caruto	Dr E. Hassler	7953	NY	NA	Paraguay	NA	NA	NA	NA
Genipa americana var.caruto	Prof Pedro Jorgensen	3694	NY	NA	Paraguay	NA	NA	NA	NA
Genipa americana var.caruto	Elbert L. Little Jr.	8239	NY	NA	Colombia	NA	Intendencia del Meta	N 3° 45' 0"	W 76° 30' 0"
Genipa americana var.caruto	R. Romero-Castaneda	9655	NY	NA	Colombia	Bolivar	aldrededores de Since	NA	NA
Genipa americana var.caruto	J. H. E. Rova	2388	GB	187582	Panama	Panama	NA	08° 42' N	79° 52' W
Genipa americana var.caruto	T. G. Yuncker	8646	NY	NA	Honduras	Yoro	Coyoles	NA	NA
Genipa americana var.caruto	Delprete P. G:	10763	GB	NA	French Guiana	NA	NA	NA	NA

Species	Collector name	Collector number	Herbarium code	Catalogue number	Country	State	Municipality	Lat	Long
Genipa infundibuliformis	Alexandre Antonelli	406	GB	NA	Brasil	Sao Paulo	Campinas	22° 54' 20" S	47° 3' 39" W
Genipa spruceana	C. Persson	1612	GB	NA	Ecuador	Orellana	NA	00° 40' 54.6"	76°21'49.2"
Genipa spruceana	C. Persson	606	GB	NA	Peru	Loreto	Maynas	03° 48' S	73° 25' W
Genipa spruceana	C. Persson	1802	GB	NA	Ecuador	Sucumbios	Canton Cuyabeno	00° 00' 02.8"	76°10'47.9"
Genipa spruceana	C. Persson	674	GB	NA	Peru	Loreto	Maynas	03° 45' N	73° 16' W
Genipa spruceana	C. Persson	1959	GB	055699 16.02.96	French Guiana	Crique Tibourou	NA	04° 29' N	52°21' W
Genipa spruceana	A. Antonelli	246	GB	NA	Brasil	Amazonas	NA	00° 07' S	67° 04' E
Genipa spruceana	L.V. Ferreira	19	GB	185410	Brasil	Para	Santarem	02° 31' S	55° 00' W
Genipa spruceana	M. Tirado	308	GB	NA	Ecuador	Esmeraldas	Eloy Alfaro	00° 43' N	78° 53' W
Genipa spruceana	F. Ayala	5685	GB	NA	Peru	Loreto	Maynas	S 23° 23' 59"	W 57° 25' 56"
Genipa spruceana	X. Cornejo	7505	GB	NA	Ecuador	Sucumbios	NA	0° 18' 0" S	70° 9' 0" W
Genipa spruceana	J. Pedrol	4856	GB	NA	Venezuela	Territorio Federal Amazonas	NA	3° 27' N	65° 27' W
Genipa spruceana	J. Brandbyge	36008	GB	155889	Ecuador	Napo	NA	00° 01' S	76° 11' W
Genipa spruceana	H. L. Clark	7917	GB	NA	Venezuela	Territorio Federal Amazonas	San Carlos de Rio Negro	1° 56' N	67° 03' W
Genipa spruceana	R. Vasquez	8580	GB	165975	Peru	Loreto	Maynas	04° 29' S	73° 35' W
Genipa spruceana	R. Vasquez	1257	GB	165976	Peru	Loreto	Maynas	4° 15' S	73° 10' W
Genipa spruceana	R. Vasquez	16142	GB	166592	Peru	Loreto	Maynas	03° 20' S	72° 55' W
Genipa spruceana	C. Grandez	1731	GB	165974	Peru	Loreto	Maynas	03° 10' S	73° 20' W
Genipa spruceana	W. Palacios	9001	GB	172977	Ecuador	Sucumbios	Lago Argio Canton	00°00 ' S	76° 11' W
Genipa spruceana	W. Palacios	8097	GB	170255	Ecuador	Sucumbios	Lago Argio Canton	00° 33' N	75° 16' " W
Genipa spruceana	Carlos E. Ceron	4903	GB	128906	Ecuador	Napo	Canton Aguarico	01° 01' S	75° 47' W
Genipa spruceana	M. J. Jansen-Jacobs	3881	U	072417 03.10.97	Guyana	Upper Takutu-Upper Essequibo	Rupununi Distr.	02° 49' N	59°31' W
Genipa spruceana	M. J. Jansen-Jacobs	3906	U	NA	Guyana	Rupununi	NA	02° 49' N	59° 31' W
Genipa spruceana	R. S. Cowan	38761	U	54216 B	French Guiana	Roura	Montagne de Kaw	N 4° 0' 0"	W 53° 0' 0"
Genipa spruceana	W. C. Steward	164	NY	NA	Brasil	Amazonas	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	S. R. Hill	13201	NY	Na	Brasil	Amazonas	NA	3° 23' S	57° 44' W
Genipa spruceana	C.A. Cid Ferreira	8131	NY	Na	Brasil	Amazonas	Presidente Figueredo	01° 30' -2°00' S	59° 30' - 60° 00'W
Genipa spruceana	R. Spruce	Sine numero	NY	NA	Brasil	Para	NA	NA	NA
Genipa spruceana	Fleury M.	1131	U	081613 29.03.99	French Guiana	Saint-Laurent-du-Moroni	Taluwen <Talhuwen>	N 4° 0' 0"	W 53° 0' 0"
Genipa spruceana	Clarke, D.	6586	U	105410 14.07.03	Guyana	U. Takutu	U. Essequibo	03°10' N	58° 40' W
Genipa spruceana	M.J. Jansen-Jacobs	2307	U	040917 26.03.94	Guyana	Rupununi District	Kuyuwini Landing	02°10' N	59° 15' W
Genipa spruceana	C.D.K. Cook	75	U	36790B 6 Oct 1962	Guyana	NA	NA	N 16° 15' 0"	W 61° 35' 0"
Genipa spruceana	Clarke, D.	2690	U	082070 10.07.99	Guyana	Potaro-Siparuni	NA	04°43' N	58° 42' W
Genipa spruceana	R. Spruce	1735	NY	Na	Brasil	Barra	na	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	Ducke	1356	NY	Na	Brasil	Amazonas	Boa Vista	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	I. L. Amaral	187	NY	Na	Brasil	Amazonas	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	W. A. Egler	47663	NY	Na	Brasil	Amapa	Rio Oiapoque	03° 12' N	52° 19' W
Genipa spruceana	J. M. Pires	50408	NY	Na	Brasil	Amapa	Rio Araguari	01° 26' N	51° 58' W
Genipa spruceana	L. O. A. Teixeira	1037	NY	Na	Brasil	Amazonas	Humaita	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	J. M. Pires	51588 A	NY	Na	Brasil	Amapa	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	Shunsuke Tsugaru	B-1007	NY	Na	Brasil	Amazonas	100km upstream Manaus	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	L. O. A. Teixeira	104.773	NY	Na	Brasil	Amazonas	Humaita	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	G. T. Prance	5762	NY	NA	Brasil	Rondonia	NA	NA	NA
Genipa spruceana	Ducke	581	NY	NA	Brasil	Amazonas	Manaus	NA	NA
Genipa spruceana	J. M. Pires	52451	NY	NA	Brasil	Amapa	NA	1° 45' N	50° 58' W
Genipa spruceana	Bassett Maguire	60165	NY	NA	Brasil	Amazonas	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	Bassett Maguire	60165	NY	NA	Brasil	Amazonas	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	P. J. M. Maas	5436	U	NA	Guyana	NA	NA	6° N	57° 50' W
Genipa spruceana	P. J. M. Maas	5443	U	NA	Guyana	NA	NA	6° N	57° 50' W
Genipa spruceana	Hoffman, B.	2384	U	NA	Guyana	Cuyuni-Mazaruni	NA	06° 09' N	60° 15' W
Genipa spruceana	C. Persson	649	GB	NA	Peru	Loreto	Mariscal Ramon Castilla	03° 19' S	71° 49' W
Genipa spruceana	H.S. Irwin	47420	NY	NA	Brasil	Amapa	NA	03° 48' N	51° 53' W
Genipa spruceana	R. Spruce	1935	NY	NA	Brasil	Barra	Prov. Rio Negro	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	McDowell T.	2579	U	NA	Guyana	Cuyuni-Mazaruni	NA	05° 53' N	60° 37' W
Genipa spruceana	Henkel T. W.	3391	U	NA	Guyana	U. Takutu-U Essequibo	South Rupununi	02° 53' N	59° 18' W
Genipa spruceana	C. Persson	641	GB	NA	Peru	Loreto	Mariscal Ramon Castilla	03° 19' S	71° 49' W
Genipa spruceana	J. Lanjouw	906	U	NA	Surinam	NA	NA	NA	NA
Genipa spruceana	I. L. Amaral	8346	NY	Na	Brasil	Amazonas	Maraa	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	C. Persson	604	GB	NA	Peru	Loreto	Maynas	03° 48' S	73° 25' W
Genipa spruceana	R. Spruce	2495	NY	NA	Brasil	Amazonas	NA	S 17° 47' 10"	W 63° 10' 52"
Genipa spruceana	E. Asanza	30534	GB	157859	Ecuador	Napo	NA	00° 03' N	76° 10' W
Genipa spruceana	Wallschlagel	796	NY	NA	NA	NA	NA	NA	NA
Genipa spruceana	Hoffman, B.	2027	U	NA	Guyana	Cuyuni-Mazaruni	NA	06° 09' N	60° 17' W

Appendix 2

Key To the Genus *Genipa*

1a. Leaves glabrous except occasional trichomes at base of midvein abaxial side, interior calyx tube walls glabrous, minute fringe of hairs at base of interior calyx tube, interior corolla hairs dense at base, spreading at margin, tertiary leaf venation reticulate and fruit shrivelled and dry…………………………………………………………………………..….G.spruceana
1b. Leaves not glabrous, interior calyx not glabrous ……………..……………2

2a. Juvenile leaves lobed, flower tube taller than calyx, petal lobes patent to spreading calyx triangular acute lobes to 4mm long………………. ……………………….G. infundibuliformis
2b.Leaves entire, flower tube shorter than calyx, petals lax, calyx tube truncate or slightly lobed rarely dentate…………………………………………… …………3

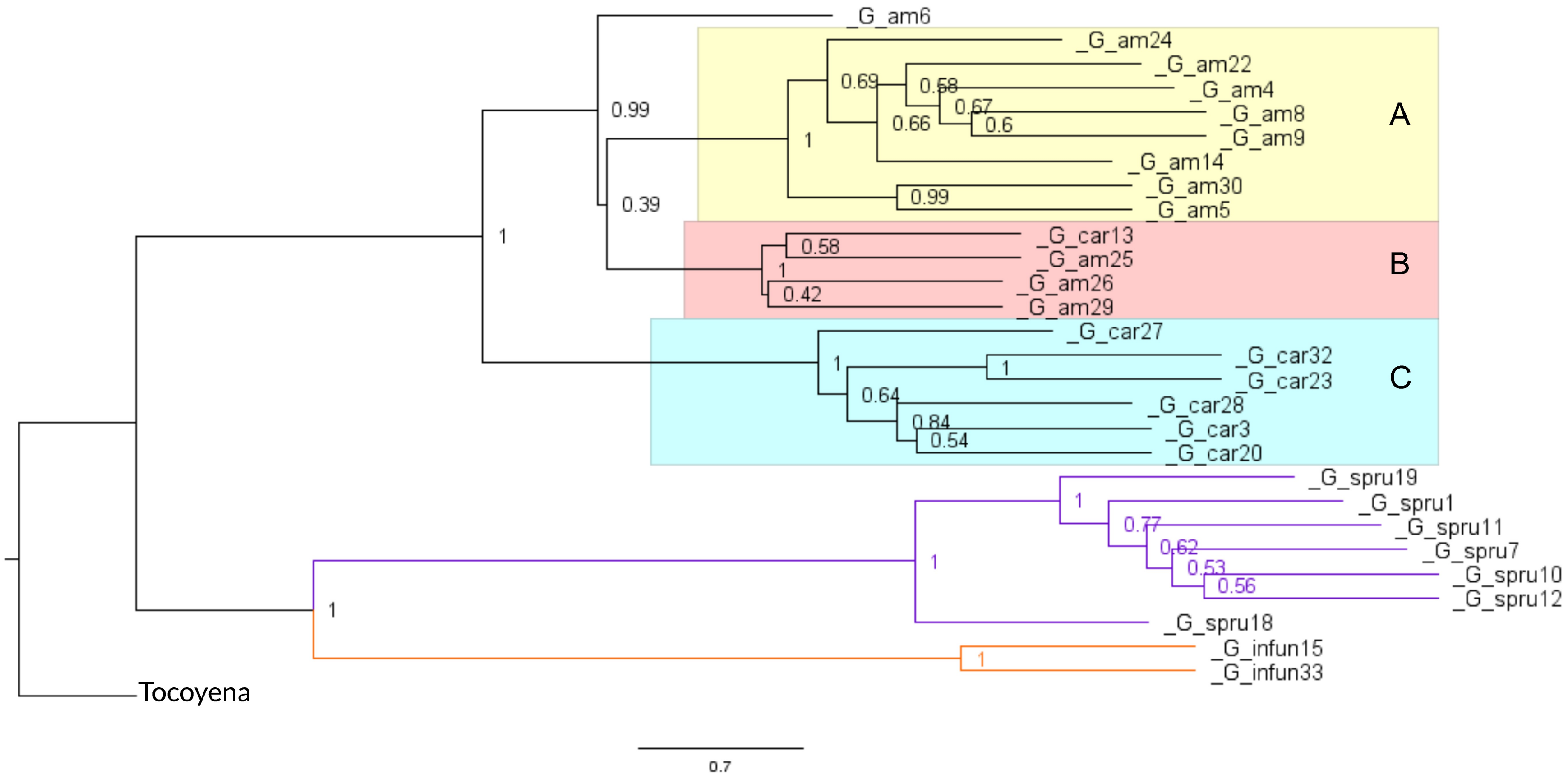
3a. Abaxial surface of leaf lamina, primary and secondary veins densely soft tomentose, ……………………………………………. …………………………...G. americana var caruto
3b. Abaxial surface of leaf lamina glabrous, only primary and secondary veins pubescent, inner calyx tube walls supposed sericious……..……………….. ……….......G. americana var americana

Appendix 3

Summary Table of Morphological Traits

Continuous Traits (mm)			Categorical Trait (character states)		
1.	LL	Longest leaf length, petiole included	1	Marg	Leaf margin (entire 0; lobed 1)
1.	LW	Widest leaf width	1	LInd	Abaxial leaf indumentum (glabrous/nearly glabrous 0; veins only 1; lamina and veins 2)
1.	LDis	Distance from widest point to leaf tip	1	InCol	Indumentum colour (white/beige/cream/straw 0; orange 1; dark brown/black 2)
1.	FrNo	Fruit number	1	LCol	Adaxial and abaxial leaf similar colour (no 0; yes 1)
1.	FrW	Fruit width at widest point	1	Bi	Flowers bisexual (yes bisexual 0; staminate 1; carpellate 2)
1.	FrDis	Fruit Distance from widest to tip	1	CoEx	Corolla indumentum exterior (glabrous 0; pubescent 1; tomentose 2)
1.	FrL	Fruit length	1	CoIn	Corolla indumentum interior (glabrous 0; pubescent 1; tomentose 2; short above, long below 3)
			2	CxEx	Calyx exterior indumentum (glabrous 0; pubescent 1)
			3	CxIn	Calyx interior indumentum (glabrous 0; glabrous above minute hairs below1; sericious 2; sericious above tomentose below 3; tomentose 4)

Appendix 4 Astral-III phylogeny



Cladogram produced using ASTRAL-III, of 28 *Genipa* samples, based on 38 nuclear loci, tree rooted on *Tocoyena pittieri,* with ASTRAL local posterior branch support shown.

Appendix 5
Morphological results table

Morphological characters: LL:leaf length, LW: leaf width, LDis: Distance from widest point to leaf tip, FrL: fruit length, FrW: fruit width at widest point, FrDis: Fruit Distance from widest to tip, FrNo: fruit number, Marg: leaf margin, LInd: leaf indumentum, InCol: indumentum colour, LCol: leaf abaxial colour differs from adaxial colour, Bi: flower bisexual, CoEx: corolla exterior indumentum, CoIn: corolla interior indumentum, CxEx: calyx exterior indumentum, CxIn: calyx interior indumentum.

Taxon	Collector Number	LL	LW	LDis	FrL	FrW	FrDis	FrNo	Marg	LInd	InCol	LCol	Bi	CoEx	CoIn	CxEx	CxIn
Genipa americana	46937	219	85	90	50	40	19	1	entire	primary and secondary veins only	na	yes	na	na	na	na	na
Genipa americana	55	257	85	102	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	21187	384	166	156	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	8954	231	89	94	50	50	25	1	entire	primary and secondary veins only	orange	yes	na	na	na	na	na
Genipa americana	5220	277	100	108	NA	NA	NA	NA	entire	primary and secondary veins only	orange	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	4960	241	95	105	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	staminate	tomentose	tomentose	glabrous	na
Genipa americana	2450	NA	129	110	NA	NA	NA	NA	entire	primary and secondary veins only	orange	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	59509	155	50	68	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	pubescent	na
Genipa americana	30432	296	103	127	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	staminate	tomentose	tomentose	glabrous	na
Genipa americana	38926	137	65	48	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	pubescent	sericeous
Genipa americana	2552	212	59	100	30	46	14	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana	4750	172	44	70	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	3030	160	71	67	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	12405	319	116	140	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana	10113	204	39	92	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	28024	173	44	71	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	31126	237	73	109	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	13653	195	55	91	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	staminate	tomentose	short hairs above and long below	glabrous	na
Genipa americana	10754	114	54	52	41	27	25	1	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	632	162	81	55	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	carpellate	tomentose	tomentose	glabrous	sericeous
Genipa americana	49	225	73	75	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	3507	223	78	106	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	1212	206	105	79	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	270	322	90	135	47	46	20	1	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	12.085	212	64	71	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	na	tomentose	na	glabrous	na
Genipa americana	na	180	70	85	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana	11917	200	98	100	57	45	20	2	entire	primary and secondary veins only	dark brown/black	no	na	na	na	na	na
Genipa americana	36537	NA	NA	NA	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	pubescent	tomentose	glabrous	tomentose
Genipa americana	3844	NA	NA	NA	NA	64	NA	1	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	4135	NA	NA	NA	45	33	NA	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	glabrous	glabrous
Genipa americana	2266	NA	NA	NA	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	pubescent	na	glabrous	tomentose
Genipa americana	20925	NA	NA	NA	54	44	NA	2	entire	primary and secondary veins only	orange	yes	na	na	na	na	na

Taxon	Collector Number	LL	LW	LDis	FrL	FrW	FrDis	FrNo	Marg	LInd	InCol	LCol	Bi	CoEx	CoIn	CxEx	CxIn
Genipa americana	612	314	96	101	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	no	bisexual	glabrous	tomentose	glabrous	na
Genipa americana	1865	276	142	126	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	1866	258	112	101	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	staminate	pubescent	na	glabrous	na
Genipa americana	342	235	104	109	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	na	pubescent	na	glabrous	sericeous
Genipa americana	2372	275	96	126	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	staminate	pubescent	tomentose	glabrous	sericeous
Genipa americana	306a	471	202	217	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	NA	na	na	na	na	na
Genipa americana	306b	208	57	86	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	pubescent	na	glabrous	sericeous
Genipa americana	2143	273	70	113	85	59	43	1	entire	glabrous/nearly glabrous	orange	yes	na	na	na	na	na
Genipa americana	16465	225	116	97	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	pubescent	pubescent	glabrous	na
Genipa americana	231	145	48	49	60	44	26	1	entire	primary and secondary veins only	orange	na	na	na	na	na	na
Genipa americana	225	256	108	90	70	55	48	1	entire	glabrous/nearly glabrous	na	yes	na	na	na	glabrous	glabrous
Genipa americana	141	270	84	112	NA	55	NA	1	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana	1695	255	70	92	48	35	32	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	pubescent	pubescent	glabrous	na
Genipa americana	24304	385	160	145	32	39	22	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	sericeous
Genipa americana	11263	NA	133	NA	30	25	NA	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	glabrous	glabrous
Genipa americana	1633	190	65	77	66	54	NA	1	entire	primary and secondary veins only	orange	yes	na	na	na	glabrous	sericeous
Genipa americana	1336	285	96	107	56	34	35	1	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	3465	328	90	132	48	48	21	1	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	5820	187	59	60	79	60	37	1	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	30	190	72	85	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	pubescent	sericeous
Genipa americana	44922	295	108	145	32	30	21	1	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	529	415	150	195	34	34	13	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	71518	299	137	112	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	na	tomentose	na	glabrous	sericeous
Genipa americana	152	280	124	100	21	20	7	1	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	20286	275	102	95	58	31	20	1	entire	primary and secondary veins only	orange	yes	na	na	na	na	na
Genipa americana	5366	300	110	90	50	44	20	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana	23043	170	67	56	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	133	380	120	141	35	35	15	1	entire	primary and secondary veins only	orange	no	na	na	na	na	na
Genipa americana	209	282	96	123	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	bisexual	pubescent	tomentose	glabrous	sericeous
Genipa americana	5123	273	85	101	NA	65	NA	1	entire	primary and secondary veins only	white/cream/beige/straw	NA	na	na	na	na	na
Genipa americana	2535	320	95	105	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	pubescent	na	glabrous	sericeous
Genipa americana	1852	222	76	95	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	carpellate	tomentose	tomentose	glabrous	sericeous
Genipa americana	701	258	110	98	46	40	20	1	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana	80	294	116	136	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	tomentose	pubescent	glabrous	sericeous
Genipa americana	10394	149	70	80	37	35	NA	1	entire	lamina and veins	orange	no	na	na	na	na	na
Genipa americana	5805	232	107	93	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	7949	260	100	114	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana	8389	290	118	85	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	9916	197	64	79	NA	NA	NA	1	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana	1439	190	54	85	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	bisexual	glabrous	tomentose	glabrous	sericeous
Genipa americana	2646	395	160	165	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	13845	201	79.5	76	NA	NA	NA	NA	entire	glabrous/nearly glabrous	orange	yes	bisexual	na	short hairs above and long below	glabrous	na
Genipa americana	26586	203	62	71	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	pubescent	tomentose	glabrous	sericeous
Genipa americana	1649	247	67	85	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana	3776	185	84	76	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	4855	322	116	144	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	1552	230	80	100	NA	NA	NA	NA	entire	primary and secondary veins only	orange	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana	11.306	205	95	94	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana	11720	192	80	87	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	tomentose
Genipa americana	683	152	57	55	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana	2997	146	47	51	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	tomentose	tomentose	pubescent	na
Genipa americana	770	252	101	102	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	tomentose	na	glabrous	sericeous

Taxon	Collector Number	LL	LW	LDis	FrL	FrW	FrDis	FrNo	Marg	LInd	InCol	LCol	Bi	CoEx	CoIn	CxEx	CxIn
Genipa americana var.caruto	S-959 B	257	85	100	51	37	20	3	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana var.caruto	S-959 A	208	86	70	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	pubescent	tomentose	glabrous	na
Genipa americana var.caruto	5849	148	42	52	60	52	52	1	entire	lamina and veins	orange	yes	na	na	na	na	na
Genipa americana var.caruto	2402	277	128	95	52	45	29	1	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana var.caruto	3680	238	110	80	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	staminate	tomentose	tomentose	glabrous	na
Genipa americana var.caruto	1976	306	128	125	NA	NA	NA	NA	entire	lamina and veins	na	yes	na	na	na	na	na
Genipa americana var.caruto	4031	171	68	70	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa americana var.caruto	12403	175	58	74	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	short hairs above and long below	glabrous	na
Genipa americana var.caruto	24850	232	65	100	58	42	29	1	entire	NA	orange	yes	na	na	na	na	na
Genipa americana var.caruto	11280	NA	NA	NA	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	NA	bisexual	pubescent	tomentose	glabrous	sericeous
Genipa americana var.caruto	1608	139	40	40	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	pubescent	na	glabrous	glabrous above pubescent at base
Genipa americana var.caruto	418	200	92	92	34	32	20	1	entire	lamina and veins	orange	no	na	na	na	na	na
Genipa americana var.caruto	NA	340	142	115	NA	45	NA	3	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana var.caruto	344	216	100	95	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	tomentose
Genipa americana var.caruto	1007	245	109	95	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	short hairs above and long below	glabrous	na
Genipa americana var.caruto	3657	252	147	75	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	na	tomentose	na	glabrous	sericeous
Genipa americana var.caruto	10147	180	91	85	47	44	21	1	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana var.caruto	8155	170	69	65	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	pubescent	glabrous	na
Genipa americana var.caruto	na	224	92	89	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	na	tomentose	na	glabrous	na
Genipa americana var.caruto	290	243	98	100	NA	NA	NA	NA	entire	lamina and veins	orange	no	na	tomentose	na	glabrous	sericeous
Genipa americana var.caruto	5418	358	177	126	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	442	169	46	70	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	glabrous	na
Genipa americana var.caruto	7488	55	59	55	NA	NA	NA	NA	entire	primary and secondary veins only	orange	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	3211	116	90	41	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	glabrous	na
Genipa americana var.caruto	6315	270	104	109	NA	NA	NA	NA	entire	lamina and veins	orange	yes	na	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	7374	158	71	58	NA	NA	NA	NA	entire	lamina and veins	orange	yes	staminate	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	1065	135	76	55	NA	NA	NA	NA	entire	lamina and veins	na	no	bisexual	tomentose	short hairs above and long below	glabrous	na
Genipa americana var.caruto	480	240	99	90	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	staminate	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana var.caruto	1416	142	79	48	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	2183	195	80	87	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana var.caruto	1384	224	114	71	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	staminate	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana var.caruto	3952	188	80	78	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	short hairs above and long below	pubescent	sericeous
Genipa americana var.caruto	409	245	93	111	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana var.caruto	4419	182	84	86	30	23	21	1	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	short hairs above and long below	glabrous	na
Genipa americana var.caruto	20626	220	50	65	38	35	18	2	entire	lamina and veins	orange	no	na	na	na	na	na
Genipa americana var.caruto	8539	248	108	79	45	52	25	2	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana var.caruto	37848	254	76	109	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	pubescent	short hairs above and long below	pubescent	tomentose
Genipa americana var.caruto	7234	220	82	75	NA	NA	NA	NA	entire	lamina and veins	orange	yes	bisexual	tomentose	tomentose	glabrous	tomentose
Genipa americana var.caruto	773	373	209	174	39	34	15	1	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana var.caruto	800	221	86	105	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	86	314	120	142	12	20	5	2	entire	lamina and veins	white/cream/beige/straw	yes	na	na	na	na	na
Genipa americana var.caruto	2644	166	55	71	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	pubescent	sericeous
Genipa americana var.caruto	577	241	134	86	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	yes	na	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	2029	208	106	80	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	pubescent	na
Genipa americana var.caruto	8646	230	100	70	45	39	23	NA	entire	lamina and veins	white/cream/beige/straw	no	na	na	na	na	na
Genipa americana var.caruto	1896	185	66	66	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	1793	241	122	87	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	pubescent	na
Genipa americana var.caruto	51455	320	179	110	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	glabrous	na
Genipa americana var.caruto	1385	211	140	91	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	glabrous	sericeous
Genipa americana var.caruto	7851	228	80	80	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	short hairs above and long below	glabrous	sericeous
Genipa americana var.caruto	7953	131	37	65	NA	NA	NA	NA	entire	lamina and veins	white/cream/beige/straw	no	bisexual	tomentose	tomentose	pubescent	sericeous
Genipa americana var.caruto	3694	115	42	57	NA	NA	NA	NA	entire	lamina and veins	orange	no	bisexual	tomentose	tomentose	pubescent	tomentose

Taxon	Collector Number	LL	LW	LDis	FrL	FrW	FrDis	FrNo	Marg	LInd	InCol	LCol	Bi	CoEx	CoIn	CxEx	CxIn
Genipa infundibuliformis	406	340	NA	120	NA	NA	NA	NA	lobed	glabrous/nearly glabrous	na	na	staminate	tomentose	tomentose	glabrous	na
Genipa spruceana	1612	168	50	75	20	20	10	4	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa spruceana	606	294	112	126	22	18	11	3	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa spruceana	1802	260	67	100	20	19	10	3	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa spruceana	674	NA	NA	NA	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	1959	270	125	95	40	40	12	2	entire	glabrous/nearly glabrous	na	yes	na	na	na	na	na
Genipa spruceana	246	215	94	96	NA	NA	NA	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	19	212	82	72	26	30	16	3	entire	glabrous/nearly glabrous	na	yes	na	na	na	na	na
Genipa spruceana	308	320	158	120	64	55	30	1	entire	primary and secondary veins only	orange	NA	na	na	na	na	na
Genipa spruceana	5685	241	82	87	NA	NA	NA	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	tomentose	short hairs above and long below	glabrous	na
Genipa spruceana	7505	223	90	89	23	24	12	3	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	glabrous	glabrous
Genipa spruceana	4856	370	150	159	17	25	9	2	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	36008	245	79	76	35	34	18	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	na	na	na	na	glabrous	na
Genipa spruceana	7917	402	165	180	23	24	11	1	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa spruceana	8580	270	102	112	25	32	14	4	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	1257	270	103	112	27	21	12	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	glabrous	na
Genipa spruceana	16142	360	147	161	38	36	12	3	entire	glabrous/nearly glabrous	dark brown/black	yes	na	na	na	na	na
Genipa spruceana	1731	267	85	110	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	na	na	pubescent	na	glabrous	glabrous
Genipa spruceana	9001	255	102	92	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	glabrous	na
Genipa spruceana	8097	216	57	87	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	na	short hairs above and long below	glabrous	na
Genipa spruceana	4903	291	120	126	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	NA	na	na	na	glabrous	glabrous
Genipa spruceana	3881	151	57	60	55	50	17	1	entire	glabrous/nearly glabrous	na	yes	na	na	na	na	na
Genipa spruceana	3906	NA	NA	NA	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	no	bisexual	pubescent	short hairs above and long below	glabrous	na
Genipa spruceana	38761	290	93	141	40	45	10	1	entire	glabrous/nearly glabrous	na	yes	na	na	na	na	na
Genipa spruceana	164	334	135	142	28	24	11	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	13201	132	56	41	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	carpellate	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	8131	303	130	114	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	short hairs above and long below	glabrous	glabrous
Genipa spruceana	ne numero	197	57	83	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	tomentose	na	glabrous	glabrous
Genipa spruceana	1131	182	64	80	28	27	15	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	6586	220	82	60	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	tomentose	tomentose	glabrous	na
Genipa spruceana	2307	206	64	56	30	32	12	3	entire	glabrous/nearly glabrous	na	na	na	na	na	na	na
Genipa spruceana	75	120	50	45	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	staminate	tomentose	tomentose	glabrous	na
Genipa spruceana	2690	142	55	50	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	no	bisexual	pubescent	tomentose	glabrous	na
Genipa spruceana	1735	235	72	106	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	na	tomentose	na	glabrous	glabrous above pubescent at base
Genipa spruceana	1356	150	68	61	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	tomentose	tomentose	glabrous	glabrous above pubescent at base
Genipa spruceana	187	231	77	104	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	staminate	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	47663	330	114	95	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	tomentose	glabrous	glabrous above pubescent at base
Genipa spruceana	50408	300	116	112	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	tomentose	na	glabrous	glabrous above pubescent at base
Genipa spruceana	1037	273	120	122	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	51588 A	490	128	168	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	B-1007	220	79	85	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	104.773	274	120	115	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	pubescent	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	5762	330	145	144	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	bisexual	tomentose	pubescent	glabrous	glabrous above pubescent at base
Genipa spruceana	581	166	75	89	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	yes	na	pubescent	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	52451	200	72	70	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	pubescent	short hairs above and long below	glabrous	glabrous
Genipa spruceana	60165	470	200	204	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	bisexual	pubescent	pubescent	glabrous	na
Genipa spruceana	60165	452	200	219	NA	NA	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	na	na
Genipa spruceana	5436	280	102	140	NA	NA	NA	NA	entire	glabrous/nearly glabrous	na	no	staminate	tomentose	tomentose	glabrous	glabrous above pubescent at base
Genipa spruceana	5443	240	96	100	23	22	13	2	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	na	na	na	na	na
Genipa spruceana	2384	227	61	87	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	no	bisexual	pubescent	short hairs above and long below	glabrous	glabrous above pubescent at base
Genipa spruceana	649	NA	NA	NA	24	21	NA	NA	entire	primary and secondary veins only	white/cream/beige/straw	yes	na	na	na	glabrous	tomentose
Genipa spruceana	47420	175	71	71	NA	NA	NA	NA	entire	glabrous/nearly glabrous	white/cream/beige/straw	yes	bisexual	pubescent	short hairs above and long below	glabrous	glabrous above pubescent at base

Appendix 6

Contingency table of categorical traits

		americana (N=87)	caruto (N=52)	infundibuliformis (N=1)	spruceana (N=51)	Total (N=191)	p value
LInd							< 0.001
	N-Miss	0	1	0	0	1	
	glabrous/nearly glabrous	29 (33.3%)	0 (0.0%)	1 (100.0%)	44 (86.3%)	74 (38.9%)	
	lamina and veins	12 (13.8%)	47 (92.2%)	0 (0.0%)	0 (0.0%)	59 (31.1%)	
	primary and secondary veins only	46 (52.9%)	4 (7.8%)	0 (0.0%)	7 (13.7%)	57 (30.0%)	
InCol							< 0.001
	dark brown/black	1 (1.1%)	0 (0.0%)	0 (0.0%)	1 (2.0%)	2 (1.0%)	
	na	18 (20.7%)	2 (3.8%)	1 (100.0%)	20 (39.2%)	41 (21.5%)	
	orange	18 (20.7%)	16 (30.8%)	0 (0.0%)	1 (2.0%)	35 (18.3%)	
	white/cream/beige/straw	50 (57.5%)	34 (65.4%)	0 (0.0%)	29 (56.9%)	113 (59.2%)	
LCol							< 0.001
	N-Miss	2	1	0	2	5	
	na	1 (1.2%)	0 (0.0%)	1 (100.0%)	3 (6.1%)	5 (2.7%)	
	no	18 (21.2%)	32 (62.7%)	0 (0.0%)	9 (18.4%)	59 (31.7%)	
	yes	66 (77.6%)	19 (37.3%)	0 (0.0%)	37 (75.5%)	122 (65.6%)	
Bi							0.033
	bisexual	34 (39.1%)	29 (55.8%)	0 (0.0%)	17 (33.3%)	80 (41.9%)	
	carpellate	2 (2.3%)	0 (0.0%)	0 (0.0%)	1 (2.0%)	3 (1.6%)	
	na	41 (47.1%)	19 (36.5%)	0 (0.0%)	30 (58.8%)	90 (47.1%)	
	staminate	10 (11.5%)	4 (7.7%)	1 (100.0%)	3 (5.9%)	18 (9.4%)	
CoEx							0.072
	glabrous	2 (2.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (1.0%)	
	na	33 (37.9%)	13 (25.0%)	0 (0.0%)	25 (49.0%)	71 (37.2%)	
	pubescent	12 (13.8%)	4 (7.7%)	0 (0.0%)	9 (17.6%)	25 (13.1%)	
	tomentose	40 (46.0%)	35 (67.3%)	1 (100.0%)	17 (33.3%)	93 (48.7%)	
CoIn							< 0.001
	na	41 (47.1%)	17 (32.7%)	0 (0.0%)	28 (54.9%)	86 (45.0%)	
	pubescent	3 (3.4%)	1 (1.9%)	0 (0.0%)	2 (3.9%)	6 (3.1%)	
	short hairs above and long below	3 (3.4%)	11 (21.2%)	0 (0.0%)	15 (29.4%)	29 (15.2%)	
	tomentose	40 (46.0%)	23 (44.2%)	1 (100.0%)	6 (11.8%)	70 (36.6%)	
CxEx							0.123
	glabrous	56 (64.4%)	32 (61.5%)	1 (100.0%)	33 (64.7%)	122 (63.9%)	
	na	27 (31.0%)	13 (25.0%)	0 (0.0%)	18 (35.3%)	58 (30.4%)	
	pubescent	4 (4.6%)	7 (13.5%)	0 (0.0%)	0 (0.0%)	11 (5.8%)	
CxIn							< 0.001
	glabrous	4 (4.6%)	0 (0.0%)	0 (0.0%)	6 (11.8%)	10 (5.2%)	
	glabrous above pubescent at base	0 (0.0%)	1 (1.9%)	0 (0.0%)	16 (31.4%)	17 (8.9%)	
	na	46 (52.9%)	27 (51.9%)	1 (100.0%)	28 (54.9%)	102 (53.4%)	
	sericeous	34 (39.1%)	20 (38.5%)	0 (0.0%)	0 (0.0%)	54 (28.3%)	
	tomentose	3 (3.4%)	4 (7.7%)	0 (0.0%)	1 (2.0%)	8 (4.2%)	

Popular Science Summary

Tats and Taxonomy

Genipa are trees (Photo 1) in the coffee family (Rubiaceae) from Central, South America and the Caribbean. The flowers (Photo 2) are large and showy and bear resemblance to the closely related *Gardenia* popular with European gardeners. *Genipa* is well known and it is important for cultural and economic reasons. It is used for timber, food, drinks and medicine. The most notable use is for Jagua tattoos as shown in Photo 3. An ink extracted from the unripe *Genipa* fruit (Photo 4) that has been used by indigenous groups in the Amazon and Caribbean for hundreds of years. The tattoos can be for coming of age ceremonies, purification rituals or more pragmatic reasons such as insect protection. The largest indigenous group in the Amazon – the Ticuna people, are actually named after the use of *Genipa*, Ticuna translating as “men, painted black”.

Photo 1 Genipa americana tree

Photo 2 Genipa flower showing male stamen and female stigma

Photo 3: Jagua body paint

Photo 4: Fruit oxidised black

Despite the importance of *Genipa* there are many gaps in our knowledge. One crucial gap is that there is no consensus on how many species of *Genipa* exist. In this study I attempt to determine the number of species and how the species are related to one another primarily using genomic data complemented with information about distribution and morphology.