

Universidad Nacional de Costa Rica
Facultad de Ciencias Exactas y Naturales
Escuela de Ciencias Biológicas
Licenciatura en Biotecnología

Informe Escrito Final

**Genetic diversity and structure of *Acrocomia aculeata* (Jacq.) Lodd. ex Mart.
(Arecaceae) using microsatellite DNA markers in Costa Rica**

**Diversidad y estructura genética de *Acrocomia aculeata* (Jacq.) Lodd. ex
Mart. (Arecaceae) utilizando marcadores de ADN tipo microsatélite en Costa
Rica**

**Artículo científico presentado como requisito parcial para optar por el grado
de Licenciatura en Biotecnología**

María Valeria Navarro Cascante

Campus Omar Dengo
Heredia, 2022

Este trabajo de graduación fue aprobado por el Tribunal Examinador de la Escuela de Ciencias Biológicas de la Universidad Nacional, como requisito parcial para optar por el grado de Licenciatura en Biotecnología.

ANGEL FRANCISCO HERRERA ULLOA (FIRMA)
Firmado digitalmente por ANGEL FRANCISCO HERRERA ULLOA (FIRMA)
Fecha: 2023.06.27 11:26:55 -06'00'

Dr. Ángel Herrera Ulloa

Representante, Decano, quién preside

LUIS ADRIAN VEGA CORRALES (FIRMA)
Firmado digitalmente por LUIS ADRIAN VEGA CORRALES (FIRMA)
Fecha: 2023.07.06 14:04:33 -06'00'

M.Sc. Luis Vega Corrales

Representante Unidad Académica

EMMANUEL MAURICIO ARAYA VALVERDE (FIRMA)
Firmado digitalmente por EMMANUEL MAURICIO ARAYA VALVERDE (FIRMA)
Fecha: 2023.06.27 14:17:04 -06'00'

M.Sc. Emmanuel Araya Valverde

Tutor

STEFANY SOLANO GONZALEZ (FIRMA)
PERSONA FISICA, CPF-01-1361-0067.
Fecha declarada: 06/07/2023 12:02:40 p. m.
Esta es una representación gráfica únicamente,
verifique la validez de la firma.

Dra. Stefany Solano González

Asesora

RAMON MOLINA BRAVO (FIRMA)
PERSONA FISICA, CPF-08-0134-0437.
Fecha declarada: 28/06/2023 01:19:26 PM
Esta es una representación gráfica únicamente,
verifique la validez de la firma.

Dr. Ramón Molina Bravo

Asesor

FRANK MARTIN SOLANO CAMPOS (FIRMA)
Firmado digitalmente por FRANK MARTIN SOLANO CAMPOS (FIRMA)
Fecha: 2023.07.06 12:09:33 -06'00'

Dr. Frank Solano Campos

Invitado especial

Agradecimientos

En primera instancia, quiero agradecer a mi familia, principalmente a mi madre, Giselle Cascante, mi abuela Carmen Esquivel y mi hermana Daniela Arroyo, por ser la voz de esperanza y darme el abrazo que me reconfortó el alma cuando más lo necesité. A mi sobrino Mateo, el aguacate de mi corazón que, aunque aún no ha mirado el sol, alegra mi alma y la llena de ilusión. A Katherine Porras, mi mejor amiga y mi alma gemela, quien me prestó su hombro para llorar y su mano para levantarme, aun cuando yo no podía estar de pie sola. A mi novio, Ariel Vega, porque me inspira todos los días a luchar por mis sueños y ve en mí a la humana y profesional que deseo ser algún día. Y, sobre todo, quiero agradecer a Dios, por alumbrarme el camino cuando se puso tan oscuro que no sabía por dónde seguir.

Este proyecto se llevó a cabo gracias al esfuerzo conjunto de un gran equipo de profesionales. Cada uno de ellos formó parte fundamental para concluir satisfactoriamente con el estudio que marca la línea base de investigación en relación con coyol como potencial fuente bioenergética. Por tanto, quiero agradecer al señor Msc. Emmanuel Araya Valverde, coordinador del área de genómica del Centro Nacional de Innovaciones Biotecnológicas – CENIBiot y mi maestro desde hace cuatro años. Quien en todo momento me apoyó y me inspiró a descubrir el mundo de la ciencia y la investigación. Le agradezco por todo el conocimiento que me heredó, y, sobre todo, porque creyó en mí, cuando apenas empezaba mi camino profesional y confió en lo que soy capaz de lograr con esfuerzo y constancia.

A la investigadora Msc. Mónica Rojas Gómez, colega y compañera en este proyecto, quien me colaboró de forma atenta e incondicional con todo lo correspondiente a los análisis estadísticos. Quien, con su paciencia y amor por la ciencia, me enseñó a utilizar herramientas sumamente valiosas para mi futuro como profesional y siempre se esmeró para lograr un trabajo de calidad.

A la PhD. Stephany Solano González y PhD. Ramón Molina Bravo, mis asesores de tesis que con su conocimiento y disponibilidad me brindaron su apoyo en el proceso de corrección y ajustes del contenido de mi proyecto.

A la Vicerrectoría de Investigación y Extensión (VIE) del Instituto Tecnológico de Costa Rica quien aportó los fondos para el financiamiento de este estudio bajo el proyecto 1510136 denominado “Proceso de domesticación del coyol (*Acrocomia aculeata*) como alternativa bioenergética en Costa Rica, Fase II”. Y finalmente, al Centro Nacional de Innovaciones Biotecnológicas (CENIBiot-CeNAT) quienes me ofrecieron sus instalaciones tecnológicas para llevar a cabo la mayoría de los ensayos técnicos, y que también formó parte fundamental para el apoyo financiero del proyecto (FP-026-2020). Además, a la Escuela de Biología de la Universidad de Costa Rica por el préstamo de su equipo analizador genético ABI 3130 (Applied biosystems) el cual fue de suma importancia para culminar esta investigación. A la Escuela de Ciencias Biológicas de la UNA por el proceso de enseñanza y excelencia que recibí mientras me formaba en esa institución.

A todos quienes en algún momento sirvieron de soporte para que esta investigadora soñadora lograra llegar a donde está, gracias por creer en mí y su apoyo incondicional.

Dedicatoria

*“La felicidad se puede hallar hasta en los más oscuros momentos,
si somos capaces de usar bien la luz.”*
— *Albus Percival Wulfric Brian Dumbledore.*

He atravesado caminos tan oscuros en este tiempo, y en algún momento creí que mi esfuerzo no era suficiente, pero la realidad es, que no podría haber completado este viaje sin las luces que iluminan mi camino a diario y me hacen feliz, mi madre Giselle Cascante y mi abuela Carmen Esquivel, a ellas que me enseñaron todo lo que sé acerca de la valentía, la perseverancia y la esperanza, les dedico este proyecto y les dedico mi amor hasta que mi corazón deje de latir.

Índice

Agradecimientos.....	I
Dedicatoria	III
Índice.....	IV
Índice de cuadros.....	V
Índice de figuras	VI
Abreviaturas o acrónimos.....	VII
Resumen	VIII
Abstract.....	IX
Genetic Diversity and Structure of <i>Acrocomia aculeata</i> (Jacq.) Lodd. ex Mart. (Arecaceae) Using Microsatellite DNA Markers in Costa Rica	1
Introduction	2
Materials and Methods	3
<i>Biological Material and Study Area</i>	3
<i>DNA Isolation and Microsatellite Genotyping Using Multiplex PCR</i>	5
<i>Microsatellite Informativeness, Genetic Diversity and Structure of <i>A. aculeata</i></i>	6
Results	7
<i>Microsatellite Informativeness</i>	7
<i>Genetic Diversity and Structure of <i>A. aculeata</i></i>	8
Discussion	12
<i>Genetic Diversity of <i>Acrocomia aculeata</i></i>	12
<i>Genetic Structure of <i>Acrocomia aculeata</i></i>	13
<i>Need for <i>Acrocomia aculeata</i> Genetic Characterization for Uses and Conservation Uses</i> .	14
References	17
Conclusiones.....	22
Recomendaciones.....	24
Anexos	25

Índice de cuadros

Table 1. Primer sequences with its respective fluorescent dye label of 12 microsatellites and the final concentration used for the multiplex PCR amplification in DNA of <i>A. aculeata</i>	6
Cuadro 1. Secuencias nucleotídicas 12 de los cebadores tipo microsatélite con su respectivo marcador de colorante fluorescente y la concentración final utilizada para la amplificación por PCR multiplex de los fragmentos de ADN en <i>A. aculeata</i>	6
Table 2. Genetic diversity parameters for nine polymorphic microsatellites amplified 185 DNA samples of <i>A. aculeata</i> . NA: number of alleles, He: expected heterozygosity, Ho: observed heterozygosity, PIC: Polymorphic Information Content and Fst as genetic differentiation coefficients.....	8
Cuadro 2. Parámetros de diversidad genética para nueve microsatélites polimórficos amplificados en el ADN de 185 muestras de <i>A. aculeata</i> . NA: número de alelos, He: heterocigosidad esperada, Ho: heterocigosidad observada, PIC: Contenido de Información Polimórfica y Fst como coeficiente de diferenciación genética.....	8
Table 3. Genetic diversity parameters estimated in the 18 sampling sites of <i>A. aculeata</i> in Costa Rica. He: expected heterozygosity, Ho: observed heterozygosity, Na: number of alleles, Ne: number of effective alleles, Np: number of private alleles.....	9
Cuadro 3. Parámetros de diversidad genética estimados para 18 sitios de muestreo de <i>A. aculeata</i> en Costa Rica. He: heterocigosidad esperada, Ho: heterocigosidad observada, Na: número de alelos, Ne: número de alelos efectivos, Np: número de alelos privados.....	9

Índice de figuras

Figure 1. Distribution of 18 sampled sites of *Acrocomia aculeata*, representing three clusters obtained thorough Structure. Sampled sites 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa. Green squares, red triangles and blue circles represent Cluster A, Cluster B and Cluster C, respectively. 4

Figura 1. Distribución de 18 sitios de muestreo de *Acrocomia aculeata*, representados en tres grupos según su estructura genética. Sitios muestreados 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa. Los cuadrados verdes, los triángulos rojos y los círculos azules representan el Grupo A, el Grupo B y el Grupo C, respectivamente..... 4

Figure 2. Population structure of the 185 samples of *A. aculeata* collected in Costa Rica. Clusters assignment of samples (k=3) is based on the Bayesian analysis on Structure Software using microsatellite markers genotyping..... 10

Figura 2. Estructura poblacional de las 185 muestras de *A. aculeata* colectadas en Costa Rica. La asignación de grupos de muestras (k = 3) fue basado en el análisis bayesiano usando el software STRUCTURE utilizando los datos de genotipado de marcadores de microsatélites. 10

Figure 3. UPGMA dendrogram of *Acrocomia aculeata*, estimated according to Nei's genetic distance. The dendrogram reflects the same distribution of 18 sampled sites represented in figure 1. Sampled sites 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa..... 11

Figura 3. Dendrograma UPGMA de *Acrocomia aculeata*, estimado según la distancia genética de Nei. El dendrograma refleja la misma distribución de 18 sitios muestreados representados en la figura 1. Sitios muestreados 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12:

Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa. 11

Abreviaturas o acrónimos

DNA	Ácido Desoxirribonucleico
PCR	Reacción en Cadena de la Polimerasa
CTAB	Bromuro de Hexadeciltrimetilamonio
DMSO	Dimetil Sulfóxido
MgCl ₂	Cloruro De Magnesio.
SSRs	Repeticiones de Secuencia Simple o microsatélites
bp	Pares de bases
Na	Número de alelos
Ne	Número de alelos efectivos
Np	Número de alelos privados
He	Heterocigosidad esperada
Ho	Heterocigosidad observada
PIC	Contenido de Información Polimórfica
Fst	Índice de fijación. Es la varianza de las frecuencias alélicas entre las poblaciones y la probabilidad de identidad por descendencia
Fis	Coefficiente de endogamia. Mide la reducción en la heterocigosidad individual debido a las desviaciones de los cruzamientos al azar.
AMOVA	Análisis de la Varianza Molecular
DPCA	Análisis Discriminante de Componentes Principales
UPGMA	Método de Agrupamiento de Pares Sin Ponderar con Media Aritmética
IBD	Aislamiento por distancia
RAPD	ADN polimórfico amplificado al azar
SNPs	Polimorfismos de Nucleótido Único

Resumen

Acrocomia aculeata es una especie de palmera arbórea ampliamente distribuida en el Neotrópico, desde México hasta Argentina, incluidas algunas islas del Caribe. El aceite de la fruta se utiliza para la producción de biodiesel, consumo humano y combustible de biomasa. A pesar de la necesidad de conocer mejor esta especie para su uso comercial y conservación, el conocimiento de la diversidad genética y la estructura de la misma es escaso en Mesoamérica. En este estudio, se usaron marcadores microsátélites para analizar 185 muestras de *A. aculeata*, de 18 sitios de muestreo, que cubrieron esencialmente la vertiente pacífica de Costa Rica. Así, se encontró una baja diversidad genética en los sitios de muestreo, con valores de H_e por debajo de 0,50, que es inferior a los niveles de diversidad encontrados en América del Sur. Curiosamente, las muestras recolectadas en los parques nacionales Santa Rosa y Rincón de la Vieja exhibieron la mayor diversidad genética ($H_e = 0.42$ y $H_e = 0.50$, respectivamente). Los 18 sitios de muestreo se estructuraron en tres conglomerados. El grupo A, abarcó muestras recolectadas en el Pacífico noroeste inferior, el Pacífico central y la región central de Costa Rica. El grupo B consistió en muestras de sitios del noroeste del Pacífico superior. El conglomerado C incorporó muestras de la región del Pacífico sureste, estando en gran parte separado de los conglomerados A y B, lo cual fue respaldado por los coeficientes de diferenciación más altos. Este es el primer estudio de diversidad genética de *A. aculeata* en Mesoamérica; por lo tanto, nuestros resultados sirven como referencia para futuros estudios sobre diversidad de germoplasma en esta región y como línea de base para futuros esfuerzos de conservación y manejo de la especie.

Palabras clave: Coyal, PCR multiplex, Mesoamérica, conservación, diferenciación y recursos genéticos, poblaciones discontinuas.

Abstract

Acrocomia aculeata is a tree palm species widely distributed throughout the Neotropics, from Mexico to Argentina including some Caribbean islands. The oil from the fruit is used for biodiesel production, human consumption, and biomass fuel. Despite the need to better understand its genetic diversity for commercial use and conservation, genetic diversity and structure knowledge of this species is scarce in Mesoamerica. We used microsatellite markers to analyze 185 samples of *A. aculeata*, from 18 sampling sites, that essentially covered geographic distribution of this species in Costa Rica. We found low genetic diversity across sampling sites, with H_e values below 0.50 which, is lower than diversity levels found in South America. Interestingly, samples collected in the national parks Santa Rosa and Rincón de la Vieja National exhibited the highest genetic diversity ($H_e = 0.42$ and $H_e = 0.50$, respectively). The 18 sampling sites were structured in three clusters. Cluster A contains samples collected in the lower northwestern Pacific, Central Pacific, and Central Region of Costa Rica. Cluster B consisted of samples from sites in the upper northwestern Pacific. Cluster C contains samples from the southeastern Pacific region, being largely separated from Clusters A and B which was supported by the highest differentiation coefficients. This is the first large genetic diversity study of *A. aculeata* in Mesoamerica; therefore, our results serve as reference for future studies on germplasm diversity studies in this region and a baseline for future conservation and management efforts of *A. aculeata*.

Keywords: Coyol, multiplex PCR, Mesoamerica, conservation, differentiation genetic resources, discontinuous populations.

Genetic Diversity and Structure of *Acrocomia aculeata* (Jacq.) Lodd. ex Mart. (Arecaceae) Using Microsatellite DNA Markers in Costa Rica

Valeria Navarro-Cascante^{1,8}, Elizabeth Arnáez-Serrano², Mónica Rojas-Gómez¹, Ileana Moreira², Guillermo Vargas-Hernández³, Nelson A. Zamora⁴, Elemer Briceño-Elizondo², Jonathan Morales-Marroquín^{5,6}, Pamela Sevilla-Cortés^{1,2}, Jesús Antonio Oviedo-Ulate^{1,7}, Emanuel Araya-Valverde^{1,2*}

¹ Centro Nacional de Innovaciones Biotecnológicas (CENIBiot), CeNAT-CONARE, 1174-1200, Pavas, San José, Costa Rica

V. Navarro-Cascante: vale_nc96@outlook.com, [ORCID: 0000-0003-3031-8020](https://orcid.org/0000-0003-3031-8020)

M. Rojas-Gómez: mrojas@cenat.ac.cr; morogo27@gmail.com [ORCID: 0000-0002-0245-2911](https://orcid.org/0000-0002-0245-2911)

E. Araya-Valverde: earaya@cenat.ac.cr, [ORCID: 0000-0002-0025-5420](https://orcid.org/0000-0002-0025-5420)

²Escuela de Biología, Instituto Tecnológico de Costa Rica, 159-7050, Cartago, Costa Rica

E. Arnáez-Serrano: earnaez@itcr.ac.cr, [ORCID: 0000-0003-4058-4429](https://orcid.org/0000-0003-4058-4429)

I. Moreira González: ilea2757@gmail.com, [ORCID: 0000-0001-94260986](https://orcid.org/0000-0001-94260986)

E. Briceño-Elizondo: ebriceno@uned.ac.cr, [ORCID: 0000-0003-1053-7330](https://orcid.org/0000-0003-1053-7330)

P. Sevilla-Cortés: sevillacortesp@gmail.com, [ORCID: 0000-0003-3853-4174](https://orcid.org/0000-0003-3853-4174)

³Estación Experimental Fabio Baudrit Moreno, Universidad de Costa Rica. Alajuela, 11501-2060, Costa Rica

G. Vargas-Hernández: guillermo.vargas@ucr.ac.cr, [ORCID: 0000-0003-1410-3898](https://orcid.org/0000-0003-1410-3898)

⁴Escuela de Ingeniería Forestal, Instituto Tecnológico de Costa Rica, 159-7050, Cartago, Costa Rica

N. Zamora: nzamora@itcr.ac.cr, [ORCID: 0000-0002-6242-3725](https://orcid.org/0000-0002-6242-3725)

J. Morales-Marroquín: jmorales1089@gmail.com, [ORCID: 0000-0002-7946-577X](https://orcid.org/0000-0002-7946-577X)

⁵Institute of Biology, Department of Genetics, University of Campinas UNICAMP, 13083-970, Campinas, Brazil. P.O.Box: 6109

⁶Centro de Pesquisa de Recursos Genéticos Vegetais, Instituto Agrônômico-IAC, Campinas, SP, Brazil.

⁷Universidad EARTH, PO box 4442-1000, Limón, Costa Rica.

JA Oviedo-Ulate: joviedou@earth.ac.cr, [ORCID: 0000-0002-4971-9556](https://orcid.org/0000-0002-4971-9556)

⁸Universidad Nacional, Escuela de Ciencias Biológicas, Heredia, Costa Rica.

V. Navarro-Cascante: vale_nc96@outlook.com, ORCID: 0000-0003-3031-8020

*Corresponding author: Emanuel Araya-Valverde

¹Centro Nacional de Innovaciones Biotecnológicas (CENIBiot) CeNAT-CONARE

1174-1200, San José, Costa Rica.

Phone + (506) 2519-5700 ext. 6090

ORCID identifier: 0000-0002-0025-5420

earaya@cenat.ac.cr

DOI: 10.1007/s10722-022-01501-8

<https://doi.org/10.1007/s10722-022-01501-8>.

Introduction

Acrocomia aculeata is widely distributed in the Neotropics from Mexico to Argentina, Bolivia, and Paraguay and also in the Caribbean (Henderson et al. 1995; Grayum 2003; Dransfield et al. 2008). Its center of origin is presumed to be in South America and the highest diversity appears to be in southeastern Brazil (Scariot et al. 1995; Lanes et al. 2015; de Lima et al. 2018). This conclusion is also supported by Diaz et al. (2021) who found the highest genetic diversity of *A. aculeata* in Brazil and the lowest diversity in samples from Central America. Moreover, fossil records demonstrated that human introduction of this palm tree to Central America came from South America (Morcote-Ríos and Bernal 2001).

In Costa Rica, *A. aculeata* frequently inhabits the Pacific lowlands, mostly from northwestern Guanacaste to the central Pacific slope and occasionally to the Osa Península and Punta Burica. It also occurs in the Guatuso plains in the northern Caribbean slope (Janzen 1983; Grayum 2003; Cornejo et al. 2012). Particularly, the occurrence of this species was studied in the Guanacaste where individuals are dispersed in the dry forest and pasturelands (Henderson et al. 1995; Esquivel-Mimenza et al. 2011). Cattle movement through pasturelands may facilitated dispersion in Costa Rica (Scariot 1998; Harvey et al. 2011).

The province of Guanacaste suffered accelerated deforestation during some 170 years, converting forest into livestock pastures. This land usage, in turn, promoted pastureland management that inhibited *A. aculeata* regeneration (Harvey et al. 2011; Calvo-Alvarado et al. 2009). In addition, the long-time practice of cutting down adult palm trees to obtain the sap, and fermenting into a beverage, also threatens the species. This practice primarily occurs throughout Central American, including Costa Rica with prevalence in Guanacaste (Janzen 1983; Balick 1990; Chizmar-Fernandez 2009; Díaz et al. 2021).

The *A. aculeata* fruit contains high oil content and good quality vegetable oil (Navarro-Díaz et al. 2014; Da Conceição et al. 2015; Lescano et al. 2015; Lieb et al. 2019) used to produce biodiesel, edible vegetable oil, and pharmaceutical products for the cosmetics industry. It also serves as a novel sustainable biomass feedstock and the post-fruit processing residues can be used as biomass fuel (Evaristo et al. 2016; Plath et al. 2016; Falasca et al. 2017). It has the potential to become a new commercial oilseed crop similar to that of oil palm if managed and bred properly (Colombo et al. 2018).

Studies in Costa Rica have quantified carotenoid and tocopherol content in fruits (Schex et al. 2018), oil content and fatty acid profile (Lieb et al. 2019), and variability of fruit morphology and oil composition in three locations (Alfaro-Solís et al. 2020). Despite these efforts, there are no studies yet that describe the genetic diversity of *A. aculeata* at large scale in Costa Rica. Molecular markers such as microsatellites and SNPs have played an important role in describing genetic diversity, structure and mating system of natural populations and germplasm collections of *A. aculeata* (Abreu et al. 2012; Lanes et al. 2015, 2016; Mengistu et al. 2016b; Silva et al. 2017; Coelho et al. 2018; Diaz et al. 2021; Laviola et al. 2022). Therefore, the aim of this study is to estimate the genetic diversity and to determine the structure of *A. aculeata* from samples collected in the range of geographic occurrence in Costa Rica, using microsatellite markers to generate a baseline for future use and conservation of this promising species in Costa Rica.

Materials and Methods

Biological Material and Study Area

The study area comprised four regions in Costa Rica i.e., Brunca, Central, Chorotega, and Huetar Norte with a total of 18 sampling sites (Supplementary Table 1). We collected 185 samples in the following 18 sites: (1) La Cruz, (2) Guanacaste National Park (hereinafter referred as Guanacaste), (3) Santa Rosa National Park (hereinafter referred as Santa Rosa), (4) Liberia, (5) Rincón de la Vieja National Park (hereinafter referred as Rincón de la Vieja), (6) Caño Negro National Wildlife Refuge (hereinafter referred as Caño Negro), (7) Santa Cruz, (8) Nicoya, (9) Barra Honda, (10) Cañas, (11) Tilarán, (12) Abangares, (13) Chomes, (14) Orotina, (15) Atenas, (16) Turrubares, (17) Paso Real, and (18) Península de Osa (Figure 1). This sampling essentially covered almost all of Costa Rica's current geographic occurrence of *A. aculeata*. The number of collected individuals per site is displayed in the Supplementary Table 1.

Each sampling site consisted of areas where mature reproductive trees (either flowering or fruiting) were growing relatively dispersed. The minimum distance between trees in each site was ca. 40-100 m, depending on the total number of trees. A leaflet lamina fragment of about 20 cm in length from each palm

tree sample was collected in conical plastic tubes of 50 mL containing table salt to avoid sample degradation. Upon arrival at the laboratory, the table salt was carefully removed, and samples were lyophilized (CHRIST BETA 1-8LD plus, Germany) prior to DNA extraction.

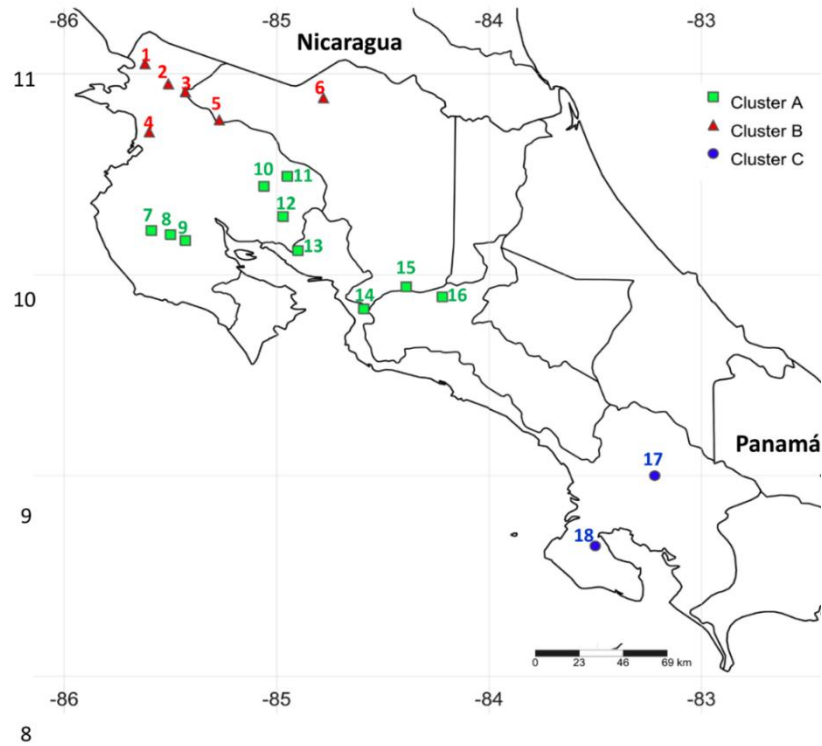


Figure 1. Distribution of 18 sampled sites of *Acrocomia aculeata*, representing three clusters obtained thorough Structure. Sampled sites 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa. Green squares, red triangles and blue circles represent Cluster A, Cluster B and Cluster C, respectively.

Figura 1. Distribución de 18 sitios de muestreo de *Acrocomia aculeata*, representados en tres grupos según su estructura genética. Sitios muestreados 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa. Los cuadrados verdes, los triángulos rojos y los círculos azules representan el Grupo A, el Grupo B y el Grupo C, respectivamente.

DNA Isolation and Microsatellite Genotyping Using Multiplex PCR

For DNA isolation, 50 mg of dried leaflet lamina were ground in a homogenizer Fastprep-24 (MP Biomedicals, USA) using lysing matrix A (MP Biomedical, USA). A CTAB-based method was employed for DNA isolation (Quirós-Guerrero et al. 2019). The DNA concentration was measured with a spectrophotometer (NanoDrop, Thermo Scientific, USA) and diluted (~30 ng/ μ L) for PCR amplification.

Twelve microsatellite markers previously developed for *A. aculeata* (Nucci 2007; Nucci et al. 2008) and *Astrocaryum aculeatum* (Ramos et al. 2012) were used. The selection criteria of microsatellites included the level of polymorphism, informativeness previously reported (Lanes et al. 2015, 2016; Mengistu et al. 2016b; Coelho et al. 2018) and expected allele size. The forward primer of each microsatellite was labeled with any of four fluorescent dye labels, i.e., 6-FAM, VIC, NED, and PET for multiplex PCR assays. Several primer combinations and concentrations were iteratively performed to obtain three PCR multiplexes (Table 1). Each multiplex PCR reaction (final volume 25 μ L) contained PCR Master Mix (1X DreamTaq Master Mix, ThermoFisher Scientific, USA), bovine serum albumin (0.01 mg, Sigma, USA), DMSO (2%, Sigma, USA), and DNA (~60 ng). Additionally, MgCl₂ (0.5 mM, Thermo Scientific, USA) and DNA Taq polymerase (1 Unit, ThermoFisher Scientific, USA) were added. The primer concentrations varied according to the optimized multiplexes (Table 1).

The amplification was performed in a thermal cycler (Veriti™, Applied Biosystems, USA) with the following profile: one cycle at 95 °C for 5 min followed by 35 cycles at 95 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, with a final extension at 72 °C for 8 min. Multiplex PCR products were visualized by capillary electrophoresis in the ABI 3130xl Genetic Analyzer (Applied Biosystems, USA) with the combination of 1.5 μ L of the multiplex PCR product, 0.4 μ L of GeneScan 600LIZ™ (Applied Biosystems, USA), and 8.5 μ L of Hi-Di Formamide (Applied Biosystems, USA). After capillary electrophoresis, allele binning and genotyping were manually inspected with GeneMapper™ v4.0 software (Applied Biosystems, USA) for further data analysis.

Table 1. Primer sequences with its respective fluorescent dye label of 12 microsatellites and the final concentration used for the multiplex PCR amplification in DNA of *A. aculeata*.

Cuadro 1. Secuencias nucleotídicas 12 de los cebadores tipo microsatélite con su respectivo marcador de colorante fluorescente y la concentración final utilizada para la amplificación por PCR multiplex de los fragmentos de ADN en *A. aculeata*.

Microsatellite name (SSR)	Fluorescent dye- Sequence of forward primer (5' - 3')	Sequence of reverse primer (5' - 3')	Final concentration (µM)	Multiplex PCR	Reference
Aacu10	NED- TGCCACATAGAGTGCTTGCT	CTACCACATCCCCGTGAGTT	0.15	1	Nucci et al. 2008
Aacu45	6FAM- CAGACTACCAGGCTTCCAGC	TCATCATCGCAGCTTGACTC	0.20	1	Nucci 2007
Aacu74	VIC- TACTGTTGTGCCAAGTCCCA	GAGCACAAGGGGGATATCAA	0.20	1	Nucci 2007
Aacu07	VIC- ATCGAAGGCCCTCCAATACT	AAATAAGGGGACCCTCCAA	0.19	2	Nucci et al. 2008
Aacu12	NED- GAATGTGCGTGCTCAAAAATG	AATGCCAAGTGACCAAGTCC	0.13	2	Nucci et al. 2008
Aacu26	6FAM- ACTTGCAGCCCCATATTCAG	CAGGAACAGAGGCAAGTTC	0.19	2	Nucci et al. 2008
Aacu30	6FAM- TGTGGAAGAAACAGGTCCC	TCGCCTTGAGAAATTATGGC	0.19	2	Nucci et al. 2008
Aac09	VIC- AGCGTAATAGCATCCCTTGC	AGTCCAGTGTTCTTCCCCTG	0.13	2	Ramos et al. 2012
Aac06	6FAM- TCTGATCCATCTGGTTGTCTAA	TGCATGGTGCTAGAGTAATCC	0.20	3	Ramos et al. 2012
Aac14	6FAM- GGCCAGTTGTGTTGATGAAA	TTAAGCCCCTGGTGAAAACA	0.20	3	Ramos et al. 2012
Aac04	VIC- GCATTGTCATCTGCAACCAC	GCAGGGGCCATAAGTCATAA	0.20	3	Ramos et al. 2012
Aac12	NED- GCTCTGTAATCTCGGCTTCT	TCCAGTTCAAGCTCTCTCAGC	0.20	3	Ramos et al. 2012

Microsatellite Informativeness, Genetic Diversity and Structure of A. aculeata

Genotyping data matrix was run in the R package *adegenet* (Jombart 2008) version 2.1.2 to create a genind class object with the data organized according to sampling sites. Each sampling site was

considered a putative population. Further, the R package *hierfstat* (Goudet 2005) version 0.04-22 was used to analyze the genetic diversity through the estimation of number of alleles per population, H_e , H_o , number of effective alleles (N_e), and private alleles (N_p). With the aid of the R package *pegas* (Paradis 2010) version 0.12, the inbreeding coefficient through F Statistic (F_{st}) was estimated (Wright 1965). Finally, the Polymorphic Information Content (PIC) was estimated using the R package *polysat* (Clark and Jasieniuk 2011) version 1.7. The intra- and interpopulation differentiation were also examined with an AMOVA using *poppr* (Kamvar et al. 2014) version 2.8.3 and *ade4* (Dray and Dufour 2007) version 1.7-13.

In order to study the genetic structure, the genotyping table was used in the software Structure version 2.3.4 (Pritchard et al. 2000) configured with eight repetitions for each possible number of groups (K from 1 to 8), burn-in of 50,000 and 100,000 repetitions of the Markov Chain Monte Carlo. The Structure results were used as input to estimate the best K number with Structure Harvester (Earl and von Holdt 2012) according to Evanno et al. (2005). This represents the most probable number of populations based on the inferred ancestry of each individual.

The genetic distance matrix of Nei was obtained with the R package *mmod* (Winter 2012) version 1.3.3. Besides, the genetic distance matrix was used in a UPGMA analysis to construct dendrograms in *MEGA6* (Tamura et al. 2013) and *DAPC* using *adegenet* (Jombart 2008) to identify the separation between clusters. Finally, the Mantel test was performed using the package *ade4* (Dray and Dufour 2007) to study the correlation between the matrix of genetic distances and geographic distances which results were visualized with an isolation by distance figure made with *MASS* (Venables and Ripley 2002) version 7.3-51.5.

Results

Microsatellite Informativeness

Three of 12 SSRs were monomorphic (Aac12, Aacu45, and Aacu74). Table 2 displays the parameter of informativeness for nine polymorphic SSRs that accounted for a total of 39 different alleles. The number of alleles at each locus ranged from 3–7 with an average of 4.33 alleles per locus. Three SSRs (Aac04, Aacu10, and Aac14) exhibited the highest number of alleles (N_a) and expected heterozygosity (H_e). The least number of different alleles was observed in Aacu07, Aacu12, and Aacu26. Five SSRs (Aac04, Aac14, Aac09, Aac06, and Aacu26) were highly informative ($PIC > 0.5$), while the remaining four (Aacu10, Aacu12, Aacu07, and Aacu30) were reasonably informative ($0.5 > PIC > 0.25$) (Botstein et

al. 1990) (Table 2). On the other hand, F_{st} values (Table 2) indicate that there is moderate to high levels of differentiation among populations. However, typically, a $F_{st} = 0.334$, is considered high for most tropical plants.

Table 2. Genetic diversity parameters for nine polymorphic microsatellites amplified 185 DNA samples of *A. aculeata*. NA: number of alleles, He: expected heterozygosity, Ho: observed heterozygosity, PIC: Polymorphic Information Content and F_{st} as genetic differentiation coefficients.

Cuadro 2. Parámetros de diversidad genética para nueve microsatélites polimórficos amplificados en el ADN de 185 muestras de *A. aculeata*. NA: número de alelos, He: heterocigosidad esperada, Ho: heterocigosidad observada, PIC: Contenido de Información Polimórfica y F_{st} como coeficiente de diferenciación genética.

Microsatellite	Allele range (bp)	Na	He	Ho	PIC	Fst
Aac04	199-217	7	0,47	0,38	0,71	0,33
Aac06	130-136	4	0,29	0,29	0,53	0,47
Aac09	329-335	4	0,37	0,29	0,53	0,34
Aac14	336-353	5	0,33	0,33	0,54	0,41
Aacu07	132-138	3	0,25	0,35	0,36	0,18
Aacu10	147-166	6	0,33	0,37	0,43	0,10
Aacu12	168-172	3	0,19	0,26	0,35	0,26
Aacu26	264-268	3	0,27	0,23	0,50	0,41
Aacu30	247-252	4	0,15	0,15	0,40	0,50
Average		4,33	0,29	0,29	0,48	0,33

Genetic Diversity and Structure of A. aculeata

The average number of alleles per population was 18.89. The lowest genetic diversity ($He < 0.19$) was found in samples collected in Península de Osa, Orotina and Abangares, which also had a low number of effective alleles (Table 3). La Cruz, Atenas, Tilarán, Barra Honda, Chomes, Santa Cruz, and Nicoya 0.23–0.28 registered an intermediate diversity. The sites with the highest diversity ($He > 0.30$) were observed in samples from Guanacaste, Liberia, Cañas, Caño Negro, Turrubares, Paso Real, Santa Rosa and Rincón de la Vieja. Santa Rosa and Rincón de la Vieja also displayed a high number of alleles when compared with other sites (Table 3). Regarding private alleles, we only found five private alleles of 147 bp

(Aacu10), 136 bp (Aac06), 250 bp (Aacu30), 331 bp (Aac09), and 264 bp (Aacu26), which amplified in samples of Caño Negro, Chomes, Liberia, Santa Cruz, and Tilarán, respectively (Table 3).

Table 3. Genetic diversity parameters estimated in the 18 sampling sites of *A. aculeata* in Costa Rica. He: expected heterozygosity, Ho: observed heterozygosity, Na: number of alleles, Ne: number of effective alleles, Np: number of private alleles.

Cuadro 3. Parámetros de diversidad genética estimados para 18 sitios de muestreo de *A. aculeata* en Costa Rica. He: heterocigosidad esperada, Ho: heterocigosidad observada, Na: número de alelos, Ne: número de alelos efectivos, Np: número de alelos privados.

Site of collection	Na	Ne	He	Ho	Np
Abangares	16	1,30	0,19	0,17	0
Atenas	20	1,35	0,24	0,22	0
Barra Honda	18	1,42	0,26	0,30	0
Cañas	19	1,55	0,32	0,27	0
Caño Negro	17	1,59	0,33	0,36	1
Chomes	19	1,42	0,26	0,22	1
Guanacaste	18	1,59	0,31	0,40	0
La Cruz	15	1,34	0,23	0,31	0
Liberia	25	1,62	0,32	0,28	1
Nicoya	24	1,52	0,28	0,24	0
Orotina	14	1,26	0,18	0,20	0
Paso Real	17	1,75	0,35	0,41	0
Península de Osa	19	1,23	0,17	0,09	0
Rincón de la Vieja	24	1,96	0,50	0,48	0
Santa Cruz	20	1,50	0,28	0,24	1
Santa Rosa	20	1,84	0,42	0,43	0
Tilarán	16	1,50	0,26	0,28	1
Turrubares	19	1,65	0,33	0,30	0
Average	18,89	1,52	0,29	0,29	

The AMOVA indicated that 20.82% of the total genetic variation was explained between regions (Brunca, Central, Huetar Norte, and Chorotega) and 20.39% was explained between sampling sites within regions. The least variation (3.13%) was observed between samples within each site and the greatest differentiation (55.65%) occurred within samples (Supplementary Table 2). Bayesian analysis revealed that the genetic structure among *A. aculeata* populations is best described by three distinct clusters according to the highest ΔK value (182.86) (Supplementary Figure 1). Cluster A is comprised of samples collected at Barra Honda, Nicoya, Santa Cruz, Cañas, Tilarán, Bagaces, Chomes, Orotina, Atenas, and Turubares. Cluster B comprised samples from La Cruz, Guanacaste, Santa Rosa, Liberia, Rincón de la Vieja, and Caño Negro. Cluster C grouped samples of Paso Real and Península de Osa (Figure 2). The assignment of each individual within each of the clusters in Figure 2 is supported by the percentage assignment given by the *Structure* software (Supplementary Table 3). Figure 1 also shows the clustering observed in the structure analysis.

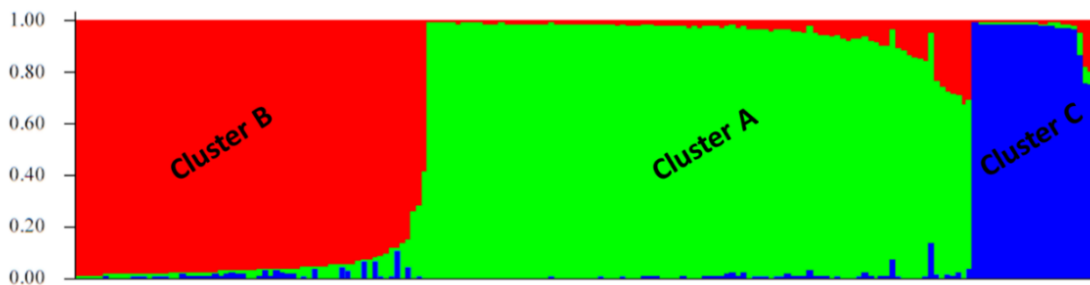


Figure 2. Population structure of the 185 samples of *A. aculeata* collected in Costa Rica. Clusters assignment of samples ($k=3$) is based on the Bayesian analysis on Structure Software using microsatellite markers genotyping.

Figura 2. Estructura poblacional de las 185 muestras de *A. aculeata* colectadas en Costa Rica. La asignación de grupos de muestras ($k = 3$) fue basado en el análisis bayesiano usando el software STRUCTURE utilizando los datos de genotipado de marcadores de microsatélites.

The clustering obtained with *Structure* (Figure 2) was the same with that found with DPCA analysis (Supplementary Figure 2). Cluster C largely separated into Península de Osa and Paso Real from the remaining sites (Supplementary Figure 2A). This result is also consistent with the UPGMA that showed a separated branch for Península de Osa and Paso Real with a high bootstrap value (Figure 3) and observed F_{st} values. Cluster C registered an F_{st} value of 0.128 compared to Clusters A and B. The differentiation of Península de Osa was moderate to high with F_{st} values between 0.104 and 0.198, with an average of 0.145. The differentiation of Paso Real was low to moderate (mean $F_{st} = 0.111$) (Supplementary Tables 4 and 5).

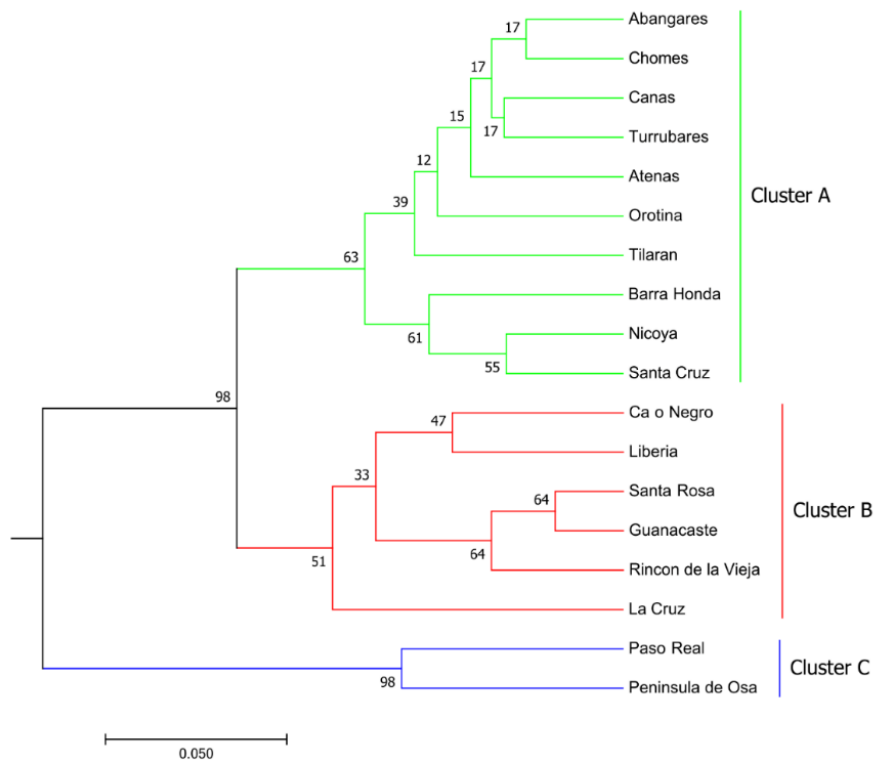


Figure 3. UPGMA dendrogram of *Acrocomia aculeata*, estimated according to Nei's genetic distance. The dendrogram reflects the same distribution of 18 sampled sites represented in figure 1. Sampled sites 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa.

Figura 3. Dendrograma UPGMA de *Acrocomia aculeata*, estimado según la distancia genética de Nei. El dendrograma refleja la misma distribución de 18 sitios muestreados representados en la figura 1. Sitios muestreados 1: La Cruz, 2: Guanacaste, 3: Santa Rosa, 4: Liberia, 5: Rincón de la Vieja, 6: Caño Negro, 7: Santa Cruz, 8: Nicoya, 9: Barra Honda, 10: Cañas, 11: Tilarán, 12: Abangares, 13: Chomes, 14: Orotina, 15: Atenas, 16: Turrubares, 17: Paso Real, 18: Península de Osa.

When Cluster C was not included in the DPCA analysis, a clearer separation of Clusters A and B was observed (Supplementary Figure 2B). The UPGMA (Figure 3) also shows a clear separation of Clusters A and B (bootstrap = 98) consistent with the DPCA and structure analyses. This clustering pattern was also supported by a moderate differentiation in both clusters with average F_{st} values of 0.065 and 0.076, respectively (Supplementary Tables 4 and 5). In Cluster A, a subcluster with samples from Barra Honda, Nicoya, and Santa Cruz was observed (Figure 3). These sampling sites are geographically close

together (Figure 1). Inside Cluster B a separation of samples from Caño Negro was resolved with the DPCA and the UPGMA; this site is relatively closer to Liberia (Figure 3). This proximity was supported with the lowest F_{st} value (0.035) between Caño Negro and Liberia (Supplementary Table 4). Also, inside Cluster B, samples collected in three national parks (Santa Rosa, Guanacaste, and Rincón de la Vieja) seem to group into a subcluster. Interestingly, in the UPGMA this subcluster also yields a bootstrap value of 64.

The correlation of genetic distances and geographical locations was explored with an isolation by distance (IBD) analysis (Supplementary Figure 3). A positive and significant correlation ($r = 0.398$; $p < 0.05$) was obtained when all 185 samples were included in the analysis. Similarly, we found a significant but weak positive correlation ($r = 0.200$; $p < 0.05$) when samples from Península de Osa and Paso Real were excluded. In both cases, genetic and geographical distances displayed dependency on each other suggesting IBD of samples collected in this study (Supplementary Figure 3B).

Discussion

Genetic Diversity of Acrocomia aculeata

The observed values of parameters such as H_e , N_e , and H_o in our study suggest that the genetic diversity of *A. aculeata* in Costa Rica is relatively low (Table 3). Studies of samples from an *ex situ* germplasm bank have shown higher genetic diversity, with almost the same set of SSRs as our work did. Lanes et al. (2015) found that with six SSRs the genetic diversity was between 0.55 and 0.74, while Mengistu et al. (2016a) reported a diversity range of 0.49–0.63 with ten SSRs. Comparison of our study results with those of *A. aculeata* studied in natural Brazilian populations support the conclusion that Costa Rica has low genetic diversity of *A. aculeata*. Higher levels of genetic diversity ($H_e > 0.50$) have been reported in natural populations in Brazil (Coelho et al. 2018; de Lima et al. 2020).

Recently, Díaz et al. (2021) observed that with SNPs markers, samples from Brazil had the highest genetic diversity when compared with samples from Central America. They also found two groups of *A. aculeata*. The northern group included samples from Costa Rica, Trinidad and Tobago, Puerto Rico, Mexico, and Colombia which had the lowest genetic diversity. Despite using microsatellite markers, our results agree with those of Díaz et al (2021).

Some researchers have proposed that the center of origin and diversification of *A. aculeata* is in South America (Lanes et al. 2015). Based on *A. aculeata* remains found from Brazil to Mexico (Morcote-Ríos and Bernal 2001), and consistent with Janzen's (1983) hypothesis that pre-Columbian Indians introduced *A. aculeata* into Costa Rica, humans may indeed have facilitated dispersion of this species from South to Central America. *A. aculeata* may also have been dispersed from the Caribbean islands, where

some samples have been collected to estimate genetic diversity (Diaz et al. 2021). The low genetic diversity found in our study supports this south to northward dispersion hypothesis.

A factor that may further explain the low genetic diversity is that this species has not been found to grow naturally in any other specific habitats other than human-disturbed areas such as paddocks, abandoned land, along roadsides, and pasturelands (Janzen 1983, Harvey et al. 2011). Despite that, *A. aculeata* is essentially the most abundant tree palm species in the Guanacaste area, where several samples were collected, its density remain very low (0.75 individuals ha⁻¹) (Harvey et al. 2011), This contrasts strongly with the 300 individuals ha⁻¹ found naturally in Brazil, where *A. aculeata* grows as aggregated clusters forming continuous forests stands (Coelho et al. 2018). In fact during the field trips, it proved difficult to locate sampling sites with groups of palm trees that met the selection criteria. This low density and fragmented distribution of *A. aculeata*, may restrict gene flow among dispersed individuals which may then explain the decreasing genetic diversity and differentiation among studied sites.

Since intense cattleranching and timber extraction in Guanacaste were principal drivers of deforestation in this region (Calvo-Alvarado et al. 2009), it is interesting to note that both Santa Rosa and Rincón de la Vieja recorded the highest genetic diversity. Both these national parks were established in Guanacaste since 1970 when cattleraising and forest clearing were prohibited. More sampling would be required to corroborate this interesting hypothesis.

Genetic Structure of Acrocomia aculeata

The structure analysis revealed three clusters (Figure 2) supported by the observed groupings in the UPGMA (Figure 3) and DPCA (Supplementary Figure 2). This structure had a moderate differentiation for Cluster A and Cluster B while for Cluster C the differentiation was higher. Furthermore, a significant IBD was observed between the three clusters (Supplementary figure 3A). The observed structure in this study may be explained by the dispersion pattern of *A. aculeata* associated with cattle ranching in Costa Rica (Harvey et al. 2011). This anthropogenic activity does not favor a continuous distribution of *A. aculeata* which probably limits gene flow.

Cluster A includes sampling sites in the lower northwestern Pacific in the Nicoya Península area, Central Pacific and, Central Region of Costa Rica. Although there are mountainous areas and river basins that delimit these regions (Gabb 2007), no differentiation was observed within sampling sites of Cluster A. Despite this, in the UPGMA (Figure 3) a subcluster containing Barra Honda, Nicoya, and Santa Cruz was identified. These three sites are located in the lower northwestern Pacific region of Nicoya Península, while

sites of Central Pacific are separated by the Guanacaste Mountain Range foothills and the Tilarán Mountain Range (Solano-Quintero 2001)

As shown in Figures 2 and 3, Cluster A and Cluster B are clearly separated. Sampling sites of Cluster B such as Liberia, Santa Rosa, Guanacaste, Rincón de la Vieja, and La Cruz are located in upper northwestern Pacific. This region is geographically located in the subregion of the Tempisque River Northeast Basin (Solano-Quintero 2001; Arroyo-Mora et al. 2005; Calvo-Alvarado et al. 2009). The first Spanish colonizers discovered that soil fertility was favorable for livestock raising in this basin. On the contrary, they found that soil characteristics and topology in the Nicoya Península were not suitable for this activity (Calvo-Alvarado et al. 2009).

During the 1800s, in fact, landowners established large farms (>5000 ha) used for livestock intensification until 1972 (Hall 1984; Calvo-Alvarado et al. 2009). In the 1970s, most forests in Guanacaste were cleared due to the accelerated deforestation and converted into pastures for livestock, rice and sugar cane crops and some timberwood forestry. Therefore, the land use of these activities favored a discontinuous occurrence of *A. aculeata* which may explain the separation of both Cluster A and B.

It is interesting to note, moreover, that the DPCA (Supplementary Figure 2) displayed a separation of Caño Negro from other sites inside this cluster. Caño Negro is located on the Caribbean slope in the lowland northern zone of Guatusos plains far enough from the rest of the sites which possibly explain genetic separation or differences between them. Finally, Cluster C (Península de Osa and Paso Real) was largely separated from the remaining sites, an observation supported by the highest differentiation coefficients (Supplementary Table 4). Even when more sampling is needed to clearly elucidate Cluster C's genetic structure, it is important to mention that the cultural use of *A. aculeata* in this region is traditionally much less than that in the northwestern Pacific region; this results in the species becoming less common and individuals more restricted or scattered across the region.

Need for Acrocomia aculeata Genetic Characterization for Uses and Conservation Uses

Our results are consistent with those of Diaz et al. (2021) that demonstrate low genetic diversity of *A. aculeata* in Costa Rica and other countries in Central America relative to genetic diversity in Brazil. Our study, therefore, serves as a useful input to conservation efforts in Costa Rica, as well as in other Mesoamerican countries. For example, in Costa Rica a germplasm bank is being established by our research group with seedlings found close to mature trees (when available) in our sampling sites. New seed and seedlings collections should consider our results to increase the number of accessions in germplasm bank collections. Once palm trees reach reproductive age, the genetic diversity impact of future outcrossings among different accessions should be estimated.

A. aculeata has tremendous potential for the biofuel and food industries as well as for other products. More understanding is needed in Mesoamerica, nonetheless, to develop commercial and conservation strategies for this species. This increased understanding depends on molecular, phenotypic, and biochemical description. For the latter to occur, traits related to oil content and the fatty acids profile require further study. Our results also contribute to future genetic diversity studies of oil-related traits in *A. aculeata* fruits in Costa Rica, similar to those studies by Da Conceição et al. (2015). A much deeper understanding of *A. aculeata* germplasm diversity will greatly enhance the species's future conservation and management.

Statements & Declarations

Funding: This research was funded by Vicerrectoría de Investigación y Extensión (VIE) of the Instituto Tecnológico de Costa Rica under the project 1510136 named “Proceso de domesticación del coyol (*Acrocomia aculeata*) como alternativa bioenergética en Costa Rica, Fase II”. The Centro Nacional de Innovaciones Biotecnológicas (CENIBiot-CeNAT) also funded the project (FP-026-2020).

Competing and Conflicts Interests: The authors declare no conflict of interest and they have no relevant financial or non-financial interests to disclose.

Author contributions: Conceptualization: E.A.S, E.A.V, I.M, N.A.Z, Methodology: E.A.V, E.A.S, V.N.C, N.A.Z, Validation: E.A.V, V.N.C, M.R.G, G.V.H, E.B.E, J.M.M, Formal Analysis: V.N.C, M.R.G, E.A.V, G.V.H, Investigation: V.N.C, P.S.C, J.A.O.U, N.A.Z, E.A.S, E.B.E, G.V.H, J.M.M, Resources: E.A.S, I.M, N.A.Z, G.V.H, E.B.E, J.M.M, Data Curation: V.N.C, M.R.G, E.A.V, Writing – Original Draft Preparation: V.N.C, E.A.V, M.R.G, Writing – Review & Editing: N.A.Z, E.A.V, E.A.S, M.R.G, I.M, J.M.M, P.S.C, J.A.O.U, Visualization: V.N.C, M.R.G, G.V.H, Supervision: E.A.V, E.A.S, I.M, Project Administration: E.A.S, E.A.V, I.M, Funding Acquisition: E.A.S, I.M

Data availability statement: The results presented in this study were obtained from the genotyping data available in the Supplementary Table 6.

Acknowledgements: The technical and administrative support of the personnel in the Instituto Tecnológico de Costa Rica and CENIBiot-CeNAT is acknowledged.

Ethical statement: The collection of samples in this research had the permission of the National Commission for the Management of Biodiversity in Costa Rica, under the permit number CM-ITCR-004-2021.

References

- Abreu AG, Priolli RHG, Azevedo-Filho JA, et al (2012) The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). *Genet Mol Biol* 35:119–121. <https://doi.org/10.1590/s1415-47572012005000002>
- Alfaro-Solís JD, Montoya-Arroyo A, Jiménez VM, et al (2020) *Acrocomia aculeata* fruits from three regions in Costa Rica: an assessment of biometric parameters, oil content and oil fatty acid composition to evaluate industrial potential. *Agrofor Syst* 94:1913–1927. <https://doi.org/10.1007/s10457-020-00511-8>
- Arroyo-Mora JP, Sánchez-Azofeifa GA, Rivard B, et al (2005) Dynamics in landscape structure and composition for the Chorotega region, Costa Rica from 1960 to 2000. *Agric Ecosyst Environ* 106:27–39. <https://doi.org/10.1016/j.agee.2004.07.002>
- Balick MJ (1990) Production of coyol wine from *Acrocomia mexicana* (Arecaceae) in Honduras. *Econ Bot* 44:84–93. <https://doi.org/10.1007/BF02861070>
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am J Hum Genet* 32:314–331
- Calvo-Alvarado J, McLennan B, Sánchez-Azofeifa A, Garvin T (2009) Deforestation and forest restoration in Guanacaste, Costa Rica: Putting conservation policies in context. *For Ecol Manage* 258:931–940. <https://doi.org/10.1016/j.foreco.2008.10.035>
- Chizmar-Fernandez C (2009) Plantas comestibles de Centroamérica, Editorial INBIO, Heredia, Costa Rica
- Ciconini G, Favaro SP, Roscoe R, et al (2013) Biometry and oil contents of *Acrocomia aculeata* fruits from the Cerrados and Pantanal biomes in Mato Grosso do Sul, Brazil. *Ind Crops Prod* 45:208–214. <https://doi.org/10.1016/j.indcrop.2012.12.008>
- Clark L V., Jasieniuk M (2011) polysat: An R package for polyploid microsatellite analysis. *Mol Ecol Resour* 11:562–566. <https://doi.org/10.1111/j.1755-0998.2011.02985.x>
- Coelho NHP, Tambarussi EV, Aguiar BI, et al (2018) Understanding genetic diversity, spatial genetic structure, and mating system through microsatellite markers for the conservation and sustainable use of *Acrocomia aculeata* (Jacq.) Lodd. *Ex Mart. Conserv Genet* 19:879–891. <https://doi.org/10.1007/s10592-018-1061-z>
- Colombo CA, Berton LHC, Diaz BG, Ferrari RA (2018) Macauba: A promising tropical palm for the

- production of vegetable oil. OCL - Oilseeds fats, Crop Lipids 25:. <https://doi.org/10.1051/ocl/2017038>
- Cornejo X, Mori SA, Aguilar R, et al (2012) Phytogeography of the trees of the Osa Península, Costa Rica. *Brittonia* 64:76–101. <https://doi.org/10.1007/s12228-011-9194-0>
- Da Conceição LDHCS, Antoniassi R, Junqueira NTV, et al (2015) Genetic diversity of macauba from natural populations of Brazil. *BMC Res Notes* 8:1–9. <https://doi.org/10.1186/s13104-015-1335-1>
- de Lima NE, Carvalho AA, Meerow AW, Manfrin MH (2018) A review of the palm genus *Acrocomia*: Neotropical green gold. *Org Divers Evol* 18:151–161. <https://doi.org/10.1007/s13127-018-0362-x>
- de Lima NE, Meerow AW, Manfrin MH (2020) Genetic structure of two *Acrocomia* ecotypes (Arecaceae) across Brazilian savannas and seasonally dry forests. *Tree Genet Genomes* 16:. <https://doi.org/10.1007/s11295-020-01446-y>
- Díaz BG, Zucchi MI, Alves-Pereira A, et al (2021) Genome-wide SNP analysis to assess the genetic population structure and diversity of *Acrocomia* species. *PLoS One* 16:1–24. <https://doi.org/10.1371/journal.pone.0241025>
- Dransfield J, Uhl NW, Asmussen CB, et al (2008) *Genera Palmarum : The Evolution and Classification of Palms*. Royal Botanic Gardens, London, UK
- Dray S, Dufour AB (2007) The ade4 package: Implementing the duality diagram for ecologists. *J Stat Softw* 22:1–20. <https://doi.org/10.18637/jss.v022.i04>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Esquivel-Mimenza H, Ibrahim M, Harvey CA, et al (2011) Dispersed trees in pasturelands of cattle farms in a tropical dry ecosystem. *Trop Subtrop Agroecosystems* 14:933–941
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Evaristo AB, Grossi JAS, Carneiro A de CO, et al (2016) Actual and putative potentials of macauba palm as feedstock for solid biofuel production from residues. *Biomass and Bioenergy* 85:18–24. <https://doi.org/10.1016/j.biombioe.2015.11.024>
- Falasca S, Ulberich A, Pitta-Alvarez S (2017) Development of agroclimatic zoning model to delimit the

- potential growing areas for macaw palm (*Acrocomia aculeata*). *Theor Appl Climatol* 129:1321–1333. <https://doi.org/10.1007/s00704-016-1850-6>
- Gabb WM (2007) On the Geology of the Republic of Costa Rica. *Rev Geológica América Cent.* 37:103-118. <https://doi.org/10.15517/rgac.v0i37.4210>
- Goudet J (2005) HIERFSTAT , a package for R to compute and test hierarchical F -statistics. *Mol Ecol Notes* 5:184–186. <https://doi.org/10.1111/j.1471-8278>
- Grayum M. (2003) Arecaceae. In: Hammel BE, Grayum MH, Herrera C, Zamora N (eds) *Manual de Plantas de Costa Rica. Gimnospermas y Monocotiledóneas (Agavaceae--Musaceae)*. Missouri Botanical Garden Press, Missouri, USA. pp 201-293
- Hall C (1984) *Costa Rica: una interpretación geo-gráfica con perspectiva historia*. Editorial Costa Rica, San Jose, Costa Rica
- Harvey CA, Villanueva C, Esquivel H, et al (2011) Conservation value of dispersed tree cover threatened by pasture management. *For Ecol Manage* 261:1664–1674. <https://doi.org/10.1016/j.foreco.2010.11.004>
- Henderson A, Galeano G, Bernal R (1995) *Field Guide to the Palms of the Americas*. Princeton University Press, New Jersey, USA
- Janzen D (1983) *Costa Rican Natural History*. The University of Chicago Press Chicago, Chicago, USA
- Jombart T (2008) Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2014:1–14. <https://doi.org/10.7717/peerj.281>
- Lanes ÉCM, Motoike SéY, Kuki KN, et al (2015) Molecular characterization and population structure of the Macaw Palm, *Acrocomia aculeata* (Arecaceae), Ex Situ Germplasm collection using microsatellites markers. *J Hered* 106:102–112. <https://doi.org/10.1093/jhered/esu073>
- Lanes ÉCM, Motoike SY, Kuki KN, et al (2016) Mating system and genetic composition of the macaw palm (*Acrocomia aculeata*): Implications for breeding and genetic conservation programs. *J Hered* 107:527–536. <https://doi.org/10.1093/jhered/esw038>
- Laviola BG, dos Santos A, Rodrigues EV, et al (2022) Structure and genetic diversity of macauba

- [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.] approached by SNP markers to assist breeding strategies. *Genet Resour Crop Evol* 69:1179–1191. <https://doi.org/10.1007/s10722-021-01295-1>
- Lescano CH, Oliveira IP, Silva LR, et al (2015) Nutrients content, characterization and oil extraction from *Acrocomia aculeata* (Jacq.) Lodd. fruits. *African J Food Sci* 9:113–119. <https://doi.org/10.5897/ajfs2014.1212>
- Lieb VM, Schex R, Esquivel P, et al (2019) Fatty acids and triacylglycerols in the mesocarp and kernel oils of maturing Costa Rican *Acrocomia aculeata* fruits. *NFS J* 14–15:6–13. <https://doi.org/10.1016/j.nfs.2019.02.002>
- Mengistu FG, Motoike SY, Caixeta ET, et al (2016a) Cross-species amplification and characterization of new microsatellite markers for the macaw palm, *Acrocomia aculeata* (Arecaceae). *Plant Genet Resour Characterisation Util* 14:163–172. <https://doi.org/10.1017/S1479262115000179>
- Mengistu FG, Motoike SY, Cruz CD (2016b) Molecular characterization and genetic diversity of the macaw palm Ex situ germplasm collection revealed by microsatellite markers. *Diversity* 8:. <https://doi.org/10.3390/d8040020>
- Morcote-Ríos G, Bernal R (2001) Remains of palms (palmae) at archaeological sites in the new world: A review. *Bot Rev* 67:309–350. <https://doi.org/10.1007/BF02858098>
- Navarro-Díaz HJ, Gonzalez SL, Irigaray B, et al (2014) Macauba oil as an alternative feedstock for biodiesel: Characterization and ester conversion by the supercritical method. *J Supercrit Fluids* 93:130–137. <https://doi.org/10.1016/j.supflu.2013.11.008>
- Nucci SM (2007) Development, characterization and analyses of the use of microsatellite markers in macaw palm population genetics. MSc. Dissertation. University of Sao Paulo.
- Nucci SM, Azevedo-Filho JA, Colombo CA, et al (2008) Development and characterization of microsatellites markers from the macaw. *Mol Ecol Resour* 8:224–226. <https://doi.org/10.1111/j.1471-8286.2007.01932.x>
- Paradis E (2010) Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Plath M, Moser C, Bailis R, et al (2016) A novel bioenergy feedstock in Latin America? Cultivation potential of *Acrocomia aculeata* under current and future climate conditions. *Biomass and Bioenergy* 91:186–195. <https://doi.org/10.1016/j.biombioe.2016.04.009>

- Pritchard J, Matthew, Donnelly P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155:945–959. <https://doi.org/10.1007/s10681-008-9788-0>
- Quirós-Guerrero L, Albertazzi F, Araya-Valverde E, et al (2019) Phenolic variation among *Chamaecrista nictitans* subspecies and varieties revealed through UPLC-ESI(-)-MS/MS chemical fingerprinting. *Metabolomics* 15:0. <https://doi.org/10.1007/s11306-019-1475-8>
- Ramos SLF, De Macêdo JLV, Lopes MTG, et al (2012) Microsatellite loci for tucumã of Amazonas (*Astrocaryum aculeatum*) and amplification in other Arecaceae. *Am J Bot* 99:2011–2013. <https://doi.org/10.3732/ajb.1100607>
- Scariot A (1998) Seed Dispersal and Predation of the Palm *Acrocomia aculeata*. *Principes* 42:5–8
- Scariot A, Lleras E, Hay JD (1995) Flowering and Fruiting Phenologies of the Palm *Acrocomia aculeata*: Patterns and Consequences. *Biotropica* 27:168. <https://doi.org/10.2307/2388992>
- Schex R, Lieb VM, Jiménez VM, et al (2018) HPLC-DAD-APCI/ESI-MSn analysis of carotenoids and α -tocopherol in Costa Rican *Acrocomia aculeata* fruits of varying maturity stages. *Food Res Int* 105:645–653. <https://doi.org/10.1016/j.foodres.2017.11.041>
- Silva LCC, Lemos R do C, Carvalho CGI, et al (2017) Genetic Diversity and Structure of Macaw Palm: Implications for Genetic Variability Sampling. *Rev Árvore* 41:. <https://doi.org/10.1590/1806-90882017000500008>
- Solano-Quintero J et al (2001) Aspectos fisiográficos aplicados a un bosquejo de regionalización geográfica climático de Costa Rica. *Top Meteor Ocean* 8:26–39
- Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, Fourth. Springer, New York, USA
- Winter DJ (2012) MMOD: An R library for the calculation of population differentiation statistics. *Mol Ecol Resour* 12:1158–1160. <https://doi.org/10.1111/j.1755-0998.2012.03174.x>
- Wright S (1965) The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *J Evol* 19:395–420. <https://doi.org/10.2307/2406450>

Conclusiones

Esta investigación representó el primer estudio en Mesoamérica en el análisis de la diversidad genética y estructura de poblaciones naturales de *Acrocomia aculeata* en función de su biogeografía, y utilizando muestras de la vertiente pacífica de Costa Rica. Al haber examinado muestras silvestres, se concluye que los resultados obtenidos representan la variación natural permanente de esta especie en el país.

Por otro lado, al utilizar un abordaje metodológico que integró la técnica PCR multiplex junto con electroforesis capilar de fluorescencia, se optimizó las condiciones para la amplificación de doce regiones microsátélites en *A. aculeata*. Lo anterior, permitió la evaluación simultánea de al menos cinco regiones microsátélite diferentes en una misma muestra. Esto, representó un importante ahorro en el uso de reactivos y un menor tiempo de duración de la investigación, aunado a la alta precisión y fiabilidad de la lectura automatizada por medio de electroforesis capilar.

De esta manera, se encontró que tres de los doce microsátélites evaluados en diferentes poblaciones del país fueron monomórficos (Aac12, Aacu45 y Aacu74), por lo que se descartaron debido a su poca información. En contraste, Aac04, Aac14, Aac09, Aac06 y Aacu26 fueron los cinco microsátélites con mayor índice de polimorfismo de acuerdo con los parámetros de diversidad identificados. Además, esta especie mostró una alta transferibilidad con los marcadores diseñados para *Astrocarium aculeatum*.

Asimismo, se observó una alta riqueza alélica poblacional, dado el número de alelos encontrados por población. Así, la mayor diversidad genética se registró en las poblaciones de Santa Rosa y Rincón de la Vieja, con altos niveles de heterocigosidad esperada y número de alelos efectivos. En contraste, las localidades de Península de Osa, Orotina y Abangares exhibieron la diversidad genética más baja, debido al bajo número de alelos efectivos y heterocigosidad esperada. Por otro lado, las únicas poblaciones que presentaron alelos privados fueron Caño Negro, Chomes, Liberia, Santa Cruz y Tilarán.

Costa Rica presentó una diversidad genética considerablemente menor que lo reportado en Suramérica previamente. Esto pudo estar relacionado con la baja densidad de las poblaciones

encontradas principalmente en sitios perturbados por la urbanización y el uso ganadero de la tierra. Lo anterior, induce la autopolinización o el cruzamiento con individuos emparentados. Además, la limitada diseminación de polen y la restringida dispersión de semillas por parte del ganado limitan el flujo de genes y provocan una deficiencia de heterocigotos explicada por el alto nivel de consanguinidad.

Por otra parte, la larga historia de uso humano, la deforestación pasada, la alta frecuencia de incendios forestales intencionales para despejar áreas tropicales secas o muy secas, así como las barreras geográficas y el aislamiento por distancia, pueden haber contribuido a la alta diferenciación genética y estructura espacial observada en este estudio.

Recomendaciones

En relación con este estudio y tomando en cuenta la optimización de los resultados obtenidos, es recomendable la homogenización en la cantidad de muestras obtenidas por cada localidad, lo mismo que la mejora en el registro individual de las coordenadas de muestreo. Esto, reduciría la posibilidad de sesgo debido a errores de muestreo en las poblaciones de estudio.

Ahora bien, dado que este estudio representa el primer abordaje en Mesoamérica encargado de comparar la diversidad genética de *Acrocomia aculeata*, se recomienda su replicación tomando en cuenta muestras de toda esta área. Lo anterior, permitiría la evaluación en conjunto de aspectos relacionados con la riqueza genética de la zona, con el fin de combinar esfuerzos para la conservación de la especie tanto en Costa Rica como otros países del área. De esta manera, se podría considerar la integración de accesiones genéticamente diferentes en el banco de germoplasma establecido en nuestro país. Esto permitiría evaluar la posibilidad de entrecruzamiento y mejoramiento genético de la especie para potenciar su cultivo para fines industriales.

Finalmente, debido a la importancia económica que representa la obtención de aceite de esta especie, es recomendable reproducir este estudio utilizando marcadores genéticos que identifiquen regiones codificantes como los Polimorfismos de Nucleótido Único (SNPs, por sus siglas en inglés). Esto permitiría la evaluación y correlación de caracteres genéticos específicos relacionados con el contenido de aceite y el perfil de ácidos grasos en los frutos y semillas de *A. aculeata* en poblaciones de diferentes áreas geográficas. De esta manera, se podría elegir con fundamentos genéticos, a las poblaciones con mejores rendimientos de producción para su cultivo.

Anexos

Supplementary table 1. Geographic and altitudinal information of the sampling sites of collection.

Tabla complementaria 1. Información geográfica y altitudinal de los sitios de muestreo.

Sampling Site	Number of palm trees collected	Region	Latitude N	Longitude W	Altitude (masl)
Abangares	16	Chorotega	10.29	-84.97	162
Atenas	11	Central	9.94	-84.39	418
Barra Honda	10	Chorotega	10.17	-85.43	176
Cañas	7	Chorotega	10.44	-85.06	160
Caño Negro	11	Huetar Norte	10.88	-84.78	37
Chomes	9	Chorotega	10.12	-84.90	103
Guanacaste	9	Chorotega	10.95	-85.51	457
La Cruz	4	Chorotega	11.05	-85.62	209
Liberia	20	Chorotega	10.71	-85.60	162
Nicoya	16	Chorotega	10.20	-85.50	178
Orotina	5	Central	9.83	-84.59	166
Paso Real	3	Brunca	9.00	-83.22	140
Peninsula de Osa	19	Brunca	8.65	-83.46	16
Rincon de la Vieja	10	Chorotega	10.77	-85.27	768
Santa Cruz	13	Chorotega	10.22	-85.59	90
Santa Rosa	8	Chorotega	10.91	-85.43	852
Tilaran	4	Chorotega	10.49	-84.95	749
Turrubares	10	Central	9.89	-84.22	127
Total of samples	185				

Supplementary table 2. Analysis of Molecular Variance (AMOVA) of 185 *Acrocomia aculeata* accessions stratified using the hierarchy of regions and sites of collection.

Tabla complementaria 2. Análisis de varianza molecular (AMOVA) de 185 accesiones de *Acrocomia aculeata* estratificadas utilizando la jerarquía de regiones y sitios de recolección.

	Df	% Variation	Sum Sq	P-value
Variations between region	3	20,823,912	242,021	0.01*
Variations between site within region	14	20,389,295	278,082	0.01*
Variations between samples within site	167	3,132,185	4,509,348	0.06*
Variations within samples	185	55,654,608	449	0.01*
Total	369	100	14,200	-

Supplementary table 3. Inferred ancestry of individuals with the percentage assignment of each sample within each of the cluster in Figure 2.

Tabla complementaria 3. Ascendencia inferida de los individuos con la asignación porcentual de pertenencia de cada muestra dentro de cada uno de los conglomerados de la Figura 2.

Sampling site	Cluster A	Cluster B	Cluster C
Abangares	0,980	0,015	0,006
Abangares	0,917	0,070	0,013
Abangares	0,984	0,010	0,006
Abangares	0,983	0,010	0,007
Abangares	0,989	0,006	0,005
Abangares	0,961	0,029	0,010
Abangares	0,973	0,020	0,007
Abangares	0,973	0,019	0,007
Abangares	0,704	0,282	0,014
Abangares	0,941	0,050	0,010
Abangares	0,982	0,012	0,006
Abangares	0,707	0,278	0,015
Abangares	0,944	0,047	0,010
Abangares	0,938	0,054	0,008
Abangares	0,986	0,009	0,005
Abangares	0,950	0,040	0,010

Atenas	0,983	0,011	0,006
Atenas	0,883	0,106	0,011
Atenas	0,939	0,051	0,010
Atenas	0,983	0,012	0,004
Atenas	0,689	0,289	0,022
Atenas	0,982	0,014	0,004
Atenas	0,814	0,048	0,138
Atenas	0,985	0,012	0,004
Atenas	0,935	0,059	0,007
Atenas	0,891	0,035	0,074
Atenas	0,987	0,008	0,005
Cañas	0,739	0,256	0,005
Cañas	0,415	0,581	0,004
Cañas	0,840	0,151	0,009
Cañas	0,963	0,024	0,013
Cañas	0,954	0,030	0,016
Cañas	0,849	0,145	0,007
Cañas	0,895	0,097	0,008
Chomes	0,861	0,132	0,007
Chomes	0,933	0,056	0,011
Chomes	0,957	0,034	0,009
Chomes	0,956	0,031	0,014
Chomes	0,912	0,081	0,007
Chomes	0,969	0,021	0,009
Chomes	0,653	0,305	0,043
Chomes	0,913	0,062	0,025
Chomes	0,985	0,011	0,004
Orotina	0,989	0,008	0,004
Orotina	0,986	0,009	0,005
Orotina	0,976	0,016	0,008
Orotina	0,989	0,007	0,004
Orotina	0,966	0,020	0,014
Tilarán	0,670	0,325	0,005
Tilarán	0,894	0,098	0,008
Tilarán	0,959	0,032	0,009
Tilarán	0,923	0,073	0,004
Turrubares	0,978	0,016	0,006
Turrubares	0,972	0,025	0,003
Turrubares	0,972	0,021	0,007
Turrubares	0,742	0,251	0,006
Turrubares	0,989	0,007	0,004
Turrubares	0,976	0,019	0,005
Turrubares	0,881	0,112	0,007

Turrubares	0,983	0,011	0,006
Turrubares	0,964	0,022	0,013
Turrubares	0,975	0,020	0,005
Barra Honda	0,964	0,015	0,021
Barra Honda	0,984	0,009	0,008
Barra Honda	0,991	0,006	0,004
Barra Honda	0,962	0,016	0,023
Barra Honda	0,983	0,011	0,006
Barra Honda	0,989	0,007	0,005
Barra Honda	0,984	0,007	0,008
Barra Honda	0,987	0,009	0,005
Barra Honda	0,944	0,021	0,036
Barra Honda	0,964	0,010	0,026
Nicoya	0,990	0,006	0,004
Nicoya	0,987	0,007	0,007
Nicoya	0,748	0,232	0,021
Nicoya	0,957	0,038	0,005
Nicoya	0,983	0,008	0,009
Nicoya	0,978	0,018	0,004
Nicoya	0,989	0,006	0,005
Nicoya	0,987	0,008	0,005
Nicoya	0,991	0,006	0,004
Nicoya	0,969	0,018	0,013
Nicoya	0,978	0,009	0,013
Nicoya	0,913	0,076	0,012
Nicoya	0,979	0,008	0,013
Nicoya	0,927	0,069	0,003
Nicoya	0,984	0,009	0,007
Nicoya	0,971	0,024	0,005
Santa Cruz	0,988	0,008	0,004
Santa Cruz	0,990	0,006	0,004
Santa Cruz	0,986	0,009	0,005
Santa Cruz	0,922	0,071	0,006
Santa Cruz	0,855	0,140	0,005
Santa Cruz	0,955	0,034	0,011
Santa Cruz	0,976	0,019	0,005
Santa Cruz	0,984	0,012	0,004
Santa Cruz	0,984	0,012	0,004
Santa Cruz	0,979	0,016	0,005
Santa Cruz	0,962	0,031	0,006
Santa Cruz	0,977	0,014	0,009
Santa Cruz	0,954	0,036	0,009
Caño Negro	0,007	0,955	0,037

Caño Negro	0,006	0,962	0,032
Caño Negro	0,010	0,947	0,043
Caño Negro	0,016	0,975	0,010
Caño Negro	0,008	0,963	0,029
Caño Negro	0,008	0,977	0,015
Caño Negro	0,025	0,960	0,016
Caño Negro	0,011	0,879	0,110
Caño Negro	0,010	0,980	0,011
Caño Negro	0,014	0,975	0,011
Caño Negro	0,008	0,982	0,009
La Cruz	0,017	0,970	0,013
La Cruz	0,016	0,976	0,009
La Cruz	0,010	0,983	0,007
La Cruz	0,012	0,980	0,008
Liberia	0,007	0,985	0,008
Liberia	0,015	0,966	0,019
Liberia	0,011	0,968	0,021
Liberia	0,009	0,972	0,020
Liberia	0,110	0,881	0,009
Liberia	0,009	0,986	0,005
Liberia	0,007	0,983	0,010
Liberia	0,007	0,983	0,010
Liberia	0,015	0,978	0,006
Liberia	0,049	0,944	0,007
Liberia	0,047	0,949	0,004
Liberia	0,044	0,951	0,005
Liberia	0,069	0,925	0,006
Liberia	0,042	0,953	0,006
Liberia	0,135	0,858	0,007
Liberia	0,026	0,964	0,010
Liberia	0,015	0,962	0,024
Liberia	0,015	0,966	0,019
Liberia	0,014	0,915	0,071
Liberia	0,036	0,958	0,006
Santa Rosa	0,007	0,987	0,006
Santa Rosa	0,272	0,718	0,011
Santa Rosa	0,015	0,977	0,007
Santa Rosa	0,028	0,963	0,009
Santa Rosa	0,007	0,989	0,005
Santa Rosa	0,059	0,934	0,006
Santa Rosa	0,013	0,982	0,006
Santa Rosa	0,017	0,978	0,005
Guanacaste	0,009	0,985	0,006

Guanacaste	0,035	0,956	0,009
Guanacaste	0,017	0,975	0,008
Guanacaste	0,010	0,984	0,006
Guanacaste	0,028	0,966	0,006
Guanacaste	0,030	0,965	0,005
Guanacaste	0,037	0,956	0,007
Guanacaste	0,088	0,906	0,006
Guanacaste	0,023	0,962	0,015
Rincón de la Vieja	0,011	0,985	0,004
Rincón de la Vieja	0,081	0,909	0,010
Rincón de la Vieja	0,106	0,847	0,047
Rincón de la Vieja	0,006	0,927	0,067
Rincón de la Vieja	0,007	0,981	0,011
Rincón de la Vieja	0,048	0,948	0,005
Rincón de la Vieja	0,020	0,946	0,034
Rincón de la Vieja	0,005	0,991	0,004
Rincón de la Vieja	0,011	0,967	0,022
Rincón de la Vieja	0,006	0,988	0,005
Paso Real	0,049	0,198	0,752
Paso Real	0,082	0,049	0,869
Paso Real	0,064	0,176	0,760
Península de Osa	0,006	0,006	0,988
Península de Osa	0,006	0,009	0,985
Península de Osa	0,007	0,008	0,985
Península de Osa	0,006	0,006	0,988
Península de Osa	0,008	0,006	0,986
Península de Osa	0,019	0,005	0,976
Península de Osa	0,004	0,004	0,992
Península de Osa	0,004	0,006	0,991
Península de Osa	0,004	0,004	0,992
Península de Osa	0,003	0,004	0,993
Península de Osa	0,004	0,004	0,992
Península de Osa	0,006	0,007	0,987
Península de Osa	0,006	0,007	0,987
Península de Osa	0,016	0,005	0,979
Península de Osa	0,004	0,004	0,992
Península de Osa	0,010	0,019	0,971
Península de Osa	0,004	0,005	0,991
Península de Osa	0,011	0,013	0,976
Península de Osa	0,011	0,013	0,976

Supplementary table 4. Differentiation coefficients between all the sampling sites estimated by pairwise FST according to Nei.

Tabla complementaria 4. Coeficientes de diferenciación entre todos los sitios de muestreo estimados usando FST por pares según Nei.

	Abangares	Atenas	Barra Honda	Cañas	Caño Negro	Chomes	Guanacaste	La Cruz	Liberia	Nicoya	Orotina	Paso Real	Península de Osa	Rincón de la Vieja	Santa Cruz	Santa Rosa	Tilarán	Turrubares
Abangares	0	0,0461	0,0631	0,0149	0,0844	0,0152	0,0546	0,1098	0,0589	0,038	0,0428	0,0986	0,1377	0,082	0,0207	0,0621	0,0149	0,0182
Atenas	0,0461	0	0,0895	0,0115	0,0861	0,0224	0,0848	0,1606	0,0632	0,0833	0,059	0,1193	0,14	0,0892	0,0642	0,0874	0,072	0,047
Barra Honda	0,0631	0,0895	0	0,0616	0,131	0,081	0,1119	0,1191	0,0994	0,0188	0,0565	0,1559	0,177	0,0957	0,0427	0,1017	0,071	0,0339
Cañas	0,0149	0,0115	0,0616	0	0,062	0,0123	0,0423	0,102	0,0262	0,0443	0,0343	0,095	0,1291	0,0482	0,0295	0,0425	0,0239	0,0097
Caño Negro	0,0844	0,0861	0,131	0,062	0	0,0627	0,0637	0,1102	0,0348	0,1181	0,1238	0,0747	0,1035	0,0728	0,1132	0,0697	0,0834	0,0952
Chomes	0,0152	0,0224	0,081	0,0123	0,0627	0	0,0571	0,1235	0,0477	0,0671	0,0642	0,0767	0,1254	0,0751	0,0471	0,0622	0,0329	0,0338
Guanacaste	0,0546	0,0848	0,1119	0,0423	0,0637	0,0571	0	0,07	0,0328	0,0907	0,0902	0,1003	0,1503	0,0252	0,0712	0,0027	0,0352	0,0502
La Cruz	0,1098	0,1606	0,1191	0,102	0,1102	0,1235	0,07	0	0,0805	0,1043	0,1334	0,1466	0,1976	0,0461	0,1102	0,069	0,0619	0,0858
Liberia	0,0589	0,0632	0,0994	0,0262	0,0348	0,0477	0,0328	0,0805	0	0,0885	0,0813	0,108	0,1252	0,032	0,0803	0,0289	0,0447	0,0508
Nicoya	0,038	0,0833	0,0188	0,0443	0,1181	0,0671	0,0907	0,1043	0,0885	0	0,0475	0,1331	0,1622	0,0908	0,0103	0,0823	0,0399	0,0246
Orotina	0,0428	0,059	0,0565	0,0343	0,1238	0,0642	0,0902	0,1334	0,0813	0,0475	0	0,1645	0,1915	0,093	0,0442	0,0825	0,0405	0,0109
Paso Real	0,0986	0,1193	0,1559	0,095	0,0747	0,0767	0,1003	0,1466	0,108	0,1331	0,1645	0	0,0598	0,1019	0,1213	0,0988	0,1113	0,1205
Península de Osa	0,1377	0,14	0,177	0,1291	0,1035	0,1254	0,1503	0,1976	0,1252	0,1622	0,1915	0,0598	0	0,1411	0,1527	0,1412	0,1649	0,1582
Rincón de la Vieja	0,082	0,0892	0,0957	0,0482	0,0728	0,0751	0,0252	0,0461	0,032	0,0908	0,093	0,1019	0,1411	0	0,079	0,0024	0,0541	0,0525
Santa Cruz	0,0207	0,0642	0,0427	0,0295	0,1132	0,0471	0,0712	0,1102	0,0803	0,0103	0,0442	0,1213	0,1527	0,079	0	0,0584	0,0216	0,0234
Santa Rosa	0,0621	0,0874	0,1017	0,0425	0,0697	0,0622	0,0027	0,069	0,0289	0,0823	0,0825	0,0988	0,1412	0,0024	0,0584	0	0,0409	0,0457
Tilarán	0,0149	0,072	0,071	0,0239	0,0834	0,0329	0,0352	0,0619	0,0447	0,0399	0,0405	0,1113	0,1649	0,0541	0,0216	0,0409	0	0,0184
Turrubares	0,0182	0,047	0,0339	0,0097	0,0952	0,0338	0,0502	0,0858	0,0508	0,0246	0,0109	0,1205	0,1582	0,0525	0,0234	0,0457	0,0184	0

Supplementary table 5. Summary by average of Fst coefficients between each sampling site with the remaining sites.

Tabla complementaria 5. Resumen promedio de coeficientes Fst entre cada sitio de muestreo con los sitios restantes.

Cluster	Sampling site	Average
A	Cañas	0,046
	Turrubares	0,052
	Tilarán	0,055
	Abangares	0,057
	Chomes	0,059
	Santa Cruz	0,064
	Nicoya	0,073
	Atenas	0,078
	Orotina	0,080
	Barra Honda	0,089
Average of A		0,065
B	Santa Rosa	0,063
	Liberia	0,064
	Guanacaste	0,067
	Rincón de la Vieja	0,069
	Caño Negro	0,088
	La Cruz	0,108
Average of B		0,076
C	Paso Real	0,111
	Península de Osa	0,145
Average of C		0,128

Supplementary table 6. Genotyping of 185 samples used in this study with 9 polymorphic microsatellite DNA markers. The numbers represents the size in base pairs of the alleles amplified by PCR and detected by capillary electrophoresis coupled with fluorescence.

Tabla complementaria 6. Genotipado de 185 muestras utilizadas en este estudio con 9 marcadores de ADN tipo microsatélites polimórficos. Los números representan el tamaño en pares de bases de los alelos amplificados por PCR y detectados por electroforesis capilar acoplada con fluorescencia.

Sampling site	Aac04	Aac04	Aac06	Aac06	Aac14	Aac14	Aacu26	Aacu26	Aacu10	Aacu10	Aacu07	Aacu07	Aacu12	Aacu12	Aacu30	Aacu30	Aac09	Aac09
Abangares	199	199	132	132	347	347	268	268	159	159	134	134	170	170	249	249	329	333
Abangares	199	199	132	134	347	347	266	266	159	159	134	138	170	172	249	249	333	333
Abangares	199	199	132	134	347	347	268	268	159	159	134	134	170	170	249	249	329	333
Abangares	199	206	132	134	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	134	347	347	268	268	157	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	132	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Abangares	206	206	132	134	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	206	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	134	347	347	266	268	159	159	134	134	170	170	249	249	329	333
Abangares	199	199	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	199	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Abangares	199	199	132	132	347	347	266	268	159	159	134	134	170	170	249	249	329	333
Abangares	206	206	132	134	347	347	266	268	159	159	134	134	170	170	249	249	329	333
Abangares	199	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Atenas	199	199	132	132	347	347	268	268	157	159	134	134	170	170	249	249	329	333
Atenas	199	199	132	132	347	347	266	266	157	159	134	134	170	170	249	249	329	329
Atenas	199	199	132	132	347	347	266	266	159	159	134	134	170	170	249	249	329	329
Atenas	199	201	132	132	347	347	268	268	159	159	134	134	170	170	249	249	329	329
Atenas	199	199	132	132	347	353	266	268	159	159	134	134	170	170	249	249	329	329
Atenas	199	206	132	132	347	347	268	268	159	162	134	138	170	170	249	249	329	329
Atenas	199	199	130	130	347	347	268	268	159	159	132	134	170	170	249	249	329	329
Atenas	199	201	132	132	347	347	268	268	159	159	132	134	170	170	249	249	329	329

Atenas	199	199	132	132	347	347	266	268	159	159	132	134	170	170	249	249	329	329
Atenas	199	199	130	132	347	353	268	268	159	162	132	134	170	170	249	249	329	329
Atenas	199	199	132	134	347	347	268	268	159	162	132	134	170	170	249	249	329	333
Barra Honda	199	216	132	134	347	353	266	268	159	159	134	134	170	170	249	249	333	333
Barra Honda	216	216	134	134	347	353	268	268	159	159	134	134	170	170	249	249	329	329
Barra Honda	216	216	134	134	347	347	268	268	159	159	134	134	170	170	249	249	329	333
Barra Honda	201	216	130	134	347	347	268	268	157	159	134	138	170	170	249	249	329	333
Barra Honda	199	216	132	134	347	347	266	268	157	159	134	134	170	170	249	249	329	333
Barra Honda	216	216	134	134	347	347	266	268	157	159	134	134	170	170	249	249	329	333
Barra Honda	216	216	134	134	347	353	268	268	159	159	134	134	170	170	249	249	329	333
Barra Honda	216	216	132	134	347	347	266	268	159	159	134	134	170	170	249	249	329	333
Barra Honda	216	216	130	132	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Barra Honda	216	216	130	134	347	347	268	268	159	159	134	134	170	170	249	249	333	333
Caño Negro	199	214	130	132	353	353	266	266	147	159	134	134	168	170	249	249	329	333
Caño Negro	199	199	130	130	349	349	266	266	159	160	134	134	168	170	249	249	329	329
Caño Negro	199	199	130	132	349	353	266	266	159	159	134	134	168	170	249	249	329	333
Caño Negro	199	199	132	132	349	349	266	266	147	159	134	134	170	170	249	249	333	333
Caño Negro	199	199	130	132	349	353	266	266	147	159	134	134	168	170	249	249	329	333
Caño Negro	199	199	132	132	353	353	266	266	147	147	134	134	168	170	249	249	329	333
Caño Negro	199	214	130	132	349	353	266	268	147	159	134	134	168	170	249	249	329	329
Caño Negro	199	199	130	132	353	353	266	266	147	159	134	134	168	170	249	249	333	333
Caño Negro	199	199	132	132	349	353	266	266	147	159	134	134	168	170	249	249	329	333
Caño Negro	214	214	132	132	349	353	266	266	159	159	134	134	170	170	249	249	333	333
Caño Negro	199	199	132	132	349	353	266	266	147	159	134	134	168	170	249	249	329	329
Cañas	206	214	132	132	347	347	266	266	159	159	134	134	168	170	249	249	329	333
Cañas	214	214	132	134	347	347	266	268	159	160	134	134	170	170	249	249	329	333
Cañas	199	199	132	132	347	347	266	266	159	162	134	134	170	170	249	249	329	329
Cañas	199	199	132	132	347	353	268	268	159	162	134	134	170	170	249	249	329	329
Cañas	199	206	132	132	347	353	266	268	159	162	134	138	170	170	249	249	333	333
Cañas	199	199	132	132	347	347	268	268	159	160	134	134	170	170	249	249	333	333
Cañas	199	199	132	132	347	347	266	268	159	159	134	134	170	170	249	249	329	329
Chomes	199	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	329	329
Chomes	199	199	136	134	347	347	266	266	159	162	134	134	170	170	249	249	329	329

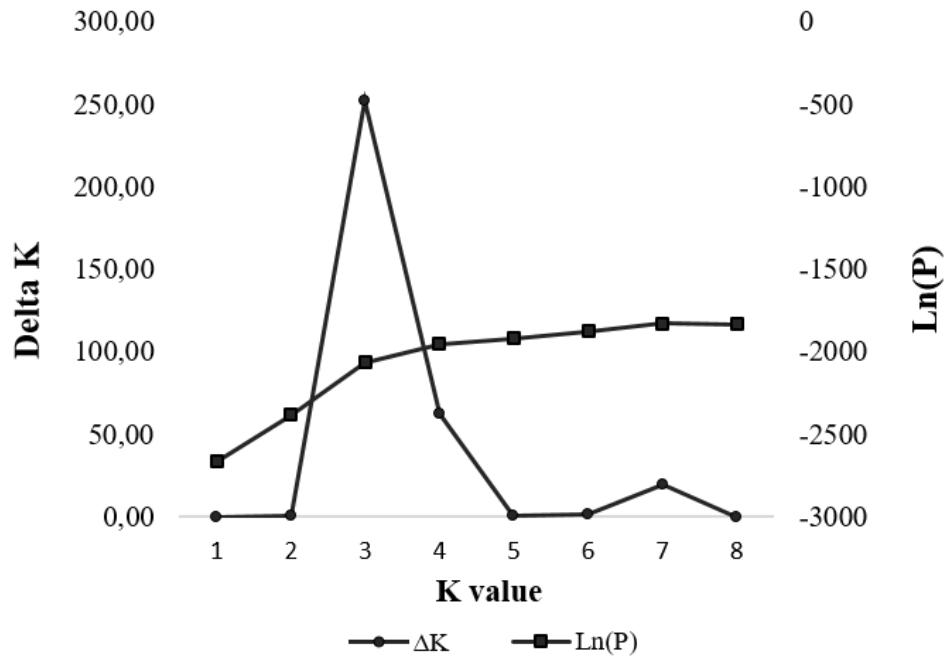
Chomes	199	199	132	132	347	347	266	268	157	159	134	134	170	170	249	249	329	333
Chomes	199	199	132	132	347	347	266	266	159	159	134	138	170	170	249	249	333	333
Chomes	199	206	132	132	347	347	266	266	157	159	134	134	170	170	249	249	329	329
Chomes	199	199	132	132	347	347	266	268	157	159	134	134	170	170	249	249	333	333
Chomes	199	199	132	132	336	347	266	266	157	159	134	134	170	170	249	249	333	333
Chomes	199	199	132	132	336	347	266	268	157	159	134	134	170	170	249	249	333	333
Chomes	199	216	132	132	347	347	268	268	157	159	134	134	170	170	249	249	329	329
La Cruz	214	214	130	132	347	349	266	266	157	159	134	138	168	170	249	249	333	333
La Cruz	214	214	130	132	347	347	266	266	159	159	134	134	168	170	249	249	329	333
La Cruz	214	214	130	132	347	347	266	266	159	159	134	138	168	168	249	249	333	333
La Cruz	214	214	130	132	347	347	266	266	159	159	134	134	168	168	249	249	333	333
Liberia	214	214	132	132	353	353	266	266	159	159	134	138	168	170	249	249	329	329
Liberia	199	214	132	132	353	353	266	266	157	159	134	134	168	170	249	249	329	329
Liberia	199	199	132	132	353	353	266	266	159	159	134	134	168	170	249	249	329	329
Liberia	199	199	132	132	353	353	266	266	159	159	134	134	168	170	249	250	329	329
Liberia	199	216	132	132	349	353	266	266	159	159	134	134	168	170	249	249	329	329
Liberia	214	214	132	132	349	349	266	266	159	159	134	134	170	170	249	249	329	329
Liberia	199	214	132	132	353	353	266	266	159	160	134	134	168	170	249	249	329	329
Liberia	199	214	132	132	353	353	266	266	159	160	134	134	168	170	249	249	329	329
Liberia	199	214	132	132	347	349	266	266	159	160	134	134	170	170	249	249	329	333
Liberia	199	199	132	132	347	347	266	266	159	160	134	134	170	170	249	249	329	329
Liberia	214	214	132	134	347	349	266	266	160	160	134	134	168	170	249	249	329	333
Liberia	199	214	132	132	347	347	266	268	159	160	134	134	168	170	250	250	329	333
Liberia	199	214	132	132	347	347	266	266	159	159	134	134	170	170	249	249	329	329
Liberia	214	214	132	132	347	347	266	266	159	159	134	134	170	170	249	249	329	333
Liberia	201	214	132	132	347	349	266	266	159	159	134	134	170	170	249	250	333	333
Liberia	199	214	132	132	347	353	266	266	159	159	134	134	170	170	249	249	329	329
Liberia	199	199	132	132	336	353	266	266	159	159	134	134	168	170	249	249	329	329
Liberia	199	214	132	132	336	353	266	266	159	159	134	134	168	170	249	249	329	333
Liberia	214	214	130	132	353	353	266	266	159	159	134	134	170	170	249	249	333	333
Liberia	199	199	132	132	347	349	266	266	159	160	132	134	168	170	249	250	329	333
Nicoya	216	216	132	134	347	347	268	268	159	159	134	134	170	170	249	249	333	333

Nicoya	199	199	134	134	347	347	268	268	159	159	134	134	170	170	249	249	333	333
Nicoya	199	206	132	132	336	347	266	266	159	159	134	138	170	170	249	249	333	333
Nicoya	214	216	132	134	347	347	266	268	157	159	134	134	170	170	249	249	333	333
Nicoya	216	216	132	134	347	353	268	268	159	159	134	134	170	170	249	249	333	333
Nicoya	214	216	134	134	347	347	268	268	157	159	134	134	170	170	249	249	333	333
Nicoya	201	217	134	134	347	347	266	268	159	162	134	134	170	170	249	249	333	333
Nicoya	217	217	134	134	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Nicoya	217	217	134	134	347	347	268	268	159	159	134	134	170	172	249	249	333	333
Nicoya	217	217	132	134	347	353	266	268	159	159	134	134	170	170	249	249	333	333
Nicoya	199	217	134	134	347	353	268	268	159	159	134	134	170	170	249	249	333	333
Nicoya	217	217	132	132	347	353	266	268	159	159	134	134	170	170	249	249	329	333
Nicoya	199	217	134	134	347	353	268	268	159	159	134	134	170	170	249	249	333	333
Nicoya	217	217	132	134	347	347	268	268	159	162	134	134	168	170	249	249	329	333
Nicoya	199	217	134	134	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Nicoya	217	217	132	132	347	347	266	268	159	159	134	134	170	170	249	249	329	333
Orotina	206	206	132	132	347	353	268	268	159	159	134	134	170	170	249	249	329	333
Orotina	199	206	132	132	347	347	268	268	157	159	134	134	170	170	249	249	329	333
Orotina	206	206	132	132	347	353	268	268	159	159	134	134	170	170	249	249	329	329
Orotina	206	206	132	132	347	347	268	268	159	159	134	134	170	170	249	249	329	333
Orotina	199	206	132	132	347	353	268	268	159	159	134	138	170	170	249	249	333	333
Santa Rosa	199	214	132	132	349	349	266	266	159	159	134	138	168	170	249	252	333	333
Santa Rosa	199	199	132	132	347	347	266	266	159	159	134	138	170	170	249	249	329	333
Santa Rosa	205	217	130	132	347	347	266	266	159	159	134	138	168	170	249	252	329	333
Santa Rosa	199	217	132	132	347	353	266	266	159	159	134	138	170	170	252	252	329	333
Santa Rosa	199	214	132	132	347	349	266	266	160	160	134	134	168	170	252	252	333	333
Santa Rosa	217	217	132	132	347	353	266	266	159	159	134	138	170	170	249	252	329	329
Santa Rosa	214	217	132	132	347	353	266	266	159	159	134	138	168	170	249	252	329	333
Santa Rosa	199	214	132	132	347	347	266	266	159	160	134	138	170	170	249	249	329	329
Guanacaste	214	214	132	132	347	353	266	266	159	159	134	138	168	170	249	252	329	333
Guanacaste	199	199	132	132	347	349	266	266	159	159	134	138	170	170	249	249	329	333
Guanacaste	199	214	132	132	347	353	266	266	159	159	134	138	170	170	249	249	329	329
Guanacaste	214	214	132	132	347	353	266	266	159	159	134	138	168	170	249	249	329	333

Guanacaste	199	205	132	132	347	347	266	266	159	159	134	138	168	170	249	249	333	333
Guanacaste	199	217	132	132	347	347	266	266	159	159	134	138	168	170	249	252	329	333
Guanacaste	199	199	132	132	347	347	266	266	159	159	134	138	168	170	249	249	329	333
Guanacaste	199	217	132	132	347	347	266	266	159	159	134	138	168	170	249	249	333	333
Guanacaste	199	199	132	132	347	353	266	266	159	159	134	138	168	170	249	249	333	333
Rincón de la Vieja	214	214	132	132	347	347	266	266	159	159	134	138	168	170	249	249	329	329
Rincón de la Vieja	199	214	130	132	347	347	266	268	159	160	134	134	168	170	249	249	329	333
Rincón de la Vieja	199	214	130	134	345	349	266	268	157	160	134	138	168	170	252	252	329	333
Rincón de la Vieja	199	214	132	132	347	349	266	266	159	160	134	138	168	170	249	252	329	335
Rincón de la Vieja	199	199	130	132	347	349	266	266	159	159	134	138	168	170	252	252	329	329
Rincón de la Vieja	217	217	132	132	347	349	266	266	159	159	134	138	170	170	249	252	329	333
Rincón de la Vieja	214	217	130	130	347	349	266	266	159	159	134	138	170	170	249	249	333	333
Rincón de la Vieja	205	214	132	132	349	349	266	266	159	160	134	138	170	170	252	252	329	329
Rincón de la Vieja	199	214	130	130	347	347	266	266	159	159	134	138	168	170	249	249	329	329
Rincón de la Vieja	214	214	130	132	347	347	266	266	160	160	134	138	168	170	249	249	329	329
Paso Real	199	199	130	132	345	353	266	266	159	159	134	138	170	170	247	249	333	333
Paso Real	199	199	130	130	336	347	266	266	166	166	134	134	170	170	247	249	333	333
Paso Real	199	199	130	132	345	353	266	266	157	159	134	138	170	170	247	249	333	333
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	247	247	333	333
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	138	170	170	247	247	333	333
Península de Osa	199	199	130	130	336	353	266	266	159	159	134	134	170	170	247	247	333	333
Península de Osa	199	199	130	130	353	353	266	266	157	159	134	134	170	170	247	247	333	333
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	247	247	333	333
Península de Osa	199	199	130	130	353	353	266	268	157	159	134	134	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	266	159	166	134	134	170	170	247	247	335	335

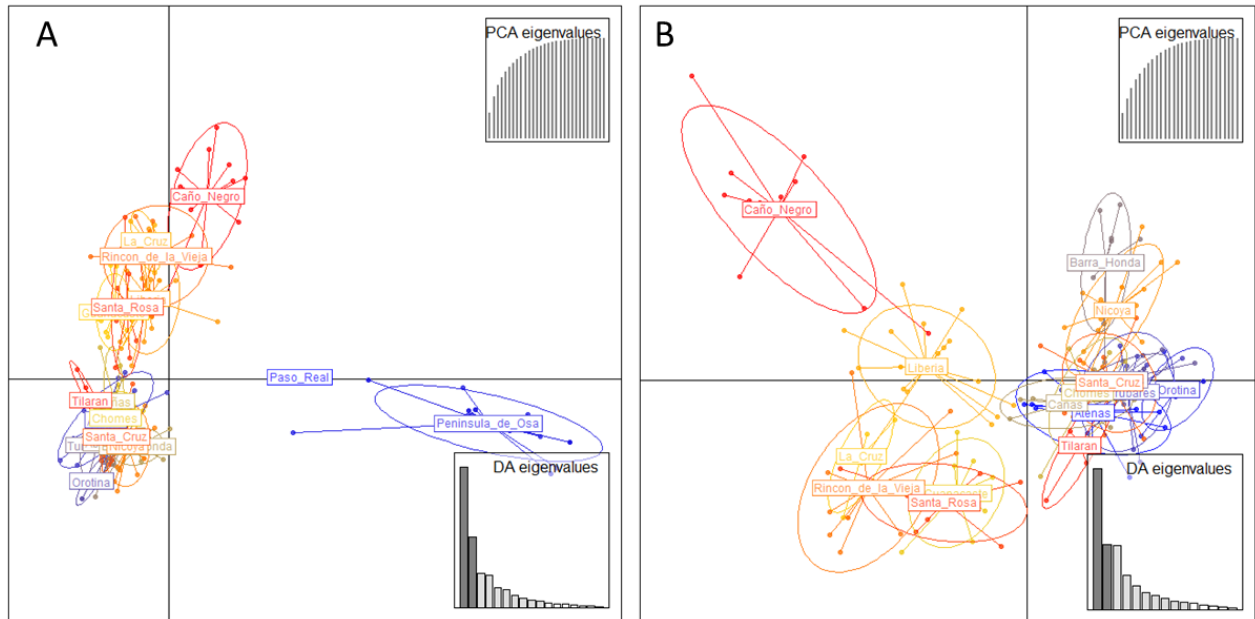
Península de Osa	199	199	130	130	353	353	266	266	159	166	134	138	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	266	166	166	134	134	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	268	159	159	134	134	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	266	157	159	134	134	170	170	247	249	335	335
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	247	249	335	335
Península de Osa	199	199	130	130	353	353	266	268	159	159	134	134	170	170	247	247	335	335
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	138	170	170	247	247	335	335
Península de Osa	199	217	130	132	347	353	266	268	159	159	134	134	170	170	247	249	333	333
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	249	249	335	335
Península de Osa	199	199	130	130	353	353	266	266	159	159	134	134	170	170	249	249	335	335
Santa Cruz	199	216	132	132	347	347	268	268	159	159	132	134	170	170	249	249	331	331
Santa Cruz	216	216	134	134	347	347	268	268	159	159	134	138	170	170	249	249	333	333
Santa Cruz	206	216	132	134	347	347	266	268	159	162	134	138	170	170	249	249	333	333
Santa Cruz	199	217	132	132	347	347	266	268	159	159	134	134	170	170	249	252	333	333
Santa Cruz	199	217	132	134	347	349	268	268	159	159	134	134	170	170	249	252	333	333
Santa Cruz	199	199	132	132	347	349	266	268	159	159	134	134	170	170	249	249	333	333
Santa Cruz	199	217	132	132	347	347	268	268	159	159	134	134	170	170	249	252	333	333
Santa Cruz	217	217	132	132	347	347	268	268	159	159	134	134	170	170	249	252	333	333
Santa Cruz	217	217	132	134	347	347	268	268	159	159	134	134	170	170	249	252	333	333
Santa Cruz	217	217	132	132	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Santa Cruz	217	217	132	132	347	347	266	266	159	159	134	134	170	170	249	252	333	333
Santa Cruz	199	199	132	134	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Santa Cruz	199	199	132	132	347	347	266	268	159	159	134	134	170	170	249	249	333	333
Tilarán	199	214	132	132	347	347	268	268	159	159	134	134	168	170	249	249	333	333
Tilarán	199	214	132	132	347	347	264	266	159	159	134	134	170	170	249	249	333	333
Tilarán	199	206	132	132	347	347	266	266	157	159	134	134	170	170	249	249	333	333
Tilarán	206	214	132	132	347	347	266	268	159	159	134	138	168	170	249	249	333	333

Turrubares	199	206	132	132	347	347	266	268	157	159	134	134	170	170	249	249	329	333
Turrubares	201	214	134	134	347	347	268	268	157	159	134	134	170	170	249	249	329	329
Turrubares	201	206	132	132	347	347	266	266	159	159	134	134	170	170	249	249	333	333
Turrubares	214	214	132	132	347	353	268	268	159	162	134	134	170	170	249	249	329	333
Turrubares	206	206	132	134	347	347	268	268	159	159	134	134	170	170	249	249	329	329
Turrubares	206	206	132	132	347	347	266	268	159	159	134	138	170	170	249	249	329	333
Turrubares	201	201	132	132	347	347	266	266	159	159	134	138	170	170	249	249	329	333
Turrubares	199	201	132	132	347	347	266	268	159	162	134	134	170	170	249	249	333	333
Turrubares	201	201	132	134	347	353	266	268	159	159	134	138	170	170	249	249	333	333
Turrubares	201	206	132	132	347	347	266	268	159	159	134	138	170	170	249	249	329	333



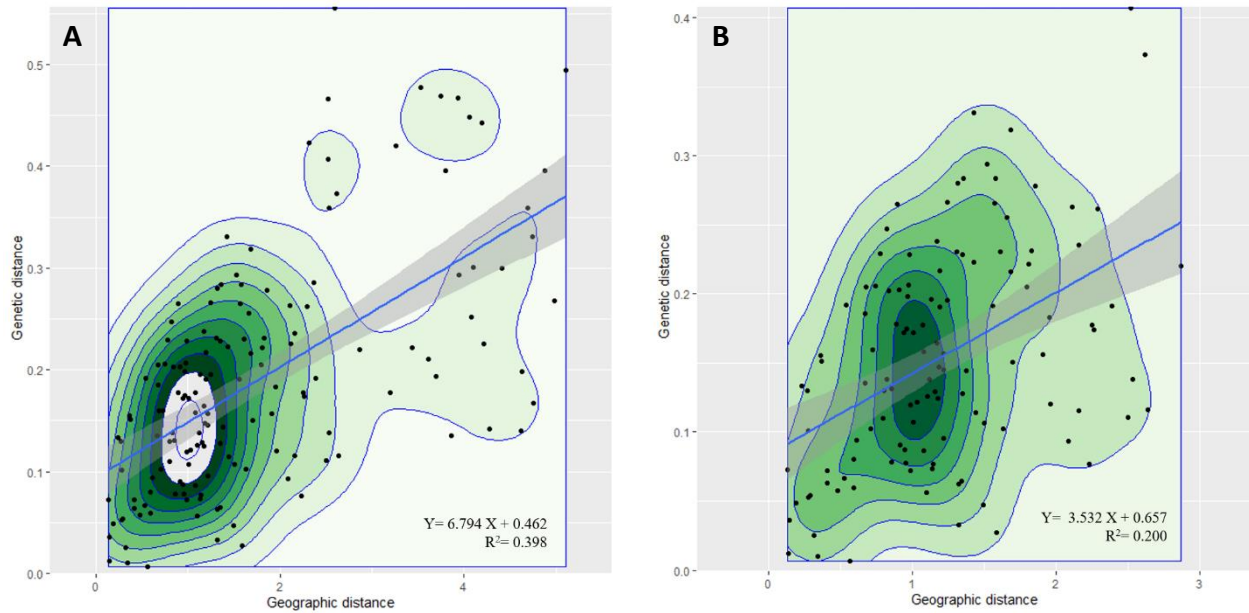
Supplementary figure 1. Population structure of 185 accessions of *Acrocomia aculeata*. Mean posterior probability values $\text{Ln}(P)$ (D) and estimated values of Δk across 8 repetitions by Bayesian analysis on Structure Software.

Figura complementaria 1. Estructura de la población de 185 accesiones de *Acrocomia aculeata*. Valores medios de probabilidad posterior $\text{Ln}(P)$ (D) y valores estimados de Δk en 8 repeticiones mediante análisis bayesiano en el software de estructura.



Supplementary figure 2. Dispersion graph of germplasm of *A. aculeata* using DPCA (Discriminant Principal Component Analysis). A: DPCA with all 185 samples which displays a clear separation of samples from Península de Osa and Paso Real (PopA) from other sampling sites of the study B: DPCA showing a clearer separation of PopB and PopC, without samples from Península de Osa and Paso Real. Eigenvalues represent the ratio of the variance between groups over the variance within groups for each discriminant function.

Figura complementaria 2. Gráfico de dispersión de germoplasma de *A. aculeata* utilizando DPCA (Análisis Discriminante de Componentes Principales). A: DPCA con las 185 muestras que muestra una clara separación de las muestras de Península de Osa y Paso Real (PopA) de otros sitios de muestreo del estudio. B: DPCA que muestra una separación más clara de PopB y PopC, excluyendo las muestras de Península de Osa y Paso Real. Los valores propios representan la relación de la varianza entre grupos sobre la varianza dentro de los grupos para cada función discriminante.



Supplementary figure 3: Isolation by distance (IBD) plot based on Nei's genetic distance matrix and geographical coordinates of sampling sites of *Acrocomia aculeata* for this study. A: IBD plot with all 185 samples of this study.

B: IBD plot without samples from Brunca region (Península de Osa and Paso Real). The intensity of colors represent the density of distribution of the samples according to genetic and geographic correlation.

Figura complementaria 3: Gráfico de aislamiento por distancia (IBD) basado en la matriz de distancia genética de Nei y las coordenadas geográficas de los sitios de muestreo de *Acrocomia aculeata* para este estudio. A: gráfico de IBD con las 185 muestras de este estudio. B: Parcela IBD excluyendo las muestras de la región Brunca (Península de Osa y Paso Real). La intensidad de los colores representa la densidad de distribución de las muestras según correlación genética y geográfica.