UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
FACULDADE DE ENGENHARIA
CAMPUS DE ILHA SOLTEIRA

LORENA FRIGINI MORO CAPO

TIMBER TRACKING OF *Jacaranda copaia* (Aubl.) D. Don. (BIGNONACEAE) FROM AMAZON FOREST USING DNA FINGERPRINT

Ilha Solteira
2022

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Ilha Solteira

POST-GRADUATION PROGRAM IN AGRONOMY

LORENA FRIGINI MORO CAPO

TIMBER TRACKING OF *Jacaranda copaia* (Aubl.) D. Don. (BIGNONACEAE) FROM AMAZON FOREST USING DNA FINGERPRINT

Doctoral thesis presented to the Engineering Faculty of Ilha Solteira – UNESP as part of the requirements to obtain the title of PhD in Agronomy.

Supervisor
Prof. Dr. Alexandre Magno Sebbenn

Ilha Solteira
2022

FICHA CATALOGRÁFICA
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

C245t

Capo, Lorena Frigini Moro.
Timber tracking of Jacaranda copaia (Aubl.) D. Don. (bignonaceae) from amazon forest using dna fingerprint / Lorena Frigini Moro Capo. -- Ilha Solteira: [s.n.], 2022
69 f. : il.

Tese (doutorado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Sistemas de Produção, 2022

Orientador: Alexandre Magno Sebbenn
Inclui bibliografia

1. DNA fingerprinting. 2. Jacaranda copaia. 3. Marker development. 4. Single nucleotide polymorphism. 5. Timber tracking. 6. Timber trade.

Raiane da Silva Santos
Supervisora Técnica de Seção
Seção Técnica de Referência, Atendimento ao usuário e Documentação
Diretoria Técnica de Biblioteca e Documentação
CRB/8 - 9999

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: TIMBER TRACKING OF Jacaranda copaia (Aubl.) D. Don. (BIGNONACEAE) FROM AMAZON FOREST USING DNA FINGERPRINT

AUTORA: LORENA FRIGINI MORO CAPO
ORIENTADOR: ALEXANDRE MAGNO SEBBENN

Aprovada como parte das exigências para obtenção do Título de Doutora em Agronomia, área: Sistemas de Produção pela Comissão Examinadora:

Pesquisador Dr. ALEXANDRE MAGNO SEBBENN (Participaçao Virtual)
Genética e Melhoramento Florestal / Instituto Florestal de São Paulo

Prof. Dr. PAULO CEZAR CERESINI (Participaçao Virtual)
Departamento de Fitossanidade, Engenharia Rural e Solos / Faculdade de Engenharia de Ilha Solteira - UNESP

Dra. PATRÍCIA FERREIRA ALVES (Participaçao Virtual)
Departamento de Fitotecnia, Tecnologia de Alimentos e Sócio Economia / Faculdade de Engenharia de Ilha Solteira

Pesquisador Dr. MIGUEL LUIZ MENEZES FREITAS (Participaçao Virtual)
Melhoramento Florestal / Instituto Florestal de São Paulo

Prof. Dr. EDSON JOSÉ VIDAL DA SILVA (Participaçao Virtual)
Ciências Florestais / Universidade de São Paulo

Ilha Solteira, 30 de setembro de 2022

DEDICATION

I dedicate this thesis to my father, Jorge Luiz Moro Capo (in memory) who supported me and made it all possible.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Alexandre Magno Sebbenn, with all my heart, because without him I would not have been able to finish my PhD. He was much more than a mentor, he was a friend and reached out to me both at work and in life. Alexandre was very professional, but above all, he was a friendly shoulder and knew how to give me space when I needed it, but he also pulled me back to reality when necessary.

I am grateful for all the support and support given by the German Federal Ministry of Food and Agriculture within the framework of the "Large-scale project on genetic verification of wood". To the Genomic and Sequencing Facility of Bordeaux, where the genotyping was carried out, (grants from the Conseil Regional d'Aquitaine n° 20030304002FA and 20040305003FA, European Union, FEDER n° 2003227 and Investissements d'avenir, n° ANR-10-EQPX-16 - 01 and CEBA: ANR-10-LABX-25-01 to UMR EcoFoG). ICMBIO permits No. 47960, 49369, CGen registration No. A16E4DA (Brazil), MAYA/VMABCCGDF/DGBAP/MEG No. 0280/2016 (Bolivia), R.D. No. 001-2016-SERNANP-DGANP, R.D. No. 001A-2015 -SERFOR-DGGSPFFS-DGSPF and Contract No. 001-2016-SERFOR-DGGSPFFS-DGSPF (Peru) were awarded for the collection, transport and genetic analysis of samples. We are grateful to colleagues at IIAP, INRA French Guiana, Station Cirad Paracou, Office National des Forêts and Institut Thünen for their assistance during field and laboratory work. In Bolivia, the samples were identified by the Museo de Historia Natural Noel Kempff Mercado. Thank you very much.

To my family and friends, I am deeply grateful for all the emotional support and strength they have given me. Every moment I had a demonstration of affection and pride for getting where I am.

And, last but not least, I thank my partner who literally fought with me daily so that I would continue and not let this dream end without my realizing it.

RESUMO

A floresta Amazônica e outras florestas tropicais ao redor do mundo estão atualmente sendo intensivamente desmatadas para e abertura de áreas para uso na agropecuaria intensiva e extração de madeira, o que em geral ocorre ilegalmente em áreas protegidas como reservas e áreas indigenas, resultando em problemas ecológicos, ambientais e econômicos. Com o objetivo de parar o desmatamento e a comercialização de madeira de corte ilegal de florestas tropicais, novas leis foram introduzidas em muitos países. Aqui investigamos a utilidade da impressão digital de DNA de marcadores SNPs nucleares e citoplasmáticos para rastrear a origem da madeira extraida e comercializada da árvore neotropical *Jacaranda copaia*. Amostras de 832 indivíduos de 43 populações da Bolívia, Brasil, Guiana Francesa e Peru foram utilizadas para investigar o poder de marcadores SNPs, sendo 113 nucleares (nSNPs), 11 cloroplastidiais (CpSNPs) e quatro mitocondriais (MtSNPs) para determinar corretamente o país, população e região dentro do Brasil e Peru de origem. A diferenciação genética (G'_{ST}) entre todas as populações, entre populações de diferentes países e entre regiões dentro dos países foi alta (0,506-0,698), especialmente para locos CpMtSNP (> 0,9), mostrando um forte padrão genético de isolamento por distância entre populações, o que é favoravel à determinação correta do local de origem de amostras de mandeira de indivíduos de *J. copaia*. Para testes de auto-atribuição foi possivel determinar corretamente, com 100% de precisão, o país, população e região de origem de todas as amostras quando foram utilizados todos os locos SNPs ou apenas os nSNPs. Os resultados mostraram que o uso de todos os marcadores SNPs ou nSNPs é uma ferramenta precisa e útil para alfândegas e policias nacionais e internacionais averiguarem se o local de extração declarado na documentação de exportação de madeira de *J. copaia* da Floresta Amazônica tem origem legal ou ilegal.

Palavras-chave: comércio de madeira; desenvolvimento de marcadores moleculares; impressão digital de DNA; *Jacaranda copaia*; polimorfismo de nucleotídeo único; rastreamento de madeira.

ABSTRACT

Amazon and other tropical forest are actually subject to strong deforestation, generally originated from illegal logging, resulting in ecological, environmental and economic problems. Aiming stop deforestation and timber commercialization of illegal logging of tropical forest, new laws has been introduced in many countries. Here we investigated the utility of DNA fingerprinting of nuclear and cytoplasmatic SNPs markers to timber tracking the intensive logged and commercialized of the Amazonian Neotropical tree *Jacaranda copaia*. Samples of 832 individuals from 43 populations from French Guiana, Brazil, Peru, and Bolivia were used to investigate the power of 113 nuclear SNPs, 11 CpSNPs and four MtSNPs loci to determine the country, population and region within Brazil and Peru origin. The genetic differentiation (G'_{ST}) among all populations, contries, and regions within coutries was generaly high (0.506-0.698), specialy for CpMtSNP (> 0.9) loci, and there is a strong isolation by distance pathern among populations, favoring the group or individual samples tracking to correct site. For self-assignment tests, we were able to 100% correct determine country, population and region site origin of all samples using all SNPs and nSNPs. Our results show that the use of all SNP or nSNP markers are suitable to correct determination of country and population site of *J. copaia* timber origin and very useful tool for customs and local and international policies.

Keywords: DNA fingerprinting, *Jacaranda copaia*; marker development; single nucleotide polymorphism; timber tracking; timber trade.

SUMMARY

1 INTRODUCTION

Much of the world's commercial logging from natural forests is illegally sourced, and even the legal source has come from unsustainable logging. This is especially true for tropical forests, where most of the world's plant and animal biodiversity lives. Both illegal and legal harvesting contribute to the loss of biodiversity, diminishing the potential for human resources to find medicinal and future sources of wood. Illegal logging is also an economic problem for the legal market, as extremely low-priced timber competes with legal logging, where costs are higher (GRAY, 2002; DEGEN *et al*., 2013). Because of this, many international casualties were established in countries around the world to prevent illegal timber imports.

To trace the origin of wood, many methods have been tested, such as chemical differences between wood species (PAREDES-VILLANUEVA *et al*., 2018), transformation cyclotron resonance mass spectrometry (DEKLERCK *et al*., 2017), wood anatomy and DNA fingerprinting (JOLIVET; DEGEN 2012; DEGEN *et al*., 2013; LOWE *et al*., 2016; CHAVES *et al*., 2018; SEBBENN *et al*., 2019). Results from different methods showed strong potential for determining species, country, and site origin, in particular using DNA fingerprinting (LOWE *et al*., 2016; CHAVES *et al*., 2018)

The current work investigated the use of the DNA fingerprint method to track the intensive and high value wood of the Neotropical pioneer tree *Jacaranda copaia* (Aubl.) D. Don. (Bignonaceae). The wood is light and used for furniture (LOUREIRO *et al*., 1979). This is a fast-growing species, with an average annual increase in diameter at breast height (DBH) of 2.05 cm and height of 1.98 m, with a great capacity for regeneration in gabs (SAMPAIO *et al*., 1989). The trees have a straight stem, reaching 106 cm in DBH and 45 m in height (VINSON *et al*., 2015a). The species occurs from northern to western South America, from Belize to Bolivia and Brazil, French Guiana, Peru (GENTRY, 1992). In Brazil, the species is found in the states of Acre, Amapá, Amazonas, Maranhão, Rondônia, Roraima, Mato Grosso and Pará and populations generally have more than one tree per hectare (VINSON *et al*., 2015a). The species is hermaphroditic, self-incompatible and about 40 species of bees, butterflies and hummingbird wasps were detected as potential pollinators, although Euglossa spp. and Centris spp. bees were detected as the main pollinators in the Tapajós National Forest, Brazil (MAUES *et al*., 2008). The fruits can have up to 250 seeds and the winged seeds are dispersed by the wind (MAUES *et al*., 2008).

In this work we investigated the potential use of chloroplast, mitochondrial and nuclear SNP markers to create a genetic reference database to track the location and country of origin of *J. copaiba*.

2 LITERATURE REVIEW

2.1 *Jacaranda* genus

The Bignoniaceae family comprises approximately 82 genera and 827 species, distributed throughout the Pantropical region (LOHMANN; ULLOA, 2011). According to Lohmann *et al*. (2020), in Brazil it has 34 genera and 420 species, dispersed in biomes such as the Amazon, Atlantic Forest, Cerrado, Caatinga, Pantanal and Pampa. Among the species found in the country is the *Jacaranda copaia*, found mainly in the northern region, in the states of Acre, Amapá, Amazonas, Maranhão, Rondônia, Roraima, Mato Grosso and Pará (VINSON *et al*., 2015b). Worldwide, the species occurs from northern to western South America, from Belize to Bolivia and Brazil, French Guiana and Peru (GENTRY, 1992).

Most *Jacaranda* species occur in savannas and dry environments around the Amazon rainforest, with the exception of *J. copaia* (Aubl.) D. Don. (Bignonaceae), which occurs throughout the Amazon rainforest, generally found with a population density > 1 tree/ha (VINSON *et al*., 2015b). *Jacaranda copaia* has self-incompatible hermaphroditic flowers pollinated by bees, beetles, butterflies, flies, hummingbirds and wasps, and its winged seeds are wind-dispersed seeds (MAUES *et al*., 2008). The trees have straight stems, fast-growing (2.05 cm in diameter at breast height (DBH) and 1.98 m in height per year), and can reach 106 cm in dap and 45 m in height (SAMPAIO *et al*., 1989; VINSON *et al*., 2015a). The species is tolerant to different types of soils, which may explain the wide geographic distribution of the species.

Considered a medium-sized tree, it has been found solitary, with bees being its main pollinators in the Tapajós National Forest, Brazil (MAUES *et al*., 2008). The fruits are encapsulated, with a rounded base and apex and can have up to 250 seeds, which are light and winged, dispersed by the wind (MAUES *et al*., 2008).

In the country, it is cited for its timber use, since the wood is light and used for furniture, toy manufacture, ceiling linings, aeronautical construction, insulating and floating material (LOUREIRO *et al*., 1979). In addition, it is used in smaller materials, such as picture frames and dividers. Its lightness added to the factor of great ease in seedling propagation, and its rapid and uniform germination (OLIVEIRA *et al*., 2013), its wood is also used in the laminate and

plywood industry (TONINI *et al*., 2008). The pulp and paper industry also benefits from its properties in the manufacture of cellulosic pulp (LORENZI, 2002).

The species is also important in the medicinal and ornamental area, in urban afforestation (MILLIKEN, 1998; ALBERT; MILLIKEN, 2009; PORTO *et al*., 2013; PAUMGARTTEN *et al*., 2018; SILVA *et al*., 2018; TATAGIBA *et al*., 2019; BRITO; PONTES, 2021). Due to its rapid growth and straight stem, it is widely used in commercial plantations and, also for this reason, the species is intensively exploited for wood production in the Amazon rainforest.

2.2 Treaties and regulations concerning timber trade

In 1975, Brazil signed a treaty in the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) so that, in this way, there is greater efficiency in the regularization of the commercialization of species of fauna and flora, preventing the extinction and the threat to international trade. In addition to Brazil, 182 other countries are in agreement with the proposal and, as a way of improving the applicability of this treaty, each producing country has its responsibility, adjusting the most effective techniques and methods to guarantee the exploitation of the species in a way that does not harm the population of this.

According to CITES, there are around 5,950 animal species and 32,800 plant species protected from rampant and harmful exploitation worldwide. In Brazil, the body responsible for controlling and implementing procedures regarding the evaluation and issuance of export and import licenses is the Brazilian Institute of Environment and Renewable Natural Resources – IBAMA. In addition to the institute, there are other bodies responsible for controlling the commercialization of fauna and flora, such as the Rio de Janeiro Botanical Garden and ICMBIO.

In 2022, a Normative Instruction was implemented (No. 8, of March 25, 2022), which "Establishes the procedures for the authorization of exportation of wood products and by-products of native species from natural or planted forests, aiming to complement, relatively to control the export of native wood loads within the scope of IBAMA, Normative Instruction No. 21, of December 24, 2014, Normative Instruction No. 17, of December 1, 2021, and Ordinance No. 8, of January 3 of 2022" and an Ordinance (No. 8, of January 3, 2022, amended by Ordinance No. 46, of January 6, 2022), which "Establishes, within the scope of Ibama, the Single Consent Platform of Brazil - PAU Brasil for use in the activities of foreign trade involving products and by-products of biodiversity."

Three lists are found in the country, according to Decree nº 3.607, of September 21, 2000, indicating the species that are already considered endangered, and those that are listed in the second list, which includes species not currently threatened, but that run serious risks, unless the traffic of commercialization is strictly observed and the laws and regulations are strictly followed. In the third annex are the species that the Brazilian agency itself has listed, so that there is greater care in the exploration, which is restricted or even blocked for better export control. The Jacaranda species is listed in the first annex (CITES, 2021).

South American lumber companies have been under pressure to provide more information, correctly issuing the name of the species and origin of the exported wood, due to highly strict policies imposed by the European Union (European Timber Regulations - EUTR), in force since the year of 2013. Following the same line of demanding complete and detailed documentation, the US, in 2008, launched the Lacey Act, which prohibits the entry of plants or products obtained illegally.

As common practices of illegal wood export are the theft of species in forests, both private and public, or in areas demarcated as protected. In addition to theft, there is also the practice of tax evasion, forgery of documents, violation of forestry regulations and smuggling. In 2014, the World Bank estimated that, of the entire supply of wood, around 25% was obtained illegally, causing losses of up to US$ 20 billion.

2.3 Illegal Timber Control

International companies invest very high amounts in wood industries in Brazil due to the high variety of forests contained here, having great exploitation for the extraction of wood and raw material for construction materials. The opening of inland roads and highways is linked to the felling of trees and the movement of foreign trade, since the easy road access contributes to the entry and exit of the forest. These already explored areas bring greater opportunities for residents to open space for pastures and cultivation of agronomic crops, causing greater deforestation (WWF).

Even though this is a typically Brazilian scenario, the illegal timber trade is not only strong in Brazil. Countries such as Papua New Guinea and Liberia had around 85% of logging operations being done illegally (STARK; PANG CHEUNG, 2006; LAWSON; MACFAUL, 2010; WIT *et al*., 2010). The biggest problem about the illegality of the trade is in the documentation of the origin of the wood, having falsified certificates and documentation and with missing or false information (CERUTTI; LESCUYER, 2011; KISHOR; LESCUYER, 2012; LESCUYER *et al*., 2014).

Taking into account the fragility regarding the control of entry and exit of wood, scientific methods were studied to help in the identification and certification of these woods. Physical methods such as dendrology and wood anatomy are based on studying the growth rings of trees, used to identify the age of the species, and on the analysis of wood structure respectively (METCALFE; CHALK, 1950; PIGOZZO, 2011). In addition to estimating its lifespan, the observation of the rings provide researchers with information about the environmental and climatic conditions in which the species was exposed during its growth. Widely used in forensic studies, physical methods help in estimating the moment of tree tipping and species differentiation in terms of ring standardization. When it comes to wood anatomy, it allows quick identification of the genus and is widely used in the field. This method is not frequently used at the species level and its results may be influenced by the environment and the genetics of the species (DORMONTT *et al*., 2015).

Chemical analysis methods are also used to facilitate legislation and surveillance of timber extraction and trade. Mass Spectrometry, Near Infrared Spectrometry and Stable Isotopes are also used to verify the geographic origin of each wood taken from its location. Near Infrared spectrometry has been used to analyze properties such as estimating the chemical composition of the tree: lignin, cellulose and extractives (TERDWONGWORAKUL *et al*., 2005), wood density and also its geographical origin (SANDAK *et al*., 2011; LI *et al*., 2019) As for Stable Isotopes, these are added during the synthesis of phytochemical compounds in the plant, revealing the availability of the species in that environment, facilitating information about the place of origin with high precision.

In addition to physical and chemical methods, genetic methods for geographic identification have been studied more closely. The DNA Barcode technique has been extensively explored and used to determine the species and trace its route, as well as products originating from that plant. Genetic Attribution also helps to determine the geographic origin of the species and, the most current of genetic methods, the DNA fingerprint is used to differentiate individuals.

Like SNP markers, the multilocus approach and barcode technique are based on nuclear microsatellites and chloroplasts. These are methods used to determine and differentiate species. Although they were initially used in animals, it is widely used in the forestry area (DEGEN; FLADUNG, 2008).

2.3.1 Genetic Assignment

Native forests have a genetic structure based on their place of origin, with their conformation based on the place where they are located or at a regional level. Based on this spatial structure, the study of population genetics has been used to determine the origin of an individual. The spatial genetic structure is given by the non-random distribution of alleles and genotypes, with individuals with greater physical proximity being those with a greater degree of kinship, that is, having a greater degree of related genes, while individuals that are more distant have lower degree of kinship (DEGEN; SEBBENN, 2014).

The crossing between trees, the form of vegetative propagation and the dispersion of seeds and pollen directly influence the breeding system of the species. In addition, population density and climatic factors such as wind speed and direction dictate the distance to the spatial genetic structure, so it is important to use molecular markers in order to identify the geographic origin and perform ecological niche modeling. For the use of markers, attention must be paid to the quality of the genetic reference database of the species and the accuracy of the distance measurement between the geographic origin and the sampling of the population of the database (DEGEN; FLADUNG, 2008).

For better characterization of species, tools such as molecular markers began to stand out in science for their quality, rapid response power and for being able to beused on a large scale. In addition, they accurately estimate allele frequency change, allele loss or fixation, and the genetic diversity of populations (SANTOS *et al*., 2020). Microsatellite markers have been used in genetic conservation studies because they are codominant and multi-allele, based on the DNA polymerase chain reaction (OLIVEIRA *et al*., 2013)..

2.3.2 DNA Fingerprinting

The DNA fingerprinting methodology has great potential for verifying forest identity at an individual level and also for tracking the origin of the wood (CHAVES *et al*., 2018; MEYER-SAND *et al*., 2017; HONÓRIO CORONADO *et al*., 2019; PAREDES-VILLANUEVA *et al*., 2020). In humans, this is the main technique used in forensics (JOBLING; GILL, 2004). This method is based on the analysis of the DNA of each microorganism present in the wood, thus it is believed that they are specific to a particular forest area (EL SHEIKHA *et al*., 2013).

The scarcity of accessible forensic methods to contribute to the detection and tracking of illegal wood ends up making it difficult to determine the geographic origin of the tree.

Therefore, having this type of analysis is of paramount importance to combat theft and illegal trade, directly affecting international trade (MEIER-AUGENSTEIN, 2019).

A set of genetic markers, such as microsatellites and SNPs, is needed to develop a DNA fingerprint. This happens because a unique pattern must be set up for each observed individual. In the area of traceability, in order to support the control of timber smuggling, it is necessary to have a database with extensive reference sample information, allowing a high degree of reliability in the results obtained (DORMONTT *et al*., 2015).

For the commercialization of wood, there is a set of laws in force, making smuggling difficult. One of the requirements for export is the chain of custody control document, the CoC, which is based on information described in a paper on the identification of the cut tree, with the aim of tracking it from the forestry conception to the trader, but this document is not proof of corruption and forgery (LOWE *et al*., 2010). The DNA fingerprint cannot be falsified, so it is used to verify the integrity of the chain of custody. This is possible because samples are collected throughout the chain of custody, to detect whether illegal wood was introduced in the shipment and, if so, at what time this operation was carried out. This technique has been widely used because it is a highly efficient and incorruptible method, since the DNA cannot be altered, and that tree always has the same genetic information.

2.4 REFERENCES

ALBERT, B.; MILLIKEN, W. Urihi a: a terra-floresta Yanomami. São Paulo: Institut de Recherche Pour le Développement, 2009.

BRITO, A. F. S.; PONTES, A. N. Metabólitos secundários de plantas medicinais usadas em garrafadas populares comercializadas em feiras livres de Belém, Pará, Brasil. **Enciclopédia Biosfera**, [*s. l.*], v. 18, n. 36, p. 207-222, 2021.

CERUTTI PO, LESCUYER G. The domestic market for small-scale chainsaw milling in Cameroon: present situation, opportunities and challenges. **Center for International Forestry Research (CIFOR)**, Bogor, Indonesia, 2011.

CHAVES, C. L., DEGEN B., PAKULL B., MADER M., HONORIO E., RUAS P., TYSKLIND N., SEBBENN A. M. Assessing the ability of chloroplast and nuclear DNA gene markers to verify the geographic origin of Jatoba (*Hymenaea courbaril* L.) timber. **Journal of Heredity**, Oxford, v. 109, p. 543–552, 2018.

CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). **What is CITES?** Available on: https://www.cites.org/eng/disc/what.php. Consulted on: 6th July 2021.

DEGEN, B.; FLADUNG, M. Use of DNA-markers for tracing illegal logging. In: Degen, B. Proceedings of the international workshop “Fingerprinting methods for the identification of timber origins”, October 8–9 2007. Bonn: Landbauforschung, vTI Agriculture and Forestry Research- 2008. **Sonderheft,** [*s. l.*], v. 321, p. 6-14, 2008.

DEGEN, B.; SEBBENN, A. M. Genetics and Tropical Forests. *In*: PANCEL, L.; KOHL, M. (ed.) **Tropical Forestry Handbook**. Berlin, Heidelberg: Springer, 2014. p. 1-30.

DEGEN, B.; WARD, S. E.; LEMES, M. R.; NAVARRO, C.; CAVERS, S.; SEBBENN, A. M. Verifying the geographic origin of mahogany (*Swietenia macrophylla* King) with DNA-fingerprints. **Forensic Science International: Genetics**, Amsterdam, v. 7, i.1, p. 55-62, 2013.

DEKLERCK, V., FINCH, K., GASSON, P., VAN DEN BULCKE, J., VAN ACKER, J., BEECKMAN, H., ESPINOZA, E. Comparison of species classification models of mass spectrometry data: Kernel Discriminant Analysis vs Random Forest; A case study of Afrormosia (*Pericopsis elata* (Harms) Meeuwen). **Rapid Communication in Mass Spectrometry,** [*s. l.*], v. 31, p. 1582–1588, 2017.

DORMONTT, E. E.; BONER, M.; BRAUN, B.; BREULMANN, G.; DEGEN, B.; ESPINOZA, E.; GARDNER, S.; GUILLERY, P.; HERMANSON, J. C.; KOCH, G.; LEE, S. L.; KANASHIRO, M.; RIMBAWANTO, A.; THOMAS, D.; WIEDENHOEFT, A. C.; YIN, Y.; ZAHNEN, J.; LOWE, A. J. Forensic timber identification: It's time to integrate disciplines to combat illegal logging. **Biological Conservation**, Oxford, v. 191, p. 790-798, 2015.

EL SHEIKHA, A. F. Novel molecular fingerprinting for geographical traceability of timber. **Journal of Tropical Forest Science**, 2013.

GENTRY, A. H. Bignoniaceae – Part II (Tribe Tecomeae). **Flora Neotropica**, New York, v. 25, p. 1-370, 1992.

GRAY, J. A. **Forest concession policies and revenue systems: country experience and policy changes for sustainable tropical forestry**. [*S. l.*]: World Bank Publications. 2002.

HONORIO CORONADO, E. N.; BLANC-JOLIVET, C.; MADER, M.; GARCÍA-DÁVILA, C. R.; GOMERO, D. A.; CASTILLO, D. T. DEL.; LLAMPAZO, G. F.; PIZANGO, G. H.; SEBBENN, A. M.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; TYSKLIND, N.; TROISPOUX, V.; MASSOT, M.; CARVALHO, C.; LIMA, H. C.; CARDOSO, D.; DEGEN, B. SNP markers as a successful molecular tool for assessing species identity and geographic origin of trees in the economically important South American legume Genus *Dipteryx*. **Journal of Heredity**, Oxford, v. 111, p. 346–356, 2020.

JOBLING, M. A.; GILL, P. Encoded evidence: DNA in forensic analysis. **Nature Reviews Genetics**, London, v. 5, i.10, p. 739-751, 2004.

KISHOR N, LESCUYER G. Controlling illegal logging in domestic and international markets by harnessing multi-level governance opportunities. **International Journal of the Commons,** [*s. l.*], v. 6, p. 255–270 2012.

LAWSON S.; MACFAUL, L. **Illegal logging and related trade: Indicators of the global response**. London: Chatham House London, 2010.

LI, Y.; VIA, B.K.; YOUNG, T.; LI, Y. Visible-Near infrared spectroscopy and chemometric methods for wood density prediction and origin/species identification. **Forests,** [*s. l.*], v. 10, p. 1–19, 2019.

LOHMANN, L. G.; ULLOA, U. 2011**. Bignoniaceae in iPlants prototype Checklist**. Disponível em: http://www.iplants.org. Acesso em: 10 ago. 2022.

LORENZI, H. **Árvores brasileiras**: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. 4. ed. Nova Odessa SP: Instituto Plantarum, 2002. V. 2.

LOUREIRO, A. A.; SILVA, M. F.; ALENCAR, J. C. **Essências madeireiras da Amazônia**. Manau: INPA, 1979. 187 p.

LOWE, A. J.; WONG, K.-N.; TIONG, Y.-S.; IYERH, S.; CHEW, F.-T. A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of merbau timber from logging concession to sawmill. **Silvae Genetica**, Frankfurt, v. 59, i.6, p. 263-268, 2010.

LOWE, A. J.; DORMONTT, E. E.; BOWIE, M. J.; DEGEN, B.; GARDNER, S.; THOMAS, D.; CLARKE, C.; RIMBAWANTO, A.; WIEDENHOEFT, A. Y. Yin, Opportunities for improved transparency in the timber trade through scientific verification. **BioScience,** Washington, v. 66, p. 990–998, 2016.

JOLIVET, C. B.; DEGEN, B. Use of DNA fingerprints to control the origin of sapelli timber (*Entandrophragma cylindricum*) at the forest concession level in Cameroon. **Forensic Science International: Genetics**, Amsterdam, v. 6, i.4, p. 487-493, 2012.

LOUREIRO, A. A., SILVA, M. F., ALENCAR, J. C. **Essências madeireiras da Amazonia**. INPA, v. 1, 1979.

MAUES, M. M., OLIVEIRA, P. E. A. M., KANASHIRO, M. Pollination biology in *Jacaranda copaia* (Aubl.) D. Don. (Bignoniaceae) at the "Floresta Nacional do Tapajós", Central Amazon. **Revista Brasileira de Botanica**, São Paulo, v. 31: p. 517-527, 2008.

MEIER-AUGENSTEIN, W. From stable isotope ecology to forensic isotope ecology: isotopes' tales. **Forensic Science International**, Lausanne, v. 300, p. 89–98, 2019.

METCALFE, C. R.; CHALK, A. **Anatomy of the dicotyledons**. Oxford: Clarendon, 1500. 1950.p.

MEYER-SAND, B. R. V; BLANC-JOLIVET, C.; MADER, M.; PAREDES-VILLANUEVA, K.; TYSKLIND, N.; SEBBENN, A. M.; GUICHOUX, E.; DEGEN, B. Development of a set of SNP markers for population genetics studies of Ipe (*Handroanthus* sp.), a valuable tree genus from Latin America. **Conservation Genetic Recourses**, [*s. l.*], v. 10, p. 779–781, 2017.

MILLIKEN, W. Plantas medicinais, malária e povos indígenas: estudos etnobotânicos no norte da Amazônia. **Boletim Integrado de Roraima**, Roraima, v. 4, n. 1, p. 23-30, 1998.

OLIVEIRA, S. A. G.; LOPES, M. T. G, CHAVES, F. C. M, MARTINS, C.C., ALVES, E. U. Estimation of generic parameters of *Plukenetia volubilis* L. seed germination. **Amazonian Journal of Agricultural and Environment Sciences,** [*s. l.*], v. 56, p. 49-54, 2013

PAUMGARTTEN, A. E. A.; NETO, A. B. B.; SOUSA, V. G. de; JUNIOR, S. B.; YARED, J. A. G. Desempenho Silvicultural de *Jacaranda copaia* (Aubl.) D. Don sob Diferentes Espaçamentos no Oeste do Estado do Pará. **Scientia Forestalis,** Piracicaba, v. 46, n. 120, p. 701-707, 2018.

PIGOZZO, R. J. B. **Espectroscopia de infravermelho-próximo em madeiras neotropicais**: aplicação na identificação e predição de propriedades físicas. 2011. 97 f. Tese (Doutorado) - Universidade de São Paulo-USP, São Paulo, 2011.

PORTO, L. P. M.; BRASIL, H. M. S.; SILVA, A. C. P.; REIS, A. F. S.; OLIVEIRA, D. B.; *et al*. **Manual de orientação técnica da arborização urbana de Belém**: guia para planejamento, implantação e manutenção da arborização em logradouros públicos. Belém: Universidade Federal Rural da Amazônia, 2013.

SAMPAIO, P. T. B; BARBOSA, A. P; FERNANDES, N. P. Ensaio de espaçamento com caroba – *Jacaranda copaia* (AUBL.) D. DON. BIGNONTACEAE. **Acta Amazonica**, Manaus, v. 9, p. 383-383, 1989.

SANDAK, A.; SANDAK, J.; NEGRI, M. Relationship between near-infrared (NIR) spectra and the geographical provenance of timber. **Wood Science and Technology,** New York, v. 45, p. 35–48, 2011.

SANTOS, S. K. S.; ORTEGA J. R.; ROCHA, G. P.; MELO, A. F.; OLIVEIRA, D. A.; BRANDÃO, M. M.; ROYO, V. A.; MENEZES, E. V. Transferabilidade de marcadores SSR, diversidade e estrutura genética de *Syagrus oleraceae*. **Brazilian Journal of Development**, [*s. l.*], v. 6, p. 59931-59947, 2020.

SILVA, S.; SOUZA, R. S.; BELARMINO, K. J. A.; FERREIRA, E. C.; ALVES, C. A. B.; *et al*. **Etnobotânica de plantas medicinais na comunidade rural da Palmeira no município de Cuitegi-PB**. *In*: LUCENA, R. F. P.; LUCENA, C. M.; CARVALHO, T. K. N.; FERREIRA, E. C. Plantas e animais medicinais da Paraíba: um olhar da etnobiologia e etnoecologia. Cabedelo: Editora IESP, 2018. p. 328-373.

SEBBENN, A. M.; BLANC-JOLIVET, C.; MADER, B.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; CORONADO, E. N. H.; GARCIA-DAVILA, C.; TYSKLIND, N.; TROISPOUX, V.; DELCAMP, A.; DEGEN, B. Nuclear and plastidial SNP and INDEL markers for genetic tracking studies of *Jacaranda copaia*. **Conservation Genetic Resources**, [*s. l.*], v. 11, p. 341–343, 2019.

STARK, T.; PANG CHEUNG, S. **Sharing the Blame**: Global Consumption and China's Role in Ancient Forest Destruction. China: Greenpeace China, 2006.

TATAGIBA, S. D.; SOUSA, I. S.; OLIVEIRA, A. E. W. Etnobotânica de plantas medicinais na região de integração do Rio Tapajós, comunidade do Bairro Maria Magdalena, município de Itaituba, Pará, Brasil. **Biota Amazônia**, Manaus, v. 9, n. 4, p. 41-49. 2019.

TERDWONGWORAKUL, A.; PUNSUWAN, V.; THANAPASE, W.; TSUCHIKAWA, S. Rapid assessment of wood chemical properties and pulp yield of Eucalyptus camaldulensis in Thailand tree plantations by near infrared spectroscopy for improving wood selection for high quality pulp. **The Japan Wood Research Society,** [*s. l.*], v. 51 p. 167–171. 2005.

TONINI, H., OLIVEIRA jr, M. M. C.; SCHWENGBER, D. Crescimento de espécies nativas da Amazônia submetidas ao plantio no estado de Roraima. **Revista Ciência Florestal,** Santa Maria, v. 2, p. 151-158. 2008.

VINSON, C. C.; KANASHIRO, M.; SEBBENN, A. M.; WILLIAMS, T. C.; HARRIS, S. A.; BOSHIER, D. H. Long-term impacts of selective logging on two Amazonian tree species with contrasting ecological and reproductive characteristics: inferences from Eco-gene model simulations. **Heredity**, Edinburgh, v. 115, i.2, p. 130-139, 2015a.

VINSON, C. C.; KANASHIRO, M.; HARRIS, S. A.; BOSHIER, D. H. Impacts of selective logging on inbreeding and gene flow in two Amazonian timber species with contrasting ecological and reproductive characteristics. **Molecular Ecology**, Oxford, v. 24, i.1, p. 38-53, 2015b.

WIT, M.; VAN DAM, J.; CERUTTI, P. O.; LESCUYER, G.; KERRETT, R.; MCKEOWN, J. P.Chainsaw milling: supplier to local markets—a synthesis. *In*: ETFRN News (ed. (ETFRN) ETFRN), p. 16. **ETFRN, Wageningen, the Netherlands,** 2010.

3 DNA FINGERPRINTING USING SNPS MARKERS

ABSTRACT

Amazon and other tropical forest are actually subject to strong deforestation, generally originated from illegal logging, resulting in ecological, environmental and economic problems. Aiming stop deforestation and timber commercialization of illegal logging of tropical forest, new laws has been introduced in many countries. Samples from 43 populations from Brazil, Bolivia, French Guiana and Peru of the Neotropical tree *Jacaranda copaia* were used to investigate the power of 113 nuclear SNPs, 11 CpSNPs and four MtSNPs loci to determine the country and population origin of timber. We were able to 100% correct determine country and site origin of samples from nSPS and nCpMtSNPs. The nuclear, chloroplast and mitochondrial gene markers used are clearly efficient to be used for a timber genetic certification system and to confirm legality of timber.

Keywords: bayesian Method; DNA fingerprinting; frequency method; nearest neighbour method; single nucleotide polymorphism; timber tracking.

RESUMO

A Amazônia e outras florestas tropicais estão realmente sujeitas a um forte desmatamento, geralmente decorrente da extração ilegal de madeira, resultando em problemas ecológicos, ambientais e econômicos. Com o objetivo de parar o desmatamento e a comercialização de madeira do corte ilegal de florestas tropicais, novas leis foram introduzidas em muitos países. Amostras de 43 populações do Brasil, Bolívia, Guiana Francesa e Peru da árvore neotropical *Jacaranda copaia* foram usadas para investigar o poder de 113 SNPs nucleares, 11 CpSNPs e quatro loci MtSNPs para determinar o país e a origem populacional da madeira. Conseguimos determinar 100% corretamente a origem do país e do local de amostras de nSPS e nCpMtSNPs. Os marcadores genéticos nucleares, cloroplastos e mitocondriais utilizados são claramente eficientes para serem usados em um sistema de certificação genética de madeira e para confirmar a legalidade da madeira.

Keywords: polimorfismo de nucleotídeo único; impressão digital de DNA; acompanhamento de madeira; método bayesiano; método de frequência; método do vizinho mais próximo.

3.1 INTRODUCTION

Since 2016, deforestation in the Amazon has increased strongly, recording alarming amounts of degraded and deforested areas by clear cutting. Data collected by Imazon (2022) show an area of 4,514 km² destroyed between August 2021 and January 2022. After this period, according to the Space Research Institute (INPE), in a survey from January to May 2022, the deforestation record was the highest in the last five years, with an area of 2,744.41 km² under alert, of these, in January alone, 261 km² of forest were cut down, representing 33% more than in the same month of the previous year. The total amount represents 21% of the entire area deforested last year [oeco.org]. When we look at the states with the highest number of deforestation in this period, we have Amazonas occupying the first place, with 364.29 km² deforested. Of this total, 34% was deforestation in areas not intended for tree cutting, 27% occurred in settlements and another 27% in areas with a rural environmental registry (CAR). In second place was the state of Pará, which recorded the clearing of 225.18 km² of virgin forest, with 26.4% occurring in areas not intended for forest exploitation, 26.2% in areas of environmental protection and 13% in conservation units. Rondônia took third place, with a 197% increase in deforested area, with 24.2% occurring in areas of public forests not intended for forest exploitation, 28.5% in areas registered in the CAR, 18% in permanent preservation areas (APPs) and 12.4% in conservation units [oeco.org]. This clearly shows the large amount of timber recently illegally extracted from areas of the Brazilian Amazon.

The biggest problem about the illegality of the trade is in the documentation of the origin of the wood, having falsified certificates and documentation and with missing or false information (LESCUYER, *et al*., 2014). In 2014, the World Bank estimated that 25% of the world's entire timber supply was illegal, which was worth up to US$20 billion. According to the Brazilian Federal Police, in 2021, 90% of the wood extracted from the Amazon Forest had an illegal origin. Such timber tracking methods are required, but methods based only on documentation are sensitive to manipulation and forgery.

Therefore, exporting companies and institutions responsible for controlling the origin of imported timber, such as customs, federal police and Interpol need reliable tools to prove and confirm the declared origin of wood and its derivative products, traded internationally. To timber tracking the species, country and specific site origin of imported timber, many methods have been tested, such as chemical analysis or mass spectrometry as species timber differences (FEDERAL POLICE BRAZIL, 2011; FIDELIS *et al*., 2012, PAREDES-VILLANUEVA *et al*., 2018) near-infrared spectroscopy (BERGO, 2016) and stable isotopes (KAGAWA, LEAVITT, 2010; VLAM *et al*., 2018), wood anatomy (GASSON, 2011; MOYA *et al*., 2013), and DNA

fingerprint (CHAVES *et al*., 2019; DEGEN *et al*., 2013; LOWE *et al*., 2016; TNAH *et al*., 2010; JOLIVET *et al*., 2011; BLANC-JOLIVET *et al*., 2018; ABEELE *et al*., 2019, PAREDES-VILANUEVA *et al*., 2020). Results of different methods have been shows strong potential for species determination, country and site origin, in special using DNA fingerprint (CHAVES *et al*., 2019; CORONADO *et al*., 2020; BLANC-JOULIVET *et al*., 2018; DORMONTT *et al*., 2020). Methods such us wood anatomical, isotopic, and spectrometric methods are limited to all species, country and site origin determination, due to variations in tissue type, individual sample age, individuals and population genetic differences or environmental influences on timber composition (PE *et al*., 1997; DURAND *et al*., 1999; TNAH *et al*., 2009). DNA fingerprinting methods have successfully been applied to identify tree species, to track the country of origin (DEGEN *et al*., 2013), to verify the forest concession from which it was issued (JOLIVET; DEGEN, 2012), and to track individual trees (LOWE *et al*., 2010). Thus, molecular methods that allow correct identification of tree species and tracking of timber origin are essential for controls on the legality of timber by public authorities, industry, and trading companies (LOWE *et al*., 2016).

In the current work we investigated the use of the DNA fingerprint method to track the intensive and high value wood of the Neotropical pioneer tree *Jacaranda copaia* (Aubl.) D. Don. (Bignonaceae).

3.2 MATERIALS AND METHODS

3.2.1 Sampling

Were collected cambium or leave samples from 832 trees from 43 natural populations (2 to 31 individuals) in the Amazon rainforest of four countries (Bolivia, Brazil, French Guiana and Peru) and all trees sampled were georeferenced with GPS usage (Table 1, Figure 1). All samples were stored in a labelled plastic bag with silica gel. The samples were collected by the Institut National de la Recherche Agronomique- INRA together with the forest authorities (Office National des Forêts, ONF) in seven populations in French Guiana (2 to 30 individuals per site); In Brazil, samples were collected in national forests, extractive reserves, ecological stations, and national parks with the support of Chico Mendes Institute of Biodiversity (Brazilian governmental institution), totaling 12 sites (4 to 31 individuals per site); In Peru, samples were collected in national forests, extractive reserves, ecological stations, national parks, farms, and forest concessions, totaling 17 sites (2 to 30 individuals per site); In Bolivia samples were collected in farms and forest concessions from five sites (5 to 29 individuals per

site). However, due to the low sample size in five populations (2 trees), for genetic analysis, these individuals were grouped with the closest population (Table 1), as well as we also divided the Brazilian samples in West (six populations) and East (six populations) origin and Peru in North (nine populations) and South (eight populations) origin. The minimum distance among sampled trees within populations was 50 m and the distance among populations ranged from 23 to 2648 km. After sampling, all collected plant material was stored and dried in silica gel. All samples were registered in a database at the Thünen Institute (SampleDataBase, Grosshansdorf, Germany).

Table 1. Information on the sampled size (n), location and latitude (Lat) and longitude (Long), and group abbreviation (Group abbr).

Country	Population	n	Lat	Long	Group abbr	n
1-French Guiana	Counami	30	5,41543	-53,175	1-FG-Couna	30
2-French Guiana	Sinamary	2	5,2884	-52,916		
3-French Guiana	Piste de Paul Isnard	27	5,33216	-53,957	2-FG-Isnard	27
4-French Guiana	Acapou	2	5,27343	-54,218		
5-French Guiana	Route de Cocoa	30	4,56779	-52,406	3-FG-Cocoa	36
6-French Guiana	Regina	2	4,13118	-52,088		
7-French Guiana	Saut Maripa	28	3,87833	-51,857	4-FG-Mari	28
8-Brazil	ESEC de Maraca-RR	31	3,37032	-61,444	5-BRW-Mara	31
9-Brazil	Flona de Anauá e arredores-Rorainópolis-RR	28	-0,9339	-60,451	6-BRW-Anau	28
10-Brazil	AMATA Flona do Jamari-RO	8	-9,4014	-62,911	7-BRW-Jama	8
11-Brazil	ESEC do Jarí	15	-0,4955	-52,829	8-BRW-Jarí	15
12-Brazil	Resex Chico Mendes-Xapuri-AC (AMATA-Flona do Jamari-RO)	16	-10,504	-68,595	9-BRW-Xapu	16
13-Brazil	Resex Chico Mendes-Comunidade Cumaru-Assis-AC	15	-10,772	-69,647	10-BRW-Cuma	15
14-Brazil	FLONA Amapá-AP	20	0,52785	-51,128	11-BRE-Amap	20
15-Brazil	PARNA da Ana Avilhanas-AM	11	-2,5345	-60,837	12-BRE-Avilh	11
16-Brazil	Flona de Tapajós-PA	27	-2,8687	-54,92	13-BRE-Tapa	27

continued on next page

Table 1. (Continued).

Country	Population	n	Lat	Long	Group abbr	n
17-Brazil	Resex Tapajós-Arapins-PA	11	-3,0792	-55,278	14-BRE-Arap	11
18-Brazil	FLONA Tefé-AM	4	-3,5248	-64,972	15-BRE-Tefe	4
19-Brazil	FLONA do Carajás	23	-6,0628	-50,059	16-BRE-Cara	23
20-Peru	Dpto Loreto, Prov.Maynas, Dist. El Napo, Huiririma Native Community	26	-2,4761	-73,744	17-PEN-Huiri	26
21-Peru	Huaman Urco	27	-3,3128	-73,198	18-PEN-Urco	27
22-Peru	Dpto Loreto, Prov. Maynas, Las Amazonas, Estación Biológica Madreselva	28	-3,6312	-72,233	19-PEN-Madr	28
23-Peru	Dpto Loreto, Prov. Mayna, Dist. Iquitos, Comunidad Campesina Yarina	28	-3,827	-73,567	20-PEN-Yari	28
24-Peru	Allpahuayo	2	-3,9544	-73,422		
25-Peru	Dpto Loreto, Prov. Mariscal Ramón Castilla, Centro Poblado Unión Progresista	27	-3,9727	-70,841	21-PEN-Prog	29
26-Peru	Dpto Loreto, Prov. Requena, Jenaro Herrera Research Centre	11	-4,8966	-73,646	22-PEN-Cent	11
27-Peru	Jenaro Herrera	25	-4,9158	-73,649	23-PEN-Herre	25
28-Peru	Dpto Loreto, Prov. Alto Amazonas, Dist. Jeberos, Centro Poblado Jeberos	26	-5,2598	-76,317	24-PEN-Jebe	26
29-Peru	Shucushuyacu	27	-6,0199	-75,854	25-PEN-Shuc	27
30-Peru	Dpto Ucayali, Coronel Portillo, Concesión Forestal-Oxigeno para el Mundo	29	-8,8869	-74,034	26-PES-Portil	29
31-Peru	Dpto Ucayali, Padre Abad, Dist. Irazola, Macuya Forestry Research Station	30	-8,8766	-75,014	27-PES-Abad	30
32-Peru	Dpto Ucayali, Atalaya, Dist. Tahuania, Concesión Forestal-Javier Díaz	29	-9,9803	-73,817	28-PES-Diaz	29

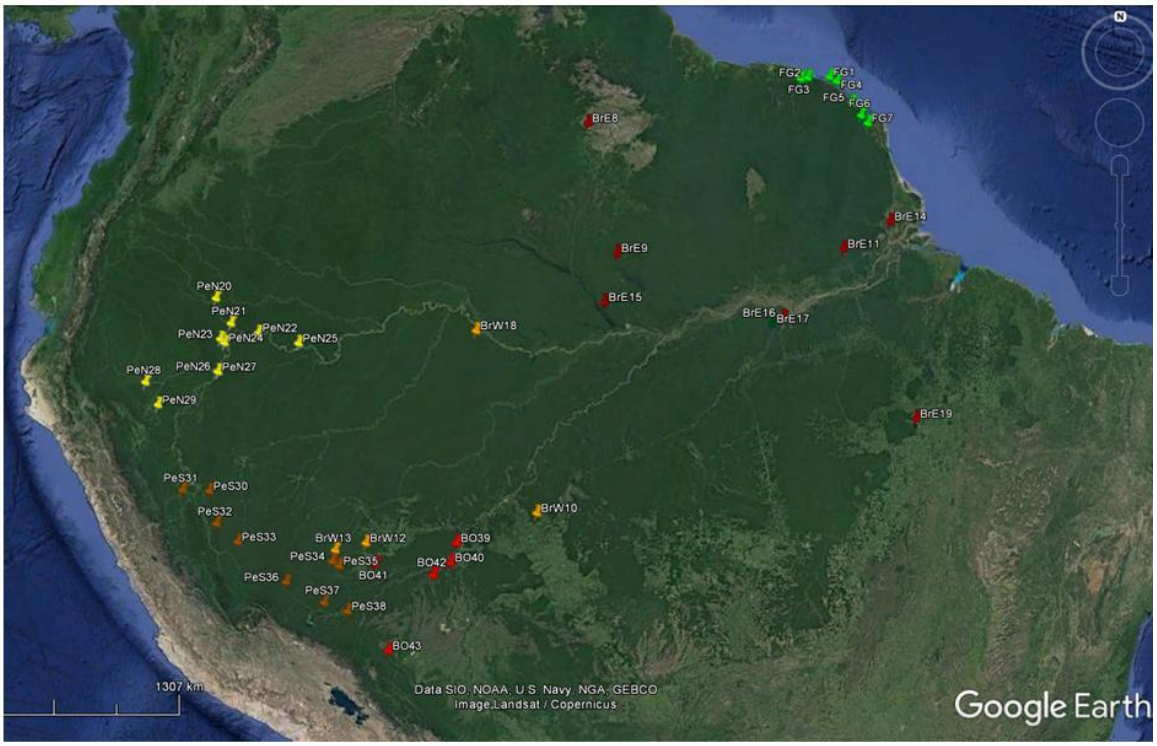
continued on next page

Table 1. (conclusion).

Country	Population	n	Lat	Long	Group abbr	n
33-Peru	Dpto Ucayali, Atalaya, Dist. Raymondi, Comunidad San Juan de Inuya	12	-10,582	-73,071	29-PES-Inuya	12
34-Peru	Dpto Madre de Dios,Tahuamanu, Dist. Iñapari Concesión Forestal Maderacre	31	-11,145	-69,758	30-PES-Made	33
35-Peru	Ibéria	2	-11,299	-69,524		
36-Peru	Dpto Madre de Dios, Parque Nacional Manu, Estación Biológica Cocha Cashu	15	-11,903	-71,403	31-PES-Cashu	15
37-Peru	Dpto Madre de Dios, Manu, Río Los Amigos, Estación Biológica Los Amigos	30	-12,565	-70,088	32-PES-Amig	30
38-Peru	Dpto Madre de Dios, Reserva Nacional Tambopata, La Torre-Sandoval	24	-12,832	-69,284	33-PES-Tamb	24
39-Bolivia	Riberalta, MABET	15	-10,442	-65,55	34-BO-MAB	15
40-Bolivia	Riberalta, El Desvelo	11	-11,093	-65,746	35-BO-Desve	11
41-Bolivia	Cobija, Road - Bella Vista (Oscar Kerdi)	13	-11,198	-68,287	36-BO-Vista	13
42-Bolivia	Riberalta, El Chorro	5	-11,514	-66,327	37-BO-Chorr	5
43-Bolivia	Rurrenabaque, Área Protegida Madidi	29	-14,162	-67,905	38-BO-Madi	29

Source: Prepared by author.

Figure 1 - Spatial distribution of samples for *Jacaranda copaia* in South America.



Source: Prepared by author.

3.2.2 DNA extraction and SNPs analysis

Cambium and leaf samples collected in Brazil were sent the Laboratory of Population Genetics and Forestry of São Paulo State University in Ilha Solteira, Brazil (UNESP-FEIS) for DNA isolation. Samples collected in French Guyana, Peru and Bolivia were sent to Thünen Institute facilities in Großhansdorf, Germany, for DNA isolation. DNA isolation from leaf and cambium was carried out according to Dumolin *et al*. (37). The samples were screened for 128 SNP and INDEL markers using the MassARRAY® iPLEX™ genotyping, where 113 were nuclear SNPs (nSNPs), 11 cloroplastidial SNPs (CpSNPs), and four mitochondrial SNPs (MtSNPs), all selected for genetic tracking analysis due to show a minimum amplification rate of 95% [38].

3.2.3 Genetic diversity and population differentiation

The genetic diversity for nSNP was characterized for each population and country by the total number of alleles (K), allelic richness (R), observed (H_o), and expected (H_e)

heterozygosity. The mean fixation index (F) was estimated to quantify the inbreeding within each population or country. The statistical significance of the F values was estimated using permutation of alleles among individuals. For CpMtSNP, the genetic diversity was characterized for each sample population, country and within Brazilian and Peru regions by K and H_e. The scandalized genetic differentiation, (G'_{ST}, HEDRICK, 2005) was estimated among all populations, countries, within Brazilian and Peru regions, and pairwise populations for all SNP, nSNP, and CpMtSNP markers. These analyses were carried out using the FSTAT 2.9.3.2. software (GOUDET, 1995). The pairwise G'_{ST} and spatial genetic distance among populations was used to investigate the isolation by distance gene dispersal, using the Spearman correlation coefficient (ρ).

3.2.4 Spatial genetic structure

We assessed SGS for all sample trees based on the coancestry coefficient (θ_{ij}) described in Loiselle *et al*. (1995), between mean pairs of individuals within 15 distance-classes, using the SPAGEDI 1.5 software (HARDY; VEKEMANS, 2002). The statistical significance of the average θ_{ij} of each distance class was derived by comparing the limits of the confidence interval at 95% probability for the average θ_{ij} for each distance class, estimated permuting (1000 times) genotypes between distances classes, using the SPAGEDI software.

3.2.5 Genetic assignment

Bayesian method (Rannala; Mountain, 1997) implemented in GeneClass 2.0 (PIRY *et al.*, 2004) was used to assign group (population or country) and individuals to its population, country, and within country region (Brazil West and East, Peru North and South) of origin. Both group and individual assignment were carried out for all SNP, nSNP, and CpMtSNP loci, and the most likely group determined by the highest score by the Bayesian criteria was used as an indicator of the power of the markers to compute the proportion of correctly assigned groups or individuals in self-assignments tests (CORNUET *et al*., 1999). Here the individuals of the reference data were self-classified to the sampled groups (populations, countries, and regions) using the leave-one-out approach (EFRON, 1983).

3.3 RESULTS

3.3.1 Genetic diversity and genetic differentiation

The total number of alleles (K) ranged among populations for 113 nSNPs from 122 to 226 alleles, for 11 CpSNPs from 10 to 15 alleles, for four MtSNPs from 4 to 6 alleles, and for 15 both CpMtSPs from 14 to 19 alleles (Table 2). Spearman correlation (ρ) was significantly higher than zero between the sample size (n) and K (ρ= 0.36, P= 0.026), between the K and H_o (ρ= 0.675, P= 0) and between K and H_e (ρ= 0.691, P= 0) (Supplementary material, Table S1). The observed heterozygosity (H_o) and expected (H_e) heterozygosity were low in all populations (H_o: ranging from 0.024-0.343; H_e: 0.029-0.349). The mean intrapopulation fixation index (F) was significantly higher than zero in 16 of the 38 populations (-0.036 to 0.533), suggestion inbreeding. The H_e values were also higher for nSNPSs (0.029-0.349) than for CpSNPs (0-0.109), MtSNPs (0-0.3), and CpMtSNPs (0-0.16). At countries level, for nSNPs, K was highest in Brazil for nSNPs (226), CpSNPs (22), MtSNPs (8) and CpMtSNPs (30), and lowest in French Guiana (nSNPs= 200; CpSNPs= 12, MtSNPs=6; CpMtSNPs= 18), and in Peru for MtSNPs (6). Bolivia presented the highest H_o (0.319) and H_e (0.359) values and lowest F (0.113), where French Guiana presented the lowest H_o (0.095) and H_e (0.192) values and the highest F (0.506). The H_e was also highest in Brazil for CpSNPs (0.282) and CpMtSNPs (0.251) and Bolivia for MtSNPs (0.336) and lowest in French Guiana for CpSNPs (0.014), MtSNPs (0.009), and CpMtSNPs (0.013). Comparing Brazil regions, west populations presented highest K, H_o, H_e and F for nSNPs, as well as the highest K and H_e for CpSNPs, MtSNPs and CpMtSNPs. Comparing Peru regions, South populations presented highest K, H_o, H_e, and F for nSNPs, as well as the highest H_e for MtSNPs and CpMtSNPs, where North populations presented highest K for CpSNPs and CpMtSNPs, as well as highest H_e for CpSNPs.

Table 2. Genetic diversity in *Jacaranda copaia* for all populations, countries, and regions within Brazil and Peru population for nSNPs (113), CpSNPs (11), MtSNPs (4), and CpMtSNPs (15).

		nSNPs				CpSNP		MtSNp		CpMtSNPs	
Populations	n	K	H_o	H_e	F	K	H_e	K	H_e	K	H_e
1-FG-Couna	30	122	0.024	0.029	0.178	12	0.017	4	0	16	0.013
2-FG-Isnard	27	190	0.109	0.197	0.445*	12	0.019	5	0.019	17	0.019
3-FG-Cocoa	36	196	0.094	0.202	0.533*	12	0.019	4	0	16	0.014
4-FG-Mari	28	195	0.148	0.254	0.417*	11	0	5	0.018	16	0.005
5-BW-Mara	31	193	0.299	0.308	0.028	11	0	5	0.109	16	0.029
6-BW-Anau	28	195	0.291	0.291	0	11	0	4	0	15	0
7-BW-Jama	8	207	0.194	0.298	0.35*	11	0	6	0.134	17	0.036
8-BW-Jarí	15	207	0.205	0.287	0.287*	12	0.018	4	0	16	0.014
9-BW-Xapu	16	204	0.275	0.315	0.127*	11	0	5	0.058	16	0.016
10-BW-Cum	15	147	0.278	0.283	0.019	11	0	4	0	15	0
11-BE-Amap	20	204	0.214	0.23	0.067	11	0	4	0	15	0
12-BE-Avilh	11	207	0.28	0.29	0.033	11	0	4	0	15	0
13-BE-Tapa	27	197	0.261	0.274	0.045	11	0	4	0	15	0
14-BE-Arap	11	183	0.24	0.279	0.14*	11	0	4	0	15	0
15-BE-Tefe	4	207	0.141	0.257	0.453*	11	0	4	0	15	0.017
16-BE-Cara	23	190	0.257	0.283	0.091*	11	0	4	0	15	0
17-PN-Huiri	26	149	0.053	0.059	0.097	12	0.091	5	0.038	17	0.077
18-PN-Urco	27	147	0.06	0.066	0.085	12	0.091	5	0.118	17	0.098
19-PN-Madr	28	147	0.054	0.066	0.179*	14	0.082	5	0.066	19	0.078
20-PN-Yari	28	147	0.059	0.064	0.087	11	0	5	0.128	16	0.037
21-PN-Prog	29	142	0.065	0.069	0.06*	13	0.054	6	0.239	19	0.103
22-PN-Cent	11	146	0.043	0.065	0.338*	11	0	4	0	15	0
23-PN-Herre	25	145	0.05	0.056	0.112	11	0.015	5	0.021	16	0.017
24-PN-Jebe	26	142	0.041	0.061	0.329*	15	0.066	4	0	19	0.049

continued on next page

Table 2. (conclusion)

25-PN-Shuc	27	147	0.047	0.053	0.105	10	0	4	0	14	0
26-PS-Portil	29	146	0.046	0.075	0.394*	12	0.012	4	0	16	0.009
27-PS-Abad	30	140	0.063	0.064	0.02	10	0	4	0	14	0
28-PS-Diaz	29	139	0.054	0.061	0.12	11	0	4	0	15	0
29-PS-Inuya	12	213	0.056	0.061	0.086	11	0	4	0	15	0
30-PS-Made	33	206	0.297	0.314	0.052	11	0	4	0	15	0
31-PS-Cashu	15	213	0.271	0.31	0.125*	11	0	4	0	15	0
32-PS-Amig	30	209	0.312	0.315	0.01	11	0	4	0	15	0
33-PS-Tamb	24	209	0.284	0.313	0.092*	11	0	4	0	15	0
34-Bo-MAB	15	207	0.321	0.324	0.007	11	0	5	0.129	16	0.034
35-Bo-Desve	11	206	0.3	0.314	0.047	11	0	5	0.136	16	0.036
36-Bo-Vista	13	205	0.313	0.314	0.003	11	0	4	0	15	0
37-Bo-Chorr	5	212	0.343	0.349	0.017	13	0.109	6	0.3	19	0.16
38-Bo-Madi	29	226	0.322	0.311	-0.036	11	0	4	0	15	0
French Guiana	121	200	0.095	0.192	0.506*	12	0.014	6	0.009	18	0.013
Brazil	209	226	0.261	0.354	0.264*	22	0.282	8	0.165	30	0.251
Peru	429	217	0.111	0.222	0.498*	17	0.068	6	0.208	23	0.105
Bolivia	73	218	0.319	0.359	0.113*	13	0.088	7	0.336	20	0.154
Brazil West	113	223	0.277	0.366	0.243*	22	0.423	8	0.268	30	0.382
Brazil East	96	215	0.24	0.292	0.176*	12	0.003	4	0	16	0.002
Peru North	227	173	0.053	0.063	0.149*	17	0.063	6	0.1	23	0.073
Peru South	202	217	0.177	0.326	0.457*	14	0.057	6	0.252	20	0.109
Total	832	183.4	0.178	0.204	0.146	11.4	0.016	4.4	0.04	15.9	0.023

*P< 0.05; n is the sample size; K is the total number of alleles; H_o is the observed heterozygosity; H_e is the expected heterozygosity; F is the fixation index.

Source: Prepared by author.

The genetic differentiation (G'_{ST}) in all samples was higher for CpMtSNps (ranging from 0 to 0.997) and all SNPs (ranging from 0.045 to 0.698) than for nSNPs (ranging from 0.013 to 0.627), with exception for some populations where the SNPs were no polymorphic,

such as French Guiana, Brazilian East, Peruvian North and South populations (Table 3). The Spearman correlation coefficient (ρ) between pairwise G'_{ST} and spatial distance was significantly higher than zero (P< 0.01) for all loci (0.466), 113 nSNPs (0.35) and 15 CpMtSNP loci (0.479), indicating a strong isolation by distance genetic (IBD) pattern (Figure 2). In concordance, the analysis of spatial genetic structure based on estimate of mean pairwise coancestry coefficient between individuals show also an IBD pattern, with higher probability of kinship between individuals located up to 737 km, decreasing for not significant or significantly lower than zero coancestry after this distance (Figure 3). Based on 128 SNPs loci, the G'_{ST} among countries was high (0.506), ranging among populations within countries from 0.158 (French Guiana) to 0.583 (Peru). Based on 128 SNPs loci, the G'_{ST} between Brazilian West and East regions was low (0.14), being higher among West (0.542) than East (0.284) populations within regions, where between Peruvian North and South regions, G'_{ST} was also low (0.227), being lower among North (0.045) than South (0.622) populations within regions. Between countries, G'_{ST} ranged from substantial (0.249) to high (0.565), as well as between Brazilian West and East and countries (0.198-0.492) and Peru Nort and South and countries (0.149-0.531) (Supplementary material, Table S2).

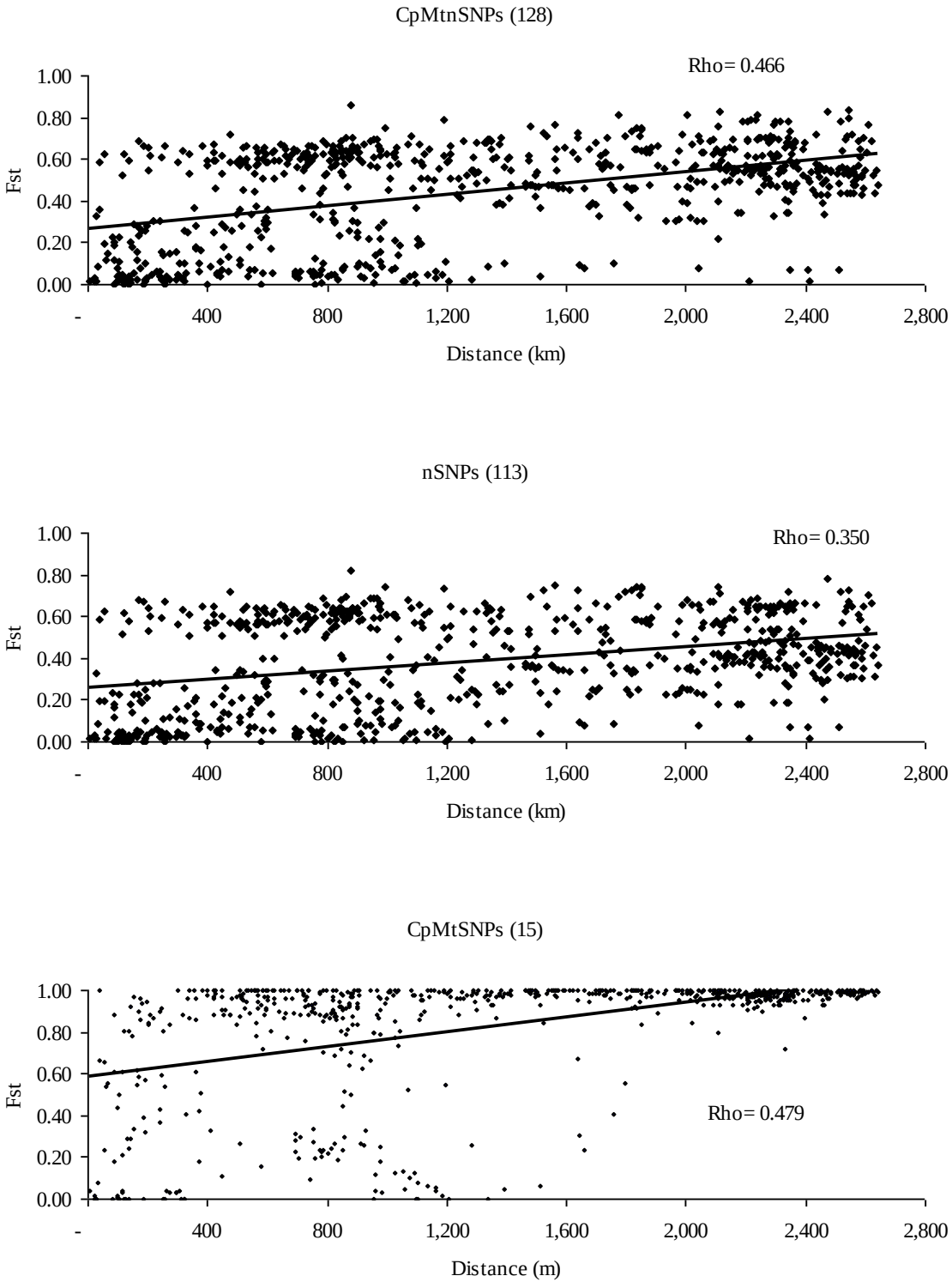
Table 3. Genetic differentiation ($G_{ST}^{'}$) among all in *Jacaranda copaia* populations, among and within countries, and regions within Brazil and Peru population for all 128 SNPs, nSNPs (113), and CpMtSNPs (15).

Sample	np	SNPs (128)	nSNPs (113)	CpMtSNPs (15)
All populations	38	0.698 (0.043)*	0.627 (0.042)*	0.997 (0.019)*
All countries	4	0.506 (0.044)*	0.416 (0.04)*	0.904 (0.137)*
French Guiana	4	0.158 (0.025)*	0.165 (0.027)*	0.003 (0.001)
Brazil	12	0.498 (0.047)*	0.399 (0.038)*	0.94 (0.099)*
Bolívia	5	0.258 (0.045)*	0.209 (0.039)*	0.263 (0.226)*
Peru	17	0.583 (0.059)*	0.57 (0.059)*	0.263 (0.219)*
Brazil West	6	0.542 (0.056)*	0.402 (0.049)*	0.949 (0.097)*
Brazil East	6	0.284 (0.041)*	0.284 (0.041)*	0
East vs West	2	0.139 (0.024)*	0.125 (0.025)*	0.271 (0.059)*
Peru North	9	0.045 (0.013)*	0.013 (0.004)*	0.085 (0.099)*
Peru South	8	0.622 (0.063)*	0.607 (0.063)*	0.269 (0.231)*
North vs South	2	0.228 (0.030)*	0.231 (0.032)*	0.218 (0.052)*

*P< 0.05; np is number of populations; Inside the parentheses is the 95% standard error, 1.96SE.

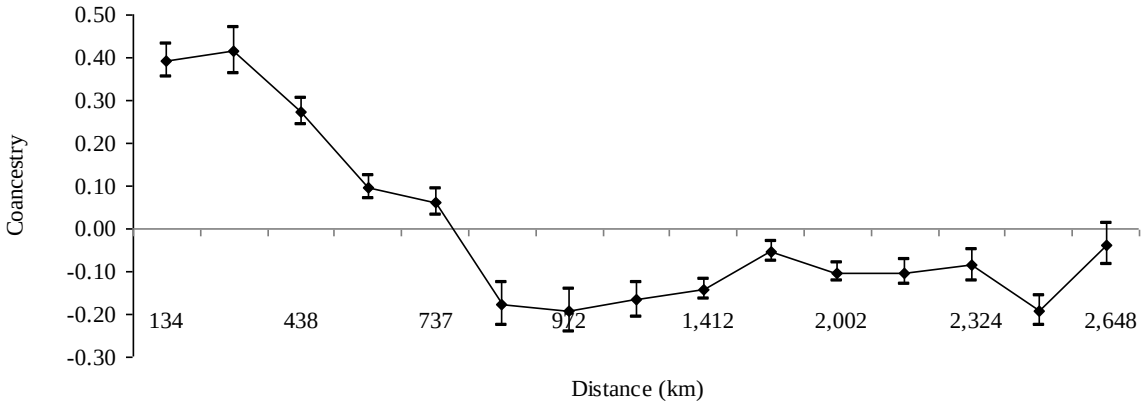
Source: Prepared by author.

Figure 2 - Pattern of isolation by distance in *Jacaranda copaia* populations. $G_{ST}^{'}$ is the pairwise genetic differentiation between populations for all 128 loci, 113 nSNPs and 15 CpMt loci. The Spearman correlation coefficient (ρ) was significantly higher than zero (P< 0.01) for all loci (0.466), 113 nSNP (0.35) and 15 CpMtSNP loci (0.479). Source: Prepared by author.



Source: Prepared by author.

Figure 3 - Correlogram of mean coancestry coefficient among individuals within 15 distance classes in *Jacaranda copaia*. Source: Prepared by author.



Source: Prepared by author.

The percentage of polymorphic loci (P) ranged among locations from 8.1 to 91.9%, with a mean among locations of 49.9% (Table 4). The total number of alleles (N_a) over all 68 SNPs loci analysed ranged among locations from 73 to 117 alleles, with a mean among locations of 94.2. In 15 of the 21 sample-populations, identical multilocus genotypes were detected, indicating a genotypic diversity (G_d) ranging from 0.27 to 1.0, with mean of 0.68. The mean observed heterozygosity (H_o) ranged from 0.01 to 0.49 (mean of 0.15), expected heterozygosity (H_e) ranged from 0.01 to 0.32 (mean of 0.13). In four of the 21 sample-populations there was a significant excess of heterozygous individuals in comparison with expected by Hardy-Weinberg equilibrium (HWE) expectations, with fixation index (F) ranging from -0.24 to -0.61, and in six sample-populations there was a significant excess of homozygous individuals, with F ranging from 0.15 to -0.23, indication inbreeding.

Table 4. Genetic diversity for species location.

Country	Location	Species	n	P(%)	N_a	G_d	H_o	H_e	F
Bolivia	Bajo Paraguá	*H. serratifolius*	20	75.7	113	0.63	0.09	0.12	0.23*
Bolivia	Concepción	*H. serratifolius*	19	16.2	77	0.5	0.05	0.04	-0.03
Bolivia	Guarayos	*H. serratifolius*	20	81.1	110	1.0	0.39	0.28	-0.32*
Bolivia	Riberalta-Mabet	*Handroanthus* sp.	13	64.9	99	0.78	0.08	0.09	0.09
Bolivia	Roboré	*H. serratifolius*	23	67.6	101	0.59	0.08	0.09	0.15*
Bolivia	Yapacaní- Naranjal	*H. serratifolius*	20	70.3	95	0.44	0.49	0.29	-0.61*
Brazil	FLONA do Jamari	*H. impetiginosus*	20	16.2	77	0.5	0.03	0.03	-0.03
Brazil	FLONA do Jamari	*H. incanus*	5	62.2	100	1.00	0.35	0.23	-0.44
Brazil	FLONA do Jamari	*H. serratifolius*	6	8.1	73	0.67	0.02	0.02	0.14
Brazil	FLONA de Tefé	*Handroanthus* sp.	14	8.1	79	0.50	0.01	0.01	0.02
Brazil	ESEC de Maracá	*Handroanthus* sp.	21	78.4	112	1.00	0.14	0.19	0.17*
Brazil	FLONA Amapá	*Handroanthus* sp.	22	86.5	116	1.00	0.40	0.30	-0.24*
Brazil	FLONA de Tapajós	*H. serratifolius*	20	91.9	117	0.81	0.14	0.19	0.19*
Brazil	FLONA do Carajás	*H. serratifolius*	32	81.1	109	0.95	0.25	0.29	0.09
Brazil	PARNA do Jaú	*Handroanthus* sp.	22	10.8	80	0.43	0.01	0.02	0.21*
Brazil	RESEX Chico Mendes	*Handroanthus* sp.	38	18.9	78	0.44	0.04	0.03	0.05
Brazil	RESEX Tapajós-Arapins	*Handroanthus* sp.	21	13.5	74	0.47	0.03	0.03	0.23*
French Guyana	Belizon	*H. serratifolius*	8	13.5	75	0.43	0.05	0.05	0.08
French Guyana	Belizon	*H. impetiginosus*	12	73.0	112	0.27	0.07	0.09	0.21
French Guyana	Paracou-Kourou-Cayenne	*H. serratifolius*	25	89.2	104	1.00	0.45	0.32	-0.35*
French Guyana	Organabo	*H. serratifolius*	18	21.6	77	0.83	0.07	0.06	-0.09
Mean			19	49.9	94.2	0.68	0.15	0.13	-0.01

*P< 0.05; n is the sample size; P is the percentage of polymorphic loci; N_a is the total number of alleles; G_d is the genotypic diversity; H_o is the observed heterozygosity; H_e is the expected heterozygosity; F is the fixation index;

Source: Prepared by author.

3.3.2 Genetic assignment

The grouped sample assignment test for all SNPs and nSNPs was able to self-assignment of 100% of population and individuals to the correct origin country and population (Table 5). The grouped sample assignment test for CpMtSNPs was also able to self-assignment of population to the correct origin country and region within Brazil and Peru (score ranging from 93.9 to 100%). Considering a score higher than 95%, only 7.9% of grouped individuals were correct assigned to population of origin for (score ranging from 98 to 100%). However, 42.1% of grouped individuals were correct assigned to population of origin (score ranging from 26.7 to 100%), 92.1% of individuals were correct assigned to country origin (score ranging from 21.5 to 100%), 58.6% and of individuals were assigned to corrected region within Brazil and Peru countries (score ranging from 25.6 to 100%). These results indicate that the use of all SNPs or only nSNP loci can be used to correct assignment for grouped samples of individuals to country, region within countries, and populations origin.

The assignment test for individuals was equal or higher for all SNPs than nSNPs for country or population determination of correct origin (Tables 5 and 6). The assignment test for individuals for all SNPs and nSNPs was able to self-assignment among 35.5 (Peru) to 96.2% of country origin, where between region within Brazil was high (> 80%) and Peru was low for both North (maximum of 29.2%) and substantial for East (64 and 58.8%, respectively). For populations, the assignment test for individuals was able to self-assignment among 0 to 100% of correct origin. The assignment test for individuals was equal low for CpMtSNPs for country, population and regions determination of correct origin (< 33%). Spearman correlation (ρ) for countries and populations was significantly higher than zero between all SNPs or nSNPs *versus* K (with exception to all SNP) *versus* H_o, and *versus* H_e, as well as ssignificantly lower than zero for population between nSNPs *versus* F (Supplementary material, Table S3).

Table 5. Score results of the self-assignment tests to the country and population in *Jacaranda copaia* of origin for grouped and individuals (mean scores) for all SNPs (All), nSNPS, and CpMtSNPs (CpMt).

	Grouped				Individuals			
		All	nSNPs	CpMt		All	nSNPs	CpMt
Assigned sample	Rank	Score (%)	Score (%)	Rank	Score (%)	Score (%)	Score (%)	Score (%)
French Gu	French Gu	100	100	French Gu	100	88.6	84.3	0
Brazil	Brazil	100	100	Brazil	100	95.7	87.2	0
Brazil West	Brazil West	100	100	Brazil West	100	95.5	95.5	0
Brazil East	Brazil East	100	100	Brazil East	93.9	94.5	82	0
Peru	Peru	100	100	Peru	100	46.6	35.5	18
Peru North	Peru North	100	100	Peru North	99.8	29.2	14.8	0
Peru South	Peru South	100	100	Peru South	100	64	58.8	4.5
Bolivia	Bolivia	100	100	Bolívia	100	96.2	96.2	0
Correct		100	100		100			

French Gu is the French Guiana; Spearman coefficient (ρ) between sample size (n) and scores for CpMtSNP of 0.357 (P= 0.028).

Source: Prepared by author.

Table 6. Score results of the self-assignment tests to the country and population in *Jacaranda copaia* of origin for individuals (mean scores) for all SNPs (All), nSNPS, and CpMtSNPs (CpMt).

	Grouped				Individuals			
		All	nSNPs	CpMt		All	nSNPs	CpMt
Assigned sample	Rank	Score (%)	Score (%)	Rank	Score (%)	Score (%)	Score (%)	Score (%)
1-FG-Couna	1-FG-Couna	100	100	3-FG-Cocoa	55.9	100	100	0
2-FG-Isnard	2-FG-Isnard	100	100	3-FG-Cocoa	41.3	18.5	18.5	0
3-FG-Cocoa	3-FG-Cocoa	100	100	3-FG-Cocoa	58.7	0	0	0
4-FG-Mari	4-FG-Mari	100	100	4-FG-Mari	79.7	57.1	57.1	0
5-BW-Mara	5-BW-Mara	100	100	5-BW-Mara	99.2	100	100	0
6-BW-Anau	6-BW-Anau	100	100	6-BW-Anau	38	100	100	0
7-BW-Jama	7-BW-Jama	100	100	34-B0-MAB	44.3	100	85.7	0
8-BW-Jarí	8-BW-Jarí	100	100	8-BW-Jarí	26.7	73.3	73.3	0
9-BW-Xapu	9-BW-Xapu	100	100	30-PS-Made	19.9	100	100	0
10-BW-Cum	10-BW-Cum	100	100	30-PS-Made	28.6	100	100	0
11-BE-Amap	11-BE-Amap	100	100	6-BW-Anau	35.5	100	100	0
12-BE-Avilh	12-BE-Avilh	100	100	6-BW-Anau	30.5	100	100	0
13-BE-Tapa	13-BE-Tapa	100	100	6-BW-Anau	37.7	96.3	96.3	0
14-BE-Arap	14-BE-Arap	100	100	6-BW-Anau	30.5	100	100	0
15-BE-Tefe	15-BE-Tefe	100	100	6-BW-Anau	21.5	75	0	0
16-BE-Cara	16-BE-Cara	100	100	6-BW-Anau	36.4	95.7	95.7	0
17-PN-Huiri	17-PN-Huiri	100	100	17-PN-Huiri	38.5	23.1	19.2	0
18-PN-Urco	18-PN-Urco	100	100	18-PN-Urco	52.4	25.9	14.8	0
19-PN-Madr	19-PN-Madr	100	100	30-PS-Made	99.9	42.9	14.3	0
20-PN-Yari	20-PN-Yari	100	100	20-PN-Yari	59.0	32.1	10.7	0
21-PN-Prog	21-PN-Prog	100	100	21-PN-Prog	100	55.2	6.9	33.3
22-PN-Cent	22-PN-Cent	100	100	27-PS-Abad	24.6	18.2	18.2	0

Table 6. (continued).

	Grouped				Individuals			
		All	nSNPs	CpMt		All	nSNPs	CpMt
Assigned sample	Rank	Score (%)	Score (%)	Rank	Score (%)	Score (%)	Score (%)	Score (%)
23-PN-Herre	23-PN-Herre	100	100	23-PN-Herre	44.2	20	12	0
24-PN-Jebe	24-PN-Jebe	100	100	24-PN-Jebe	50	15.4	11.5	0
25-PN-Shuc	25-PN-Shuc	100	100	27-PS-Abad	34.9	30.8	25.9	0
26-PS-Portil	26-PS-Portil	100	100	26-PS-Portil	44.3	37.9	35.9	0
27-PS-Abad	27-PS-Abad	100	100	27-PS-Abad	36.3	40	23.3	0
28-PS-Diaz	28-PS-Diaz	100	100	27-PS-Abad	35.3	37.9	34.5	0
29-PS-Inuya	29-PS-Inuya	100	100	27-PS-Abad	25.6	25	16.7	0
30-PS-Made	30-PS-Made	100	100	30-PS-Made	33.7	87.9	87.9	0
31-PS-Cashu	31-PS-Cashu	100	100	30-PS-Made	28.5	86.7	86.7	0
32-PS-Amig	32-PS-Amig	100	100	30-PS-Made	33.2	90	90	0
33-PS-Tamb	33-PS-Tamb	100	100	30-PS-Made	31.9	95.8	95.8	0
34-Bo-MAB	34-Bo-MAB	100	100	34-Bo-MAB	75.4	93.3	93.3	0
35-Bo-Desve	35-Bo-Desve	100	100	34-Bo-MAB	70.7	100	100	0
36-Bo-Vista	36-Bo-Vista	100	100	30-PS-Made	27.4	84.6	84.6	0
37-Bo-Chor	37-Bo-Chor	100	100	37-Bo-Chor	98	100	100	0
38-Bo-Madi	38-Bo-Madi	100	100	30-PS-Made	32.7	100	100	0
Correct		100	100		42.1			

French Gu is the French Guiana; Spearman coefficient (ρ) between sample size (n) and scores for CpMtSNP of 0.357 (P= 0.028); The correct assignment of individuals to population origin for scores > 95% is 7.9%. Source: Prepared by author.

3.3.3 Genotype self-assignment tests

The self-assignment test in the country level obtained a success of 72% with the Bayesian approach, 74% with the frequency method and 96% with the nearest neighbour (Table 7). The minimum success rate for the nearest neighbour method was 95%, while for Bayesian approach was 84% and 85% with the frequency method.

Table 7. Percent of correct self-assignment samples in country level by employing the Bayesian method, the frequency method and the nearest neighbour approach in *Jacaranda copaia* population.

Country	Sampling size	Bayesian	Frequency	Nearest neighbour
Bolivia	115	84%	85%	95%
Brazil	221	92%	95%	95%
French Guyana	63	41%	41%	100%
Mean		72%	74%	96%

Source: Prepared by author.

Self-assignment performed in location level used only sample-populations with 20 or more individuals: Bajo Paraguá Guarayos, Roboré and Yapacaní-Naranjal in Bolivia; FLONA do Jamari for *H. impetiginosus*, ESEC de Maracá, FLONA Amapá, FLONA de Tapajós, FLONA do Carajás, PARNA do Jaú, Resex Chico Mendes and Resex Tapajós-Arapiuns in Brazil; and Paracou-Kourou-Cayenne in French Guyana. The successes achieved for the self-assignment were of 77% with the Bayesian approach, the 79% with the frequency method and 85% with the nearest neighbour approach (Table 8).

Table 8. Percentile of correct self-assignment samples in location level by employing the Bayesian, frequency and the nearest neighbour method in *Jacaranda copaia* population.

Location	Sampling size	Bayesian	Frequency	Nearest neighbour
Bajo Paraguá (BO)	20	0%	0%	60%
Guarayos (BO)	20	80%	80%	80%
Roboré (BO)	23	96%	87%	48%
Yapacaní, Naranjal (BO)	20	100%	100%	100%
FLONA do Jamari (BR)	20	100%	100%	95%
ESEC de Maracá (BR)	21	86%	95%	95%
FLONA do Amapá (BR)	22	73%	95%	95%
FLONA do Tapajós (BR)	20	5%	5%	40%
FLONA do Carajás (BR)	32	72%	72%	91%
PARNA do Jaú (BR)	22	100%	100%	100%
RESEX Chico Mendes (BR)	38	95%	95%	95%
RESEX Tapajós-Arapiuns (BR)	21	100%	95%	95%
Paracau, Kourou-Cayenne (FG)	25	100%	100%	100%
Mean		77%	79%	85%

Source: Prepared by author.

The self-assignment test performed for Brazilian states level pooled together sample-populations belonging to the same state of origin: Acre (RESEX Chico Mendes), Amapá (FLONA do Amapá), Amazonas (FLONA de Tefé and PARNA do Jaú), Pará (FLONA de Tapajós, FLONA do Carajás and RESEX Tapajós-Arapiuns), Rondônia (FLONA do Jamari with all species) and Roraima (ESEC de Maracá). A total of 221 samples were analysed and the percentage of successfully self-assigned samples were of 87% with the Bayesian approach, the 89% with the frequency method and 93% with the nearest neighbour approach (Table 9).

Table 9. Percentile of correct self-assignment samples among Brazilian states level by employing the Bayesian, frequency and the nearest neighbour method in *Jacaranda copaia* population.

Brazilian States	Sampling size	Bayesian	Frequency	Nearest neighbour
Amazonas	36	100%	100%	100%
Roraima	21	86%	95%	95%
Amapá	22	95%	95%	95%
Pará	73	81%	84%	92%
Rondonia	31	65%	65%	84%
Acre	38	97%	95%	95%
Mean		87%	89%	93%

Source: Prepared by author.

3.3.4 Timber assignment test

The following analyses were developed by assigning individual timber samples to the country and Brazilian states reference levels, separately. Due to the age of the timber samples used (~3 years) the amplification rate was lower (67.3%) in comparison with cambium and leaf material (over 95%), with only one locus amplified for all 20 samples (0000846). Samples with less than 40% of analysable data (27 loci) were excluded from the analysis and a final number of eight samples were tested. In the country level the Bayesian and frequency methods achieved similar results, both with 100% of correct assignment to Brazil (Table 9). All score-values were significant (P> 0.99), except for the sample T194 that presented non-significant score-values with the Bayesian (P= 0.688) and frequency (P= 0.683) method. The nearest neighbourhood method differently from the others two methods obtained a not significant index (I_r= 38%) for samples assigned to Brazil (Table 10).

Table 10. Assignment for timber samples among countries in *Jacaranda copaia* population, using the Bayesian, frequency and nearest neighbour methods. n_l is the number of loci; Prob: probability of exclusion based on LOD1-values; E_p is the exclusion probability (Pr) based on outlier genotypes; I_r is the proportion index.

		Bayesian method				Frequency method				Nearest neighbourhood method		
Sample	n_l	Assigned country	LOG (L)	Score	Pr	Assigned country	LOG (L)	Score	Pr	Assigned country	E_p	I_r
T048	30	Brazil	5.57	0.999	0.091	Brazil	5.56	0.999	0.089	F. Guyana	0.467	0.286
T055	36	Brazil	9.7	0.992	0.626	Brazil	9.7	0.993	0.641	Brazil	0.621	0.25
T128	44	Brazil	9.41	1.0	0.241	Brazil	9.39	1.0	0.238	Brazil	0.573	0.75
T140	37	Brazil	8.26	0.994	0.427	Brazil	8.26	0.995	0.451	F. Guyana	0.489	1.0***
T188	44	Brazil	11.2	0.996	0.535	Brazil	11.2	0.996	0.503	F. Guyana	0.607	1.0***
T194	32	Brazil	9.06	0.688	0.562	Brazil	9.07	0.683	0.555	F. Guyana	0.697	0.714***
T216	41	Brazil	10.04	0.996	0.409	Brazil	10.0	0.997	0.448	F. Guyana	0.62	1.0***
T229	35	Brazil	5.48	0.998	0.028	Brazil	5.46	0.998	0.025	Brazil	0.43	0.5

***P< 0.001;

Source: Prepared by author.

The assignment test in Brazilian states level (Table 11) presented similar results as those obtained for country level assignment. *H. impetiginosus* timber samples were, for all three methods, mainly assigned to Rondônia that is the Brazilian state that presents declared *H. impetiginosus* samples. Timber samples were also assigned to the Amazonas state and Pará (correct state of origin). The samples T048 (*T. serratifolius*) and T229 (*H.impetiginosus*) were assigned by the three methods as being from Amazonas state (significant with the probability> 0.95) with low exclusion probabilities. The third sample (T194- *H. impetiginosus*) assigned to Amazonas state with the Bayesian and frequency methods presented a high score (BM: 0.995; FM: 1.0), however it also obtained a high exclusion probability significant with probability > 0.95. This sample was significantly (P< 0.99) assigned with the nearest neighbourhood approach to Rondônia with an index of 1.0 and not significant exclusion probability with at 0.90, presumably due to the species.

Table 11. Assignment for timber samples among Brazilian states by employing the Bayesian method, the frequency method and the nearest neighbour approach. Pr: probability of exclusion based on LOD1-values; Example, Pr: exclusion probability based on outlier genotypes; I_r : proportion index.

		Bayesian method				Frequency method				Nearest neighbourhood method		
Sample	n_l	Assigned state	LOG (L)	Score	Pr	Assigned state	LOG (L)	Score	Pr	Assigned state	Pr	I_r
T048	30	Amazonas	1.99	0.988	0	Amazonas	1.87	0.989	0.12	Amazonas	0.644	1.0***
T055	36	Rondônia	10.57	0.572	0.997	Rondônia	11.32	0.495	0.997	Para	0.641	0.333
T128	44	Pará	9.45	1.0	0.472	Pará	9.44	1.0	0.43	Para	0.609	1.0**
T140	37	Rondônia	5.43	0.994	0.666	Rondônia	5.41	0.999	0.683	Rondonia	0.623	0.8***
T188	44	Rondônia	8.16	1.0	0.774	Rondônia	8.19	1.0	0.773	Amazonas	0.681	0.667*
T194	32	Amazonas	6.41	0.995	0.973	Amazonas	6.55	1.0	0.978	Rondonia	0.729	1.0***
T216	41	Rondônia	7.87	0.983	0.837	Rondônia	7.95	0.988	0.794	Rondonia	0.698	0.889***
T229	35	Amazonas	3.00	0.970	0.333	Amazonas	2.88	0.971	0.32	Amazonas	0.622	0.5*

***P< 0.001; **P< 0.01.

Source: Prepared by author.

3.4 DISCUSSION

3.4.1 Genetic diversity

Illegal timber trade, eitherin the species context as well as illegally sourced has being a major problem in tropical forest. Our study shows the potential of the DNA fingerprint to track country, population, and region within country origin, as well as to follow and verify the chain of custody of timber products of *J. copaia*.

This is the first study reporting the genetic diversity and population structure in *J. copaia* for SNP markers. Our study displayed only moderate levels of genetic diversity for the nSNP, CpSNP, and MtSNP markers of *J. copaia*. Although, the observed (H_o) and expected (H_e) heterozygosity were low, the means were within the reported pattern in other Neotropical trees for SNPs markers, where H_o has been observed ranging from 0.03 to 0.39 and H_e ranging from 0.02 to 0.336 (Table S4). The genetic diversity in *J. copaia* was especialy low within some populations, due to the low sample size, as indicated by the positive Spearman correlation (ρ) significantly between the sample size (n) and total number of alleles (K), as well as between the K and H_o, and K and H_e. The H_e values were also higher for nSNPSs than for CpSNPs, MtSNPs, and CpMtSNPs. This can be attributed to the highest polymorphism of nSNP loci.

The estimate of intrapopulation fixation index (F) indicated inbreeding in some populations. Inbreeding has also been observed in other Neotropical trees for SNPs markers (Table S4). However, due to the fact that *J. copaia* is self-incompatible (MAUES *et al*., 2008) and our samples within populations were tacked from geographic distant trees, the observed inbreeding is very probably an artefact of the Wahlund effect (WAHLUND, 1928) due to the mixtures of samples from different subpopulations.

3.4.2 Population genetic differentiation

The presence of population genetic differentiation and intrapopulation spatial genetic structure (SGS) or in another terms the ocurrence of isolation by distance patterns (IBD) is key to assigned timber from different origins (TNAH *et al*., 2009; OGDEN; LINACRE, 2015) , such us countries, populations and regiosn within countries. High genetic differentiation among the different genetic groups increases the success of genetic assignment (OGDEN; LINACRE, 2015). Our results showed strong genetic diferentiation among all populations, countries, SGS and the IBD pattern. The standardised G'_{ST} for all SNPs, nSNPs, and CpMtSNps were high, especially for CpMtSNps, with exception for some populations where the SNPs were no polymorphic, such as French Guiana, Brazilian East, Peru North and South population, so

supporting a strong power of the SNPs markers for determining the correct population of timber origin. The genetic diferentiation was higher among all population than among countries. Within countries, for all SNPs markers, G'_{ST} was highest among Brazilian, and Peru population and between regions within countries, G'_{ST} was highest in Brazilian West and Peru South populations. These results indicates that due to the fact that the group/populations-specific alleles among populations be more common than among countries or between some regions within countries. These results indicate strong potential for assigned timber between populations, countries and for Brazilian West and Peru South regions. The genetic differentiation is in agreement with the structure found for other neotropical tree species, such as *Carapa* spp. (SCOTTI-SAINTAGNE *et al*., 2013a) and *J. copaia* (SCOTTI-SAINTAGNE *et al*., 2013a).

3.4.3 Genetic assignment and practical application

The results show a high power of correct group assignment in all three levels: among countries, population, and region within Brazilian and Peru countries. All three levels attained the higher success rate (> 90%) for all SNPs and nSNPs, confirming that this approach have the success expectation in imperfect grouping of reference samples. This success can be attributed the general great genetic differentiation among countries, populations and regions within countries. Even in case where the differentiation was only moderated (G'_{ST}< 0.3) (French Guiana, Bolivia, Brazil East, Brazil West *versus* East, Peru North, and Peru North *versus* South), differences in allele frequencies between groups of analyses were enough to produce high scores for the correct origin.

The results for individual assignment test for all SNPs and nSNPs were lower than group, but also show high power of correct assignment in among countries, population, and region within regions in Brazilian and Peru. Peru and its regions showed the low scores rate of assignments. According with Spearman correlation (ρ) for countries and populations for assignment of all SNPs and nSNPs *versus* K, H_o, and H_e, and for population level for assignment of nSNPs *versus* F, the H_o and H_e, following by K are the most determinant indices to have success in the individual assignment tests. Peru North population generally showed the lowest H_o and H_e values, and highest inbreeding with can explain the lows cores rate of assignment. However, in general, the results for individual assignment test for all SNPs and nSNPs support that this approach have power determination of the specific site of timber origin. For to use CpMtSNP loci markers, it is necessary the development of a great number of loci to increase both grouped and individual assignment tests.

Similar results of higher group that individual assignments tests were reported for *Hymenaea* sp. (CHAVES *et al*., 2019) and *Handroanthus* sp. (MEYER-SAND *et al*., 2018). For *Swietenia macrophylla*, based on nSSR loci, assignment test was higher at the country (82%) than population (71) level (DEGEN *et al*., 2013). For the Malaysia *Gonystylus bancanus* tree, the self-assignment rate to a set of 16 nSSRs was in lower (55%) than has been observed here on the population level (NG *et al*., 2016). For SNP data of *Entandrophragma cylindricum*, the assignment on the country level ranged from 66 to 74%, depending on the assignment method (DEGEN *et al*., 2017). Many Other studies in tropical, African and European tree species has been developed microsatellite and SNPs markers to timber tracking and the main conclusion is that due to the presence of spatial genetic structure and genetic differentiation DNA fingerprint is the most indicated tool to tracking the country and population origin (JARDINE *et al*., 2016; PAKULL *et al*., 2016; SCHROEDER *et al*., 2016; PAREDES-VILLANUEVA *et al*., 2019; TYSKLIND *et al*., 2019; PAKULL *et al*., 2020).

Finally, we suggest for the timber sector to add such genetic controls as an independent audit on top of paper-based proofs of the chain of custody (LOWE *et al*., 2016). It is important to note that the power of the genetic reference data to detect false declarations reaches 100% if more than one sample with the same declaration was tested.

3.4.4 Genotype self-assignment tests

The results obtained show a high power of correct assignment in all three levels: among countries, locations and Brazilian states. All three levels attained the higher success rate of correct self-assignment by employing the nearest neighbour approach with 96% among countries, 85% among locations and 93% among Brazilian states. This approach confirmed the success expectation in imperfect grouping of reference samples.

At the location level, the sample-population of Bajo Paraguá in Bolivia presented a 0% of corrected self-assignment rate for both allele frequency base methods (Bayesian and Frequency methods); nevertheless its samples were assigned to 90% to Roboré and 10% to Yapacaní-Naranjal, both in Bolivia. Self-assignment of samples to Brazilian states, Rondônia state had the major percentage of incorrect assignments for all three approaches. This occurred possibly due to the mix of species, due to the fact that *H. impetiginosus* and *H. incanus* might occur in other locations that were botanically identified only in the genus level.

3.4.5 Timber assignment tests

In the analysis in the country level most of the timber samples were incorrectly assigned to French Guyana. This fact can be explained by the mixed species of the timber samples and sample-populations with *H. impetiginosus*, *H. incanus* and *H. serratifolius*. Nineteen percent of the samples from French Guyana (12 individuals) are *H. impetiginosus*, against 9% of the Brazilian samples (20 individuals), associated with the small number of individuals from French Guyana (63) in comparison with Brazil (221) and Bolivia (115).

3.5 CONCLUSION

The genetic differentiation (G'_{ST}) among all populations, contries, and regions within coutries is generaly high, specialy for CpMtSNP loci, and there is a strong isolation by distance pathern among populations, favoring the group or individual samples tracking to correct site. For self-assignment tests, we were able to 100% correct determine country, population and region site origin of all samples using all SNPs and nSNPs. Our results show that the use of all SNP or nSNP markers are suitable to correct determination of country and population site of *J. copaia* timber origin and very useful tool for customs and local and international policies. The *J. copaia* reference database of our study represents a robust assignment tool available to timber companies or governmental agencies to test and validate origin declarations. It is recommended to use the method described here for other tropical native species, since it presents high efficiency when it comes to showing the origin of the wood, thus helping the police and the competent bodies in the delimitation of illegally deforested areas, as well as unsustainable extraction.

3.6 REFERENCES

ABEELE, S. V.; HARDY, O. J.; BEECKMAN, H.; ILONDEA, B. A.; JANSSENS, S. B. Genetic Markers for Species Conservation and Timber Tracking: Development of Microsatellite Primers for the Tropical African Tree Species *Prioria balsamifera* and *Prioria oxyphylla*. **Forests**, [*s. l.*], v. 10, p. 1037, 2019.

BERGO, M. C. J.; PASTORE, T. C. M.; CORADIN, V. T. R.; WIEDENHOEFT, A. C.; BRAGA, J. W. B. NIRS identification of *Swietenia macrophylla* is robust across specimens from 27 countries. **IAWA Journal**, Utrecht, v. 37, p. 420–430, 2016.

BLANC-JOLIVET, C.; YANBAEV, Y.; KERSTEN, B.; DEGEN, B. A set of SNP markers for timber tracking of *Larix* spp. in Europe and Russia. **Forestry**: An International Journal of Forest Research**,** [*s. l.*], v. 91, p. 614–628, 2018.

CHAVES, C. L; BLANC-JOLIVET, C.; SEBBENN, A. M.; MADER, M.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; CORONADO, E. N. H.; GARCIA-DAVILA, C.; TYSKLIND, N.; TROISPOUX, V.; MASSOT, M.; DEGEN, B.; Nuclear and chloroplastic SNP markers for genetic studies of timber origin for *Hymenaea* trees. **Conservation Genetics Resourses**, [*s. l.*], v. 11, p. 329–331, 2019.

CORONADO, E. N. H.; BLANC-JOLIVET, C.; MADER, M.; GARCÍA-DÁVILA, C. R.; GOMERO, D. A.; CASTILLO, D. T. DEL.; LLAMPAZO, G. F.; PIZANGO, G. H.; SEBBENN, A. M.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; TYSKLIND, N.; TROISPOUX, V.; MASSOT, M.; CARVALHO, C.; LIMA, H. C.; CARDOSO, D.; DEGEN, B. SNP markers as a successful molecular tool for assessing species identity and geographic origin of trees in the economically important South American legume Genus *Dipteryx*. **Journal of Heredity**, Oxford, v. 111, p. 346–356, 2020.

COURNUET, J. M.; PIRY, S.; LUIKART, G.; ESTOUP, A.; SOLIGNAC, M. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. **Genetics**, Austin, v. 153, p. 1989–2000, 1999.

DEGEN, B.; WARD, S. E.; LEMES, M. R.; NAVARRO, C.; CAVERS, S.; SEBBENN, A. M. Verifying the geographic origin of mahogany (*Swietenia macrophylla* King) with DNA-fingerprints. **Forensic Science International**: Genetics, Amsterdam, v. 7, i.1, p. 55-62, 2013.

DEGEN, B.; BLANC-JOLIVET, C.; STIERAND, K.; GILLET, K. A nearest neighbour approach by genetic distance to the assignment of individual trees to geographic origin. **Forensic Science International Genetics**, Amsterdam, v. 27, p. 132–141, 2017.

DORMONTT, E. E.; BONER, M.; BRAUN, B.; BREULMANN, G.; DEGEN, B.; ESPINOZA, E.; GARDNER, S.; GUILLERY, P.; HERMANSON, J. C.; KOCH, G.; LEE, S. L.; KANASHIRO, M.; RIMBAWANTO, A.; THOMAS, D.; WIEDENHOEFT, A. C.; YIN, Y.; ZAHNEN, J.; LOWE, A. J. Forensic timber identification: It's time to integrate disciplines to combat illegal logging. **Biological Conservation**, Oxford, v. 191, p. 790-798, 2015.

DUMOLIN, S.; DEMESURE, B.; PETIT, R. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. **Theoretical and Applied Genetics**, Berlin, v. 91, i.8, p. 1253-1256, 1995.

DURAND, S. R.; SHELLEY, P. H.; ANTWEILER, R. C.; TAYLOR, H. E. Trees, chemistry, and prehistory in the American Southwest. **Journal of Archaeological Science,** New York, v. 26, p. 185–203, 1999.

EFRON, B. Estimating the error rate of a prediction rule – improvement on cross validation. **Journal of American Statistic Association**, [*s. l.*], v. 78, p. 316–331, 1983.

FEDERAL POLICE BRAZIL, **GOV**, 2011. Disponível em: https://www.gov.br/pf/pt-br/search?SearchableText=madeira%20ilegal%20extra%C3%ADda%20da%20Amaz%C3%B4nia. Acesso em: 17 maio 2022.

FIDELIS, C. H. V.; AUGUSTO, F.; SAMPAIO, P. T. B.; KRAINOVIC, P. M.; BARATA, L. E. S. Chemical characterization of rosewood (*Aniba rosaeodora* Ducke) leaf essential oil by comprehensive two-dimensional gas chromatography coupled with quadrupole mass spectrometry. **Journal of Essential Oil Resources,** [*s. l.*], v. 24, p. 245–251, 2012.

GASSON, P.; BAAS, P.; WHEELER, E. Wood anatomy of Cites-listed tree species. **IAWA Journal**, Utrecht, v. 32, p. 155–198, 2011.

GOUDET, G. Fstat (Version 2.9.3.2.): a computer program to calculate *F*-statistics. **Journal of Heredity**, Oxford, v. 86, 485–486, 1995.

GRAY, J. A. Forest concession policies and revenue systems: country experience and policy changes for sustainable tropical forestry, World Bank Publications, 2002.

HARDY, O.; VEKEMANS, X. SPAGeDI: a versatile computer program to analyze spatial genetic structure at the individual or population levels. **Molecular Ecology Notes**, Oxford, v. 2, p. 618–620, 2002.

HEDRICK, P. W. A standardized genetic differentiation measure. **Evolution**, Oxford, v. 59, i.8, p. 1633–1638, 2005.

JARDINE, D. I.; BLANC-JOLIVET, C.; DIXON, R. R. M.; DORMONTT, E. E.; DUNKER, B.; GERLACH, J.; KERSTEN, B.; VAN DIJK, K.J.; DEGEN, B.; LOWE, A. J. Development of SNP markers for Ayous (*Triplochiton scleroxylon* K. Schum) an economically important tree species from tropical West and Central Africa. **Conservation Genetic Resourses,** [*s. l.*], v. 8, p. 129–139, 2016.

JOLIVET, C.; HÖLTKEN, A. M.; LIESEBACH, H.; STEINER, W.; DEGEN, B. Spatial genetic structure in wild cherry (*Prunus avium* L.): I. variation among natural populations of different density. **Tree Genetics & Genomes**, Heidelberg, v. 7, i.2, p. 271-283, 2011.

JOLIVET, C.; DEGEN, B. Use of DNA fingerprints to control the origin of sapelli timber (*Entandrophragma cylindricum*) at the forest concession level in Cameroon. **Forensic Science International Genetic,** [*s. l.*], v. 6, p. 487–493, 2012.

KAGAWA, A.; LEAVITT, S. W. Stable carbon isotopes of tree rings as a tool to pinpoint the geographic origin of timber. **Journal of Wood Science,** [*s. l.*],v. 56, p. 175–183, 2010.

LESCUYER, S.; NDOTIT, L. B. B.; NDONG, R.; TSANGA, P. O. CERUTTI, POLICY options for improved integration of domestic timber markets under the voluntary partnership agreement (VPA) regime in Gabon. **Center for International Forestry Research** (CIFOR), 2014.

LOISELLE, B. A.; SORK, V. S.; NASON, J.; GRAHAM, C. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). **American Journal of Botany**, Lancaster, v. 82, p. 1420-1425, 1995.

LOWE, A. J.; WONG, K.-N.; TIONG, Y.-S.; IYERH, S.; CHEW, F.-T. A DNA method to verify the integrity of timber supply chains; confirming the legal sourcing of merbau timber from logging concession to sawmill. **Silvae Genetica**, Frankfurt, v. 59, i.6, p. 263-268, 2010.

LOWE, A. J.; DORMONTT, E. E.; BOWIE, M. J.; DEGEN, B.; GARDNER, S.; THOMAS, D.; CLARKE, C.; RIMBAWANTO, A.; WIEDENHOEFT, A. Y. Yin. Opportunities for improved transparency in the timber trade through scientific verification. **BioScience,** Washington, v. 66, p. 990–998, 2016.

MAUES, M. M.; OLIVEIRA, P. E. A. M.; KANASHIRO, M. Pollination biology in *Jacaranda copaia* (Aubl.) D. Don. (Bignoniaceae) at the "Floresta Nacional do Tapajós", Central Amazon. **Revista Brasileira de Botany,** [*s. l.*], v 31, p. 517-527, 2008.

MEYER-SAND, B. R. V.; BLANC-JOLIVET, C.; MADER, M.; PAREDES-VILLANUEVA, L.; TYSKLIND, N.; SEBBENN, A. M.; GUICHOUX, E.; DEGEN, B. Development of a set of SNP markers for population genetics studies of Ipe (*Handroanthus* sp.), a valuable tree genus from Latin America. **Conservation Genetic Resourses,** [*s. l.*], v. 10, p. 779–781, 2018.

MOYA, R.; WIEMANN, M. C.; OLIVARES, C. Identification of endangered or threatened Costa Rican tree species by wood anatomy and fluorescence activity. **Revista de Biologia Tropical,** San José, v. 61, p. 1113–1156, 2013.

NG, K. K.; LEE, S. L.; TNAH, L. H.; NURUL-FARHANAH, Z.; NG, C. H.; LEE, C. T.; TANI, N.; DIWAY, B.; LAI, P. S.; KHOO, E. Forensic timber identification: a case study of a cites listed species, *Gonystylus bancanus* (thymelaeaceae). **Forensic Science International Genetics**, [*s. l.*], v. 23, p. 197-209, 2016.

OGDEN, R.; LINACRE, A. Wildlife forensic science: A review of genetic geographic origin assignment. **Forensic Science International Genetic**, [*s. l.*], v. 18, p. 152–159, 2018.

PAKULL, B.; MADER, M.; KERSTEN, B.; EKUE, M. R. M.; DIPELET, U. G. B.; PAULINI, M.; BOUDA, Z. H. N.; DEGEN, B. Development of nuclear, chloroplast and mitochondrial SNP markers for *Khaya* sp. **Conservation Genetic Resourses,** [*s. l.*], v. 8, p. 283–297, 2016.

PAKULL, B.; SCHINDLER, L.; MADER, M.; KERTEN, B.; BLANC-JOLIVET, C.; PAULINI, M.; LEMES, M. R.; WARD, S. E.; NAVARRO, C. M.; CAVERS, S.; SEBBENN, A. M.; DIO, O. DI; GUICHOUX, E.; DEGEN, B. Development of nuclear SNP markers for Mahogany (*Swietenia* spp.). **Conservation Genetic Resourses,** [*s. l.*], v. 12, p. 585–587, 2020.

PAREDES-VILLANUEVA, K.; ESPINOZA, E.; OTTENBURGHS, J.; STERKEN, M. G.; BONGERS, F.; ZUIDEMA, P. A. Chemical differentiation of Bolivian *Cedrela* species as a tool to trace illegal timber trade. **Forestry: An International Journal of Forest Research**, [s. l.], v. 91, p. 603–613, 2018.

PAREDES-VILLANUEVA, K.; GROOT, G.A. DE.; LAROS, I.; BOVENSCHEN, J.; BONGERS, F.; ZUIDEMA, P. A. Genetic differences among *Cedrela odorata* sites in Bolivia provide limited potential for fine-scale timber tracing. **Tree Genetics & Genomes,** v 15, p. 33, 2019.

PAREDES-VILLANUEVA, K.; BLANC-JOLIVET, C.; MADER, M.; CORONADO, E. N. H.; GARCIA-DAVILA, C.; SEBBENN, A. M.; MEYER-SAND, B. R. V.; CARON, H.; TYSKLIND, N.; CAVERS, S.; DEGEN, B. Nuclear and plastid SNP markers for tracing *Cedrela* timber in the tropics. **Conservation Genetic Resources**, v. 12, p. 239-244, 2020.

PE, M.; SANZ, J.; CABEZUDO, M. Analysis of volatile components of oak wood by solvent extraction and direct thermal desorption-gas chromatography-mass spectrometry. **Journal of Chromatography**, Amsterdam, v. 778, p. 427–434, 1997.

PIRY, S.; ALAPETITE, A.; CORNUET, J. M.; PAETKAU, D.; BAUDOUIN, L.; ESTOUP, A. GENECLASS2: a software for genetic assignment and first-generation migrant detection. **Journal of Heredity**, Cary, v. 95, p. 536–539, 2004.

RANNALA, B.; MOUNTAIN, J. L. Detecting immigration by using multilocus genotypes. **Proceedings of the National Academy of Sciences of the United State of America**, Washington, v. 94, i.17, p. 9197–9201, 1997.

SCHROEDER, H.; CRONN, R.; YANBAEV, Y.; JENNINGS, T.; MADER, M.; DEGEN, B.; KERSTEN, B. Development of molecular markers for determining continental origin of wood from white oaks (*Quercus* L. sect. *Quercus*). **PloS one,** [*s. l.*], v. 11, p. 158-221. 2016.

SCOTTI-SAINTAGNE, C.; DICK, C. W.; CARON, H.; VENDRAMIN, G. G.; GUICHOUX, E.; BUONAMICI, A.; DURET, C.; SIRE, P.; VALENCIA, R.; LEMES, M. R.; GRIBEL, R.; SCOTTI, I. Phylogeography of a species complex of lowland Neotropical rain forest trees (*Carapa*, Meliaceae). **Journal of Biogeography,** Oxford, v. 40, p. 676–692, 2013a.

SCOTTI-SAINTAGNE, C.; DICK, C. W.; CARON, H.; VENDRAMIN, G. G.; TROISPOUX, V.; SIRE, P.; CASALIS, M.; BUONAMICI, A.; VALENCIA, R.; LEMES, M. R. *et al*., Amazon diversification and cross-Andean dispersal of the widespread Neotropical tree species *Jacaranda copaia* (Bignoniaceae). **Journal of Biogeography,** Oxford, v. 40. p. 707–719. 2013b.

SEBBENN, A. M.; BLANC-JOLIVET, C.; MADER, B.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; CORONADO, E. N. H.; GARCIA-DAVILA, C.; TYSKLIND, N.; TROISPOUX, V.; DELCAMP, A.; DEGEN, B. Nuclear and plastidial SNP and INDEL markers for genetic tracking studies of *Jacaranda copaia*. **Conservation Genetic Resources**, [*s. l.*], v. 11, p. 341–343, 2019.

TNAH, L. H.; LEE, S. L.; NG, K. K. S.; TANI, N.; BHASSU, S.; OTHMAN, R. Y. Geographical traceability of an important tropical timber (*Neobalanocarpus heimii*) inferred from chloroplast DNA. **Forest Ecology and Management,** Amsterdam, v. 258, p. 1918–1923, 2009.

TNAH, L. H.; LEE, S. L.; NG, K. K. S.; FARIDAH, Q. Z.; FARIDAH-HANUM, I. Forensic DNA profiling of tropical timber species in Peninsular Malaysia. **Forest Ecology and Management**, Amsterdam, v. 259, p. 1436–1446, 2010.

TYSKLIND, N.; BLANC-JOLIVET, C.; MADER, M.; MEYER-SAND, B. R. V.; PAREDES-VILLANUEVA, K.; CORONADO, E. N. H.; GARCIA-DAVILA, C.; SEBBENN, A. M.; CARON, H.; TROISPOUX, V.; GUICHOUX, E.; DEGEN, B. Development of nuclear and plastid SNP and INDEL markers for population genetic studies and timber traceability of *Carapa* species. **Conservation Genetic Recourses,** [*s. l.*], v. 11, p. 337–339, 2019.

VLAM, M.; GROOT, G. A. DE.; BOOM, A.; COPINI, P.; LAROS, I.; VELDHUIJZEN, K. *et al*. Developing forensic tools for an African timber: Regional origin is revealed by genetic characteristics, but not by isotopic signature. **Biological Conservation,** [*s. l.*], v. 220, p. 62–271, 2018.

WAHLUND, S. Composition of populations and correlation appearances viewed in relation to the studies of inheritance. **Hereditas**, [*s. l.*], v. 11, p. 65–106, 1928.

3.7 SUPPLEMENTARY MATERIAL

Table S1. Spearman correlation coefficient (ρ) between sample size (n) and genetic diversity indices.

SNPs	Indices	ρ	P value
nSNPs	$n\ vs\ K$	0.36	0.026
CpSNPs	$n\ vs\ K$	0.119	0.477
MtSNPs	$n\ vs\ K$	-0.08	0.633
CpMtSNPs	$n\ vs\ K$	0.034	0.839
nSNPs	$n\ vs\ Ho$	-0.188	0.258
nSNPs	$n\ vs\ F$	0.058	0.729
nSNPs	$n\ vs\ He$	-0.259	0.116
CpSNPs	$n\ vs\ He$	0.163	0.329
MtSNPs	$n\ vs\ He$	0.102	0.504
CpMtSNPs	$n\ vs\ He$	-0.031	0.853
nSNPs	$k\ vs\ Ho$	0.675	0
nSNPs	$k\ vs\ F$	-0.213	0.199
nSNPs	$k\ vs\ He$	0.691	0
CpSNPs	$k\ vs\ He$	-0.297	0.07
MtSNPs	$k\ vs\ He$	-0.037	0.823
CpMtSNPs	$k\ vs\ He$	-0.171	0.306

K is the total number of alleles; H_o is the observed heterozygosity; H_e is the expected heterozygosity; F is the fixation index. Source: Prepared by author.

Table S2. Pairwise genetic differentiation ($G_{ST}^{'}$) between populations for all SNPs.

	Brazil	Peru	Bolivia
French Guiana	0.249 (0.035)*	0.415 (0.057)*	0.565 (0.06)*
Brazil		0.368 0.053)*	0.3 (0.045)*
Peru			0.324 (0.043)*
Brazil West		0.306 (0.045)*	0.198 (0.036)*
Brazil East		0.492 (0.06)*	0.458 (0.054)*
Peru North	0.504 (0.045)*		0.531 (0.046)*
Peru South	0.294 (0.047)*		0.149 (0.045)*

*P< 0.05. Source: Prepared by author.

Table S3. Mean genetic differentiation among populations and among countries. δ_G is the Gregorius δ_G -statistic; F_{ST} is the Wright's fixation index; G'_{ST} is the Herdick standardised G'_{ST} -statistic.

	Among populations			Among countries		
Locus	δ_G	F_{ST}	G'_{ST}	δ_G	F_{ST}	G'_{ST}
tas0000082	0.248	0.363	0.475	0.045	0.007	0.012
tas0000119	0.113	0.284	0.316	0.084	0.074	0.085
tas0000134	0.075	0.196	0.213	0.038	0.022	0.023
tas0000211	0.066	0.141	0.155	0.064	0.044	0.051
tas0000294	0.233	0.307	0.405	0.042	0.006	0.01
tas0000443	0.073	0.231	0.248	0.082	0.069	0.08
tas0000496	0.1	0.374	0.405	0.14	0.13	0.16
tas0000513	0.077	0.13	0.143	0.036	0.014	0.016
tas0000541	0.09	0.3	0.327	0.073	0.064	0.072
tas0000637	0.394	0.657	0.79	0.092	0.022	0.047
tas0000838	0.248	0.386	0.505	0.125	0.045	0.078
tas0000846	0.316	0.512	0.671	0.253	0.147	0.301
tas0000880	0.334	0.485	0.641	0.135	0.044	0.092
tas0000984	0.22	0.327	0.424	0.055	0.013	0.02
tas0000993	0.197	0.568	0.644	0.216	0.161	0.269
tas0001084	0.18	0.207	0.279	0.106	0.042	0.064
tas0001089	0.275	0.379	0.505	0.057	0.01	0.018
tas0001118	0.08	0.115	0.132	0.033	0.013	0.015
tas0001119	0.291	0.398	0.583	0.101	0.02	0.05
tas0001135	0.09	0.258	0.281	0.092	0.075	0.089
tas0001208	0.394	0.644	0.782	0.05	0.008	0.017
tas0001231	0.2	0.391	0.473	0.085	0.031	0.045
tas0001289	0.383	0.756	0.848	0.229	0.12	0.234
tas0001361	0.115	0.402	0.435	0.137	0.116	0.142
tas0001399	0.261	0.364	0.48	0.067	0.017	0.029
tas0001470	0.243	0.738	0.799	0.338	0.402	0.525
tas0001562	0.138	0.169	0.215	0.041	0.011	0.015
tas0001653	0.096	0.223	0.256	0.046	0.017	0.023
tas0001740	0.238	0.376	0.482	0.075	0.02	0.032
tas0001766	0.138	0.106	0.194	0.076	0.012	0.03
tas0001767	0.324	0.482	0.628	0.055	0.009	0.018
tas0001796	0.112	0.201	0.235	0.078	0.04	0.051
tas0001816	0.255	0.379	0.53	0.103	0.027	0.055
tas0001827	0.287	0.472	0.595	0.123	0.049	0.086
tas0001921	0.292	0.458	0.598	0.094	0.03	0.055
tas0001943	0.026	0.039	0.041	0.008	0.001	0.002
tas0001945	0.236	0.312	0.415	0.06	0.013	0.022

continued on next page

Table S3. (continued)

	Among populations			Among countries		
Locus	δ_G	F_{ST}	G'_{ST}	δ_G	F_{ST}	G'_{ST}
tasCp20_12116	0.498	0.888	0.944	0.329	0.248	0.432
tasCp20_16837	0.475	0.81	0.902	0.208	0.109	0.235
tasCp229_2616	0.278	0.926	0.947	0.254	0.286	0.368
tasCp229_8690	0.333	0.787	0.856	0.219	0.168	0.268
tasCp373_68	0.154	1.0	1.0	0.269	0.311	0.4
tasCp373_606	0.278	0.926	0.947	0.246	0.28	0.357
tasCp436_3740	0.336	0.798	0.865	0.219	0.168	0.268
tasCp436_10452	0.278	0.926	0.947	0.246	0.28	0.357
tasCp544_19217	0.333	0.792	0.86	0.219	0.168	0.268
tasCp544_19813	0.395	0.702	0.81	0.084	0.017	0.038
tasCp612_1890	0.278	0.926	0.947	0.246	0.28	0.357
tasCp784_4541	0.154	1.0	1.0	0.275	0.319	0.411
tasCpIN544_750	0.279	0.949	0.963	0.24	0.272	0.346
tasMt53_18528	0.176	0.117	0.261	0.141	0.018	0.052
tasMt79_14230	0.174	0.691	0.742	0.146	0.143	0.173
tasMt79_21228	0.185	0.784	0.82	0.263	0.274	0.367
tasMt173_2622	0.243	0.679	0.753	0.093	0.028	0.046
tasMt615_5988	0.41	0.704	0.827	0.449	0.488	0.739
tasMt680_8937	0.074	0.46	0.48	0.127	0.136	0.159
tasMt732_3426	0.33	0.641	0.76	0.166	0.065	0.13
tasMt967_4329	0.191	0.599	0.668	0.134	0.11	0.14
tasMt1438_1478	0.304	0.669	0.766	0.163	0.106	0.157
tasMt1580_4320	0.182	0.732	0.777	0.26	0.264	0.357
tasMt1751_2224	0.185	0.784	0.82	0.263	0.274	0.367
tasMt2011_567	0.153	0.851	0.874	0.136	0.146	0.173
tasMt2580_2831	0.185	0.784	0.82	0.257	0.265	0.355
tasMt4446_1048	0.182	0.732	0.777	0.254	0.256	0.345
tasMt4446_1728	0.174	0.694	0.743	0.145	0.142	0.172
tasMt6079_1485	0.228	0.528	0.63	0.131	0.1	0.134
tasMtIN173_1921	0.182	0.73	0.775	0.259	0.262	0.355
tasMtIN2369_2284	0.185	0.784	0.82	0.263	0.274	0.367
MEAN	0.224	0.537	0.607	0.15	0.121	0.171

Source: Prepared by author.

Table S4. Results of the self-assignment tests to the population of origin for all individuals for all SNPs, nSNPS and CpMtSNPs.

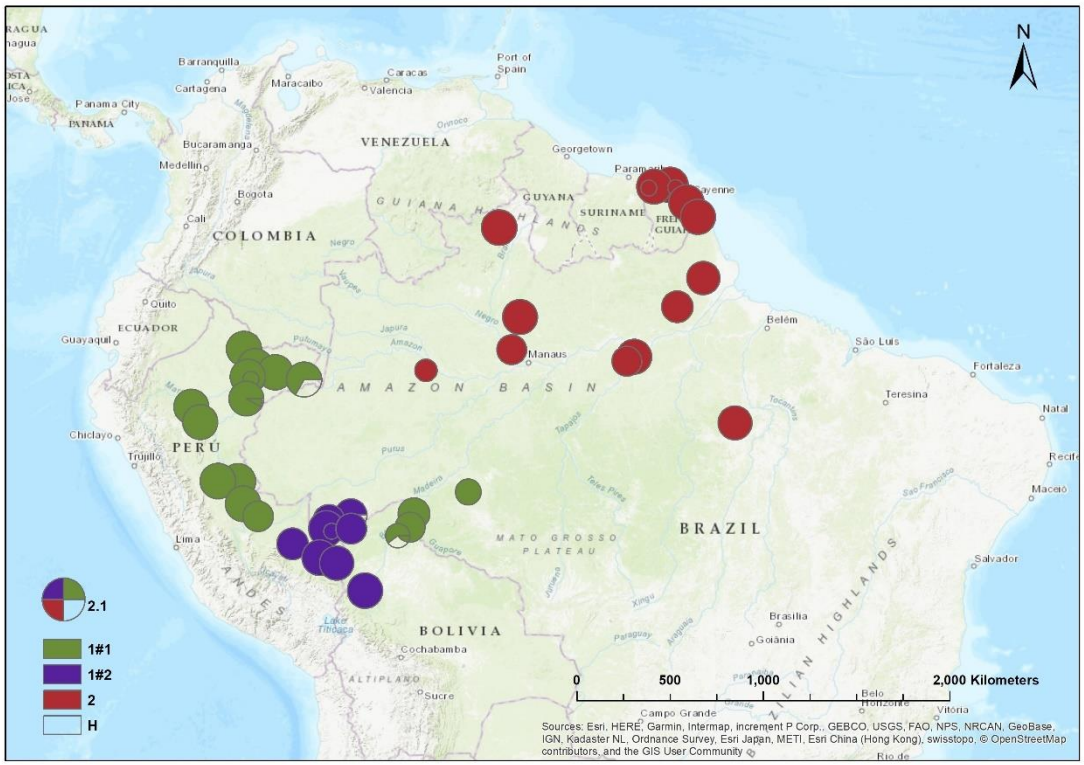
		All SNPs (128)	nSNPs (113)	CpMts (15)	
Assigned sample	Rank	Score (%)	Score (%)	Rank	Score (%)
French Guiana	French Guiana	100	100	French Guiana	100
Brazil West	Brazil West	100	100	Brazil West	100
Brazil East	Brazil East	100	100	Brazil East	93.9
Peru North	Peru North	100	100	Peru North	99.8
Peru South	Peru South	100	100	Peru South	100
Bolívia	Bolivia	100	100	Bolívia	100
Correct		100	100		100
1-FG-Counami	1-FG-Counami	100	100	3-FG-Cocoa	55.9
2-FG-Isnard	2-FG-Isnard	100	100	3-FG-Cocoa	41.3
3-FG-Cocoa	3-FG-Cocoa	100	100	3-FG-Cocoa	58.7
4-FG-Maripa	4-FG-Maripa	100	100	4-FG-Maripa	79.7
5-BRW-Maraca	5-BRW-Maraca	100	100	5-BRW-Mara	99.2
6-BRW-Anauá	6-BRW-Anauá	100	100	6-BRW-Anauá	38.0
7-BRW-Jamari	7-BRW-Jamari	100	100	34-BO-MABET	44.3

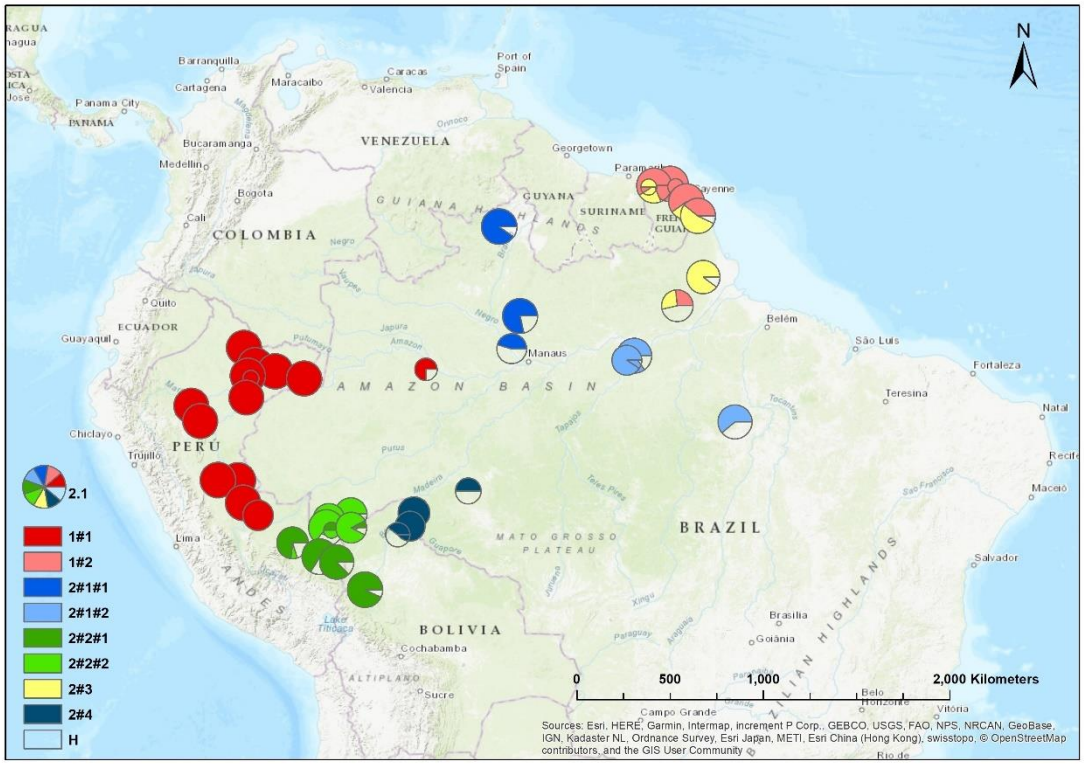
8-BRW-Jarí	8-BRW-Jarí	100	100	8-BRW-Jarí	26.7
9-BRW-Xapuri	9-BRW-Xapuri	100	100	30-PES-Madera	19.9
10-BRW-Cumaru	10-BRW-Cumaru	100	100	30-PES-Madera	28.6
11-BRE-Amapá	11-BRE-Amapá	100	100	6-BRW-Anauá	35.5
12-BRE-Avilhanas	12-BRE-Avilhanas	100	100	6-BRW-Anauá	30.5
13-BRE-Tapajós	13-BRE-Tapajós	100	100	6-BRW-Anauá	37.7
14-BRE-Arapins	14-BRE-Arapins	100	100	6-BRW-Anauá	30.5
15-BRE-Tefe	15-BRE-Tefe	100	100	6-BRW-Anauá	21.5
16-BRE-Carajás	16-BRE-Carajás	100	100	6-BRW-Anauá	36.4
17-PEN-Huiririma	17-PEN-Huiririma	100	100	17-PEN-Huiri	38.5
18-PEN-Urco	18-PEN-Urco	100	100	18-PEN-Urco	52.4
19-PEN-Madreselva	19-PEN-Madreselva	100	100	19-PEN-Madre	99.9
20-PEN-Yarina	20-PEN-Yarina	100	100	20-PEN-Yarina	59.0
21-PEN-Progresista	21-PEN-Progresista	100	100	21-PEN-Progre	100.0
22-PEN-Centre	22-PEN-Centre	100	100	27-PES-Abad	24.6
23-PEN-Herrera	23-PEN-Herrera	100	100	23-PEN-Herrera	44.2
24-PEN-Jeberos	24-PEN-Jeberos	100	100	24-PEN-Jeberos	50.0

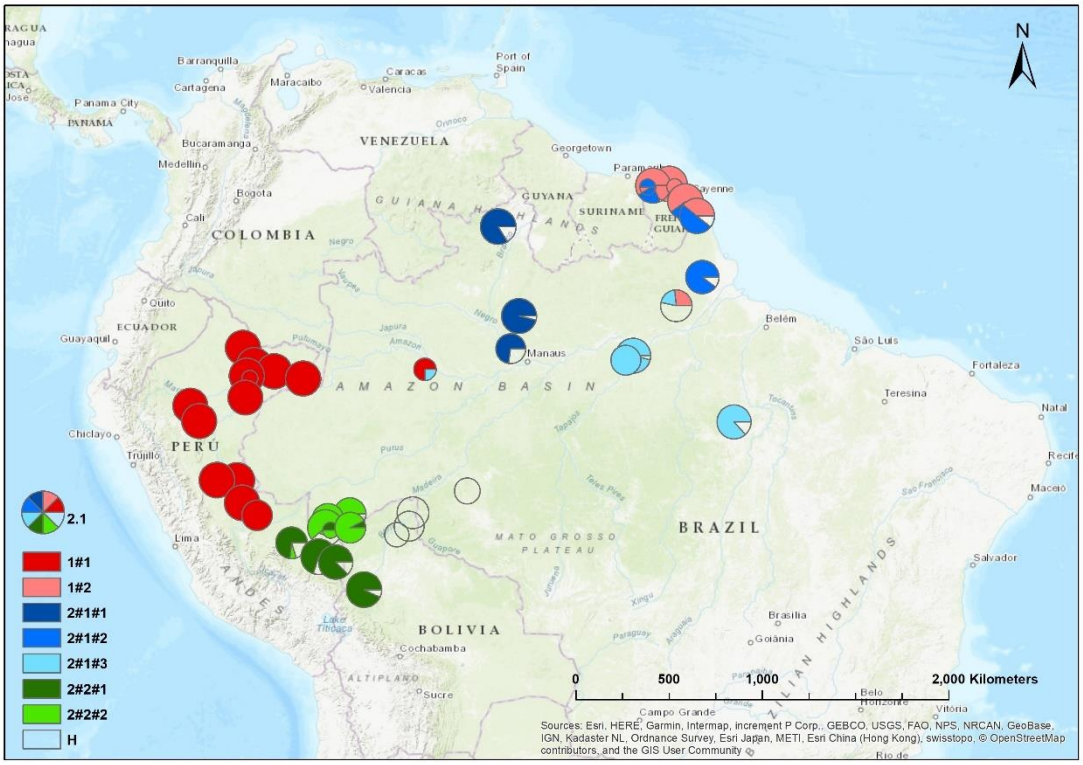
25-PEN-Shucushuyacu	25-PEN-Shucushuyacu	100	100	27-PES-Abad	34.9
26-PES-Portillo	26-PES-Portillo	100	100	26-PES-Portillo	44.3
27-PES-Abad	27-PES-Abad	100	100	27-PES-Abad	36.3
28-PES-Diaz	28-PES-Diaz	100	100	27-PES-Abad	35.3
29-PES-Inuya	29-PES-Inuya	100	100	27-PES-Abad	25.6
30-PES-Maderacre	30-PES-Maderacre	100	100	30-PES-Madera	33.7
31-PES-Cashu	31-PES-Cashu	100	100	30-PES-Madera	28.5
32-PES-Amigos	32-PES-Amigos	100	100	30-PES-Madera	33.2
33-PES-Tambopata	33-PES-Tambopata	100	100	33-PES-Sando	31.9
34-BO-35-MABET	34-BO-35-MABET	100	100	34-BO-MABET	75.4
35-BO-Desvelo	35-BO-Desvelo	100	100	34-BO-MABET	70.7
36-BO-Vista	36-BO-Vista	100	100	30-PES-Madera	27.4
37-BO-Chorro	37-BO-Chorro	100	100	37-BO-Chorro	98.0
38-BO-Madidi	38-BO-Madidi	100	100	30-PES-Madera	32.7
Correct		100	100		17

Spearman coefficient (Rho) between n and scores for CpMtSNP = 0.357 (P= 0.028). Source: Prepared by author.

Figure 1, 2 and 3 - Spatial distribution of CpMtSNPs (A), nSNPS (B) and nCpMtSNP (C) estimated by STRUCTURE (K = ?) for *Jacaranda copaia* in South America. Source: Prepared by author.







Source: Prepared by author.