CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE
FACULTY OF TROPICAL AGRISCIENCES

Genetic variation and structure of agroforestry useful trees, *Inga edulis* Mart. and *I. ingoides* (Rich.) Willd., (Fabaceae) in Amazonian Peru.

Dissertation thesis

Faculty of Tropical AgriSciences

Department of Crop Sciences and Agroforestry

Study programme: Agriculture in Tropics and Subtropics

Author: Ing. Alexandr Rollo

Supervisor: doc. Ing. Bohdan Lojka, Ph.D.

Co-supervisor: prof. Mgr. Bohumil Mandák, Ph.D.

Prague, May 2019

DISSERTATION THESIS TOPIC

Author of thesis:	Ing. Alexandr Rollo
Study programme:	Agricultural Specialization
Field of study:	Tropical and Subtropical Agriculture
Thesis supervisor:	doc. Ing. Bohdan Lojka, Ph.D.
Supervising department:	Department of Crop Sciences and Agroforestry
Language of a thesis:	English
Thesis title:	**Genetic variation and structure of agroforestry useful trees, Inga edulis Mart. and I. ingoides (Rich.) Willd., (Leguminosae: Mimosoideae) in Amazonian Peru**
Objectives of thesis:	i. To evaluate genetic variation and structure of widely cultivated neotropical tree species I. edulis in geographically different anthropogenic and natural populations from Amazonian Peru. ii. To evaluate genetic variation and structure of potentionaly useful neotropical tree species I. ingoides in geographically different natural populations from Amazonian Peru. iii. To check for putative introgression between both mentioned species.
Methodology:	The sampling will be carried out in geographically different natural forest sites and urbanized areas in Amazonian Peru. The two Inga species are going to be identified according to morphological aspects detailed in Pennington TD, (1997). The samples will be randomly selected from sexually mature trees. Voucher specimens are going to be archived in the Regional Herbarium of Ucayali IVITA-Pucallpa, Peru. Total DNA extraction from a 20 mm by 10 mm section of silica- dried young leaf material will be performed using the Invitek, Invisorb® Spin Plant Mini Kit according to the manufacture's instructions. Four microsatellite markers are going to be used, one (Pel5) previously developed by Daynandan et al. (1997) for Pithecellobium elegans Ducke, and three (Inga03, Inga08 and Inga33) by Hollingsworth et al. (2005) for I. edulis. Loci are going to be amplified individually in 10 µl reaction containing: 20 ng of template DNA, 5 µM of forward and reverse primer, 50 µM of dNTPs, 2 mM of MgCl2, 2 µl 5x GoTaq Flexi Buffer (Promega) and 1.0 U of GoTaq® Flexi DNA Polymerase (Promega). Amplifications were undertaken in Biometra® T1 Thermocycler using the following profile: 95 °C for 2 min; 95°C for 15 s, 55°C (Inga03) and 59°C (Inga08, Inga33 and Pel5) for 30 s, 72 °C for 30 s, 30 cycles; 72 °C for 15 min. Completed reactions will be loaded onto an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Forest City, CA, USA) and run according to the manufacture's protocol. Allele sizes will be determined using the ROX500 internal size standard and GeneMarker® v2.4 software. Genetic variation and structure will be determined using analysis of molecular variance and Bayesian analysis of population structure, according to the objectives of the study.
The proposed extent of the thesis:	80-120 pages
Keywords:	biodiversity conservation, domestication, edible fruits, *Inga*, introgression, microsatellite markers, shade trees

Recommended information sources:

1. Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (leguminosae). Am J Botany 84:1658-1663
2. Hollingsworth PM, Dawson IK, Goodall-Copestake WP, Richardson JE, Weber JC, Sotelo Montes C, Pennington RT (2005) Do farmers reduce genetic diversity when they domesticate tropical trees? A case study from amazonia. Mol Ecol 14:497-501
3. Koptur S (1984) Outcrossing and pollinator limitation of fruit-set: Breeding systems of neotropical inga trees (fabaceae: Mimosoideae). Evolution 38:1130-1143
4. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
5. Oliveira PJC, Asner GP, Knapp DE, Almeyda A, Galván-Gildemeister R, Keene S, Raybin RF, Smith RC (2007) Land-use allocation protects the peruvian amazon. Science 317:1233-1236. doi:10.1126/science.1146324
6. Peakall R, Smouse PE (2012) Genalex 6.5: Genetic analysis in excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537-2539. doi:10.1093/bioinformatics/bts460
7. Pennington TD (1997) The genus inga: Botany. Kew: Royal Botanic Gardens, London
8. Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293:2242-2245. doi:10.1126/science.1061421
9. Rousset F (2008) Genepop'007: A complete re-implementation of the genepop software for windows and linux. Mol Eco Resour 8:103-106. doi:doi:10.1111/j.1471-8286.2007.01931.x
10. Weir BS, Cockerham CC (1984) Estimating f-statistics for the analysis of population structure. Evolution 38:1358-1370

Expected date: 2017/18 (SS) - FTA - State Doctoral Examinations

Advisor of thesis: doc. Mgr., Bohumil Mandák, Ph.D.

Electronically approved: 29. 3. 2016
doc. Ing. Bohdan Lojka, Ph.D.
Head of department

Electronically approved: 6. 2. 2017
doc. Ing. Bohdan Lojka, Ph.D.
Chairperson of Departmental Board

Electronically approved: 28. 3. 2018
doc. Ing. Jan Banout, Ph.D.
Dean

.

DECLARATION OF AUTHORSHIP

I, Alexandr Rollo, hereby declare that this thesis entitled Genetic variation and structure of agroforestry useful trees, *Inga edulis* Mart. and *I. ingoides* (Rich.) Willd., (Fabaceae) in Amazonian Peru, submitted in partial fullfilment of the requirements for the degree of Ph.D. at the Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, and the work presented in it is entirely my own work. Information derived from the published or unpublished work has been acknowledged in the text and a list of references is given.

Prague, May 2019

Ing. Alexandr Rollo

ACKNOWLEDGEMENT

First and foremost I wish to express my deep and sincere gratitude to my supervisor, doc. Ing. Bohdan Lojka, Ph.D., head of Department of Crop Sciences and Agroforestry, Faculty of Tropical AgriSciences (FTA), and my cosupervisor, prof. Mgr. Bohumil Mandák, Ph.D., Department of Ecology, Faculty of Environmental Sciences, both from Czech University of Life Sciences Prague (CULS) for their constructive comments, and for their support throughout this work. During this work I have collaborated with many colleagues for whom I have great regard, and I wish to extend my warmest thanks to all those who have helped me with this work. Adjunct Professor, Ms.C. Maria Margarida Chagas de Ataíde Ribeiro Ph.D., population geneticist and professor at the College of Agriculture of Castelo Branco, Polytecnic Institute of Castelo Branco, Portugal, Rita Maria Lorenco Da Costa, Ph.D., molecular geneticist at Laboratory of Molecular Biology, National Institute Of Biological Resources Oeiras, Portugal. My very deep gratitude belongs to Zoyla Mirella Clavo Peralta, director of Regional Herbarium of Ucayali in Pucallpa - IVITA, National University of San Marcos Lima, Peru, who provided her expertise in plant identification. My thanks are also due to doc. Dr. Ing. Pavel Vejl, head of Department of Genetics and Breeding, Faculty of Agrobiology, Food and Natural Resources (CULS), who gave me very valuable advices. Moreover, my fieldwork in Peru would not have been possible without the local inhabitants and farmers who opened their doors and let me work on their land.

I gratefully acknowledge the folowing institutions which helped me realized the field work in Peru: National University of Ucayali, Pucallpa (UNU), National University of Agriculture of La Selva, Tingo María (UNAS). My deep gratitude belongs also to The National Service of Natural Areas Protected by Steate (SERNANP), Peruvian Ministery of Environment, which allowed me to enter to the Peruvian Protected Natural Area (ANP).

I am also indebted to the donors, who made the study possible through their financial and technical support: Ministry of Foreign Affairs of the Czech Republic through the Czech Republic Development Aid Project entitled “Sustainable use of natural resources in Peruvian Amazon Project” No. 23/MZe/B/07-10; Ministry of Education, Youth and Sport through the scholarship on (UNU), Peru; The Academy of Science of The Czech Republic (AVCR) and The National Council for Science, Technology and Technological Innovation (CONCYTEC), Peru, through the binational Project entitled “Morphological and genetic diversity of indigenous tropical trees in the Amazon – model study of *Inga edulis* Mart. in Peruvian Amazon”; Foundation “Nadace Nadání Josefa, Marie a Zdeňky Hlávkových”, Czech Republic; European Union ERASMUS Training Program and to Mobility Grant of FTA (CULS).

DEDICATION

to my family and all Amazon people.

...když jsme se vraceli cestou domů, tak jsme viděli dřevorubce,
jak porážej strom a ten strom plakal…

František Sahula, 1990

LIST OF ABREVIATIONS

A	Average number of alleles per locus
A_R	Allelic richness
AMOVA	Analysis of Molecular Variance
ANOVA	Analysis of Variance
ANP	Protected Natural Area
ARc	*I. edulis* cultivated population Antonio Raimondi
ATc	*I. edulis* cultivated population Atalaya
AVCR	The Academy of Science of The Czech Republic
bp	base pair
b.p.	before present
BRc	*I. edulis* cultivated population Bretaña
c	cultivated
CICFOR	Center of Forestry Investigation and Capacitacion Center
CONCYTEC	National Council for Science, Technology and Technological Innovation
CTc	*I. edulis* cultivated population Campo Verde-Tournavista road
CULS	Czech University of Life Sciences Prague
CVc	*I. edulis* cultivated population Campo Verde
df	degrees of freedom
DHPLC	Denaturing High Performance Liquid Chromatography
DNA	Deoxyrybonucleic Acid
EDc	*I. edulis* cultivated population El Dorado
EPc	*I. edulis* cultivated population Ex Petroleros
ERASMUS	EuRopean Community Action Scheme for the Mobility of University Students
FAO	Food and Agriculture Organization
F_{IS}	Fixation index
F_{IT}	Fitness
F_{ST}	Wright´s Genetic distance
Φ_{CT}	Differentiation among geographical regions (AMOVA)
Φ_{SC}	Differentiation among populations within the same geographical region (AMOVA)
Φ_{ST}	Differentiation among populations (AMOVA)
FTA	Faculty of Tropical AgriSciences
GPS	Global Positioning System
GR	Geographic Region
G_{ST}	Nei´s Genetic distance
H_e	Expected heterozygosity
H_o	Observed heterozygosity
H_s	Average genetic diversity
HWE	Hardy-Weinberg equilibrium
ITS	Internal Transcribed Spacers
JHc	*I. edulis* cultivated population Jenaro Herrera
INc	*I. edulis* cultivated population Indiana
K	number of genetic clusters
LAc	*I. edulis* cultivated population Lagunas
LD	linkage disequilibrium
LnP(D)	Mean Log-likelihood of K
MAc	*I. edulis* cultivated population Manacamiri
MAw	*I. edulis* wild population Macuya - Experimental Forest Macuya
MCMC	Markov chain Monte Carlo
MY	Million Years
MZc	*I. edulis* cultivated population Mazán
N	Sample size
N_a	Number of alleles per locus
N_e	Effective number of alleles
N_l	Number of legumes
NAc	*I. edulis* cultivated population Nauta
NP_a	Number of private alleles

NW	North-west
NWA	North Western Amazonian geological region
P	Level of probability
P_a	Private alleles
PCoA	Principal Coordinates Analysis
PCR	Polymerase Chain Reaction
PIc	*I. edulis* cultivated population Pichanaqui
RHS	Royal Horticulture Society
Pop.	Population
R_S	Mean allelic richness
RPI	*I. ingoides* population River Pacaya - National Reserve Pacaya Samiria
RPw	*I. edulis* wild population River Pacaya - National Reserve Pacaya Samiria
RSI	*I. ingoides* population River Samiria - National Reserve Pacaya Samiria
RSw	*I. edulis* wild population River Samiria - National Reserve Pacaya Samiria
RUI	*I. ingoides* population River Utiquinia
RUw	*I. edulis* wild population River Utiquinia
S	South
SAc	*I. edulis* cultivated population Satipo
SCc	*I. edulis* cultivated population Santa Clotylda
SD	Standard Deviation
SDw	*I. edulis* wild population Sierra del Divisor - National Park Sierra del Divisor
SERNANP	National Service of Natural Areas Protected by Steate, Peru
Site	Study site
SMc	*I. edulis* cultivated population San Martín de Pangoa
SRc	*I. edulis* cultivated population San Ramón
SS	Sum of Squares
SSc	*I. edulis* cultivated population Santa Sofía
SSRs	Single Sequence Repeats
SWA	South Western Amazonian geological region
UNAS	National University of Agriculture of La Selva, Tingo María, Peru
UNU	National University of Ucayali, Pucallpa, Peru
UNESCO	United Nations Educational, Scientific and Cultural Organization
VHc	*I. edulis* cultivated population Von Humboldt
VRc	*I. edulis* cultivated population Villa Rica
W	West
w	wild
YAc	*I. edulis* cultivated population Yarinacocha

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Inga species (Fabaceae) are important components of neotropical forests. *Inga edulis* is frequently used tree species for fruits and shade tree in the Amazon region. *Inga ingoides* is phylogenetically *I. edulis* close relative, but underutilized and poorly known tree species. Little is known about *I. edulis* species' genetic structure in the wild and cultivated populations in Amazonian Peru, as well as the degree of introgression with *I. ingoides*. The genetic structure and diversity was observed on 259 *I. edulis* trees in five wild (62 trees) and 22 cultivated (197 trees) populations in three different geographical regions (Selva Central, Ucvayali and Loreto) of Amazonian Peru. Seventy seven *I. ingoides* were sampled in three wild populations and used to asessed the degree of genetic divergence and introgression with wild *I. edulis*. Microsatellite markers, analysis of molecular variance and Bayesian analysis have been used to determine the genetic diversity and population structure of both species. Characterization descriptors for *I. edulis* have been designed for further use in hybridization programs. Legume length was measured to highlight morphological difference between wild and cultivated *I. edulis* trees.

The average legume length in cultivated trees (83 cm) was significantly larger then in wild trees (39 cm). The Loreto region cultivated *I. edulis* trees showed longest legumes and lowest allelic richeness. The expected genetic diversity and the average number of alleles was higher in the wild *I. edulis* compared to the cultivated *I. edulis* populations. Overall genetic differentiation between wild *I. edulis* and *I. ingoides* was weak and the degree of genetic variation was similar. A putatively strong introgression was detected between the two species and an intense gene flow was identified among populations. The identified intense gene flow in the past could have led to a small differentiation among populations within species. A loss of genetic diversity was confirmed in the *I. edulis* cultivated populations. The species could have been simultaneously domesticated in multiple locations, usually with local origin. The original *I. edulis* Amazonian germplasm should be maintained, and cultivated population new germplasm influx from the wild populations could increase genetic diversity, provided that fruit yield will not be compromised. Selection of natural hybrids or artificial hybridization between *I. edulis* and *I. ingoides* could be applied to improve legume size and yield in the later species, while maintaining tolerance to flooding. Improved *I. ingoides* could be used in multipurpose agroforestry on open areas along the rivers, instead of using the usual slash and burn practice to create new inland open areas.

Key words: biodiversity conservation, domestication, edible fruits, *Inga*, introgression, microsatellite markers, shade trees

ABSTRAKT

Rod *Inga* (*Fabaceae*) tvoří důležitou složku amerických tropických lesů. Druh *Inga edulis* je v Amazonii hojně využíván pro produkci jedlého ovoce a pro stínění v agrolesnických systémech. *Inga ingoides* je blízce příbuzný *I. edulis*, avšak téměř neznámý a lidmi nevyužívaný druh. Málo je známo o genetické diverzitě a struktuře těchto druhů v peruánské Amazonii, podobně jako o stupni jejich vzájemné introgrese. Genetická diverzita a struktura byla hodnocena u druhu *I. edulis* na vzorku 259 stromů v pěti divokých a 22 kulturních populacích, ve třech geograficky odlišných oblastech (Selva Central, Ucvayali a Loreto). Sedmdesát sedm jedinců druhu *I. ingoides* bylo sledováno ve třech přirozených populacích. Hodnocen byl také stupeň genetické divergence a introgrese mezi divokými populacemi obou druhů. Pro měření genetické diverzity a populační struktury obou druhů byla použita metoda mikrosatelitních markerů, analýzy molekulární variance a Bayesovská analýza. Pro budoucí šlechtitelské účely byl navržen morfologický deskriptor druhu *I. edulis*. Délka lusků divokých a kultivovaných forem *I. edulis* byla sledována pro důkaz jejich morfologické odlišnosti.

Průměrná délka lusků u pěstovaných stromů *I. edulis* (83 cm) byla výrazně větší než u divokých stromů (39 cm). Nejdelší lusky *I. edulis* a nejnižší alelická bohatost byli pozorovány v oblasti Loreto. Očekávaná genetická diverzita a průměrný počet alel byl u druhu *I. edulis* vyšší u planých populací ve srovnání s pěstovanými. Celková genetická diferenciace mezi divokými *I. edulis* a *I. ingoides* byla slabá a stupeň genetické variability podobný. Mezi oběma druhy byla pozorována silná introgrese a intenzivní genový tok, který mohl být již dříve příčinou malé diferenciace mezi populacemi. Ztráta genetické diverzity byla pozorována v kultivovaných populacích *I. edulis*. Výsledky studie naznačují, že druh *I. edulis* mohl být domestikován na více místech současně, obvykle z materiálu místního původu. Původní přirozeně se vyskytující genetický materiál druhu *I. edulis* by měl být předmětem ochrany a konzervace, navíc přikřížením divokého materiálu by za předpokladu, že nebude ohrožen výnos jedlého ovoce, mohla být zvýšena genetická diverzita u populace pěstovaných jedinců tohoto druhu. Výběr přírodních mezidruhových hybridů nebo umělá hybridizace mezi *I. edulis* a *I. ingoides* by mohli být použity pro zlepšení velikosti a výnosu lusků u druhu *I. ingoides*, při zachování tolerance vůči záplavám. Vylepšený druh *I. ingoides* by následně mohl být použit při zavádění agrolesnických systémů v příbřežních zátopových oblastech, což by mohlo zmírnit tlak na zisk nové zemědělské půdy obvyklou metodou žďáření lesa.

Klíčová slova: domestikace, *Inga*, introgrese, jedlé ovoce, mikrosatelitní markery, stinné stromy, zachování biodiverzity

TABLE OF CONTENTS

.

1 INTRODUCTION

The Amazon drainage basin containing mainly lowland rainforest habitats is a major component of the Neotropical region, with more than 8 million km^2 and about 25 million people (Junk and Piedade 2011). Increasing population density and human activity are destroying the forest landscape and inflicting a loss of biological diversity (Oliveira et al. 2007). Amazonian inhabitants have used natural resources through millennia and modified the natural environment, but how human management practices resulted in Amazonian forests domestication is not known, in particular the germplasm source (Levis et al. 2018). Moreover, the species' gene pool could have been narrowed due to farmers' selection, thus strategies for genetic resource conservation and management are needed (Dawson et al. 2009). Due to its large, relatively contiguous area the peruvian primary rainforest, has major conservation value and is considered a priority in nearly all global biodiversity inventories (Brooks et al. 2006; Oliveira et al. 2007). Despite major conservation value recognized internationally due to their uniqueness and importance, the impacts of human activities throughout the region remain poorly understood (Oliveira et al. 2007). Today, due to the continuing massive pressure exerted by farmers, cattle ranchers, and logging companies on the forests, new management concepts are urgently required to avoid the destruction of this unique forest type (Junk and Piedade 2011). The Peruvian Amazon tropical area (ca. 661,000 km2) suffered disturbance and deforestation at the average rate of 647 km^2 per year from 1999 to 2005: 75 % within legally sanctioned areas, 64 % concentrated around the Ucayali logging centre Pucallpa, and 1–2 % occurred within natural protected areas (Oliveira et al. 2007).

The genus *Inga* Mill. (Fabaceae) comprises ca. 300 species of trees restricted to tropical America. Each region has preferred edible *Inga* species sold in large quantities in markets during the fruiting season. *Inga edulis* Mart. (Fabaceae) is a lowland rain forest light-demanding species, distributed in Colombia and tropical South America east of the Andes, extending from south to north-western Argentina. The species natural altitudinal range is mostly below 750 m, though it has been occasionally recorded at 1 200 m in Roraima, Brazil. Usually occurs naturally on non-flooded or only temporarily flooded sites (Pennington 1997). It is one of the most widely distributed and economically useful species in the whole Amazon region (León 1998; Pennington and Fernades 1998). This fast growing, symbiotic nitrogen fixing tree with umbrella-like canopy, is commonly used for fruit consumption as a shade tree for cocoa, coffee, coca and tea plantations, in agroforestry systems and in "home grown" multi-purpose uses cultivations (Pennington 1997). Historical records show that this species has been cultivated in Peru for its

edible fruit since pre-Colombian time and has become a commonly used tree species in the Amazon region (Nichols and Carpenter 2006). The origin of the cultivated populations of *I. edulis* is uncertain (Pennington 1997), however León (1987) and Clement (1999a) claimed West Amazonia as a probable origin. The species genetic structure was not studied in detail, yet a reduction of allelic richness in cultivated relative to natural populations was found in *I. edulis* from Peruvian Amazon (Dawson et al. 2008; Hollingsworth et al. 2005). *Inga ingoides* (Rich.) Willd., a close relative of *I. edulis*, is used frequently in gardens and pastures for its edible fruit, and has ecological adaptability with potential use in a wide range of locations with limited conditions due to flood or poor soil drainage (Pennington 1997). This species could be considered as a multipurpose fruit tree species in agroforestry and other crop systems practiced in areas affected by periodical flooding. Production of fruit and timber from this species near rivers would be less costly, more sustainable and more forestfriendly due to: (1) easy accessibility for humans, (2) economy of transport, (3) nutrient input provided by periodical flooding, and (4) cultivation in forest buffer zones avoiding new forest sites colonization. Such use could be achieved by genetic improvement through selection of natural hybrids or artificial hybridization with *I. edulis* and backcrossing, selecting for tolerance to flooding, legume size and yield, similar to the type of breeding achieved in the genus Eucalyptus (Potts and Dungey 2004). Interspecific hybrids of Eucalyptus have been used in forestry for decades, particularly in tropical and sub-tropical forestry, with plantations initially based on outstanding spontaneous hybrids. Selection was based on phenotype, followed afterwards by breeding programs based on manipulated hybrids (Potts and Dungey 2004). A similar approach, initiated with the selection of performing hybrids, could be applied to the *Inga* species under study.

Population genetic studies of tropical trees have shown, that most of the species investigated are outcrossed and exhibit high levels of genetic diversity and gene flow, carrying much of the variation within, rather than among populations (Finkeldey and Hattemer 2007). Also, the specific evolutionary history of each species has played an important role in determining the level and distribution of genetic diversity (Hamrick et al. 1992). In tropical forests, the levels of genetic diversity within populations vary considerably among species (Finkeldey and Hattemer 2007). Genetic differentiation among populations is slightly higher for tropical forest tree species than for temperate forests tree species, probably due to higher fragmentation levels in tropical trees. Moreover, tropical tree species with abiotic seed dispersal show, on average, much higher differentiation among populations than biotic-seed dispersed species. Seed dispersal by animals (zoochory) is usually very efficient and results in low genetic differentiation among populations (Loveless 1992).

Studies in *I. edulis* and *I. vera*, using microsatellite markers, compared natural vs. planted populations to understand habitat fragmentation and to clarify the impact of species domestication and possible diversity loss (Cruz-Neto et al. 2014; Hollingsworth et al. 2005; Dawson et al. 2008). The authors of the latter study found, that diversity was lower in planted compared to natural populations, but the values were still relatively high and the genetic diversity in planted stands can, to some extent, be restored by receiving pollen from natural populations. More recently, Cruz-Neto et. al. (2014) using microsatellite markers observed high levels of genetic diversity within *I. vera* populations from the Atlantic forest of north-eastern Brazil. They concluded that cultivated populations compared to natural populations displayed reduced genetic diversity. No studies about the genetic diversity in *I. ingoides* have been published yet. Nevertheless, maintaining high levels of genetic variation within agroforestry trees are important for two main reasons: genetic variation in agricultural landscapes helps farmers to manage their inputs in more efficient ways and because they provide the ability for tree species to adjust to new environments, such as the shifting climate and weather conditions, allowing local adaptation and the migration of better-suited provenances along ecological gradients (Dawson et al. 2009). In addition, a stronger emphasis on the genetic quality of the trees planted by smallholders is needed, which means paying attention both to domestication and to the systems by which improved germplasm is delivered to farmers for the management of tree genetic resources and the livelihoods of rural communities in the tropics (Dawson et al. 2014).

In the present study the objectives were to (i) explore differences in legume length between wild and cultivated *I. edulis* trees from different geographical regions in the Peruvian Amazon; (ii) compare the cultivated and wild *I. edulis* and *I. ingoides* populations' genetic structure using microsatellite markers; (iii) observe if the cultivated populations' genetic structure reflects the different uses and cultivation practices throughout the species use history, to help designing practical measures to preserve *I. edulis* genetic resources; (iv) test if populations from three Peruvian Amazon tributary rivers, geographically separated, had diverged and accumulated substantial differentiation among populations within the *I. edulis* and *I. ingoides* species; (v) to check for putative introgression between both species; and (vi) discuss the possibility of the targeted hybridization between the two studied species, the transfer of the tolerance to flooding from *I. ingoides* to *I. edulis*, and the transfer of legume size and yield potential from the latter to *I. ingoides*.

2 LITERATURE REVIEW

2.1 PERUVIAN AMAZON

2.1.1 Ecological conditions

Amazonia lies in tropical region with Tropical rain forest climate (Köppen 1936). The region is characterised by a hot and humid climate with only slight variation throughout the year. The rainfall ranges from 1.500 to 2.100 mm (a mean of 1,546 mm in Pucallpa, with rainfall increasing to the west, eg. 3.000 mm in Tingo Maria). Wet season period is between February – May and September – November; dry season is from June, August - December – January (MINAG 2002). Plant phenology is influenced by rainy and dry seasons (Köppen 1936). The mean annual temperature is 25.7 °C, with a maximum of 31 ºC and a minimum of 19.5 ºC, with the mean annual relative humidity reaching 80 % (MINAG 2002). Monthly average temperatures are between 24 and 26 °C, with minimum between 18 and 20 °C and maximums between 33 and 36 °C. The variation of temperature oscillates through the day between 5 and 8 °C, which is more then the anual 1 or 2 °C temperature variation (Egg and Vargas 2004). However, in the last few years, probably as a result of high deforestation, the climate has changed slightly and the difference between dry and wet periods is not so sharp (Odar and Rodrígues 2004).

Peruvian Amazon is divided into two subregions according to its climatic conditions, topography and altitude. Peruvian lowland jungle (Selva baja): up to 500 m altitude. It has hot and humid climate, with heavy annual rainfall not exceeding 3,000 mm per year. Relief is almost flat with some elevation. Depending on geographic location, it can be distinguished a tropical lowland in northern regions (Loreto, San Martin) and central region (Ucayali), and a subtropical lowland in the south (Madre de Dios, Cusco and Puno). Peruvian highland jungle (Selva alta) between 500 and 1,900 m a.s.l., with hot and humid weather, heavy rainfall in the rainy season from November to April and dry season from May to October. The relief is hilly, tropical highland in north and centre of the country (regions Loreto, San Martin, Ucayali, Amazonas and Cajamarca) and a subtropical highland in the south (Madre de Dios, Ayacucho, Apurimac, Cusco and Puno) (Egg and Vargas 2004).

The soils are very heterogeneous, but all of them have their fluvial origin. Higher located terrain is characterized by well-drainned forest areas of acidic (pH 4.4), low phosphorus (2 ppm) ultisols (acrisols according to the FAO/UNESCO classification system) (Fujisaka et al. 2000). Ultisols are yellow red, acid with low natural fertility, deep, well drainaged and with content of clay (Egg and Vargas 2004). These upland soils lack sufficient essential nutrients for sustainable, repeated harvests of trees and annual crops (Weber 2001). The upland terrain is usually flat or

undulating (de Jong 1995). The drainage of the upland soils is good to moderate, with a low content of organic matter and medium to high texture. The base saturation varies from 35-40 %, while aluminium saturation is 30 % to 70 % (de Jong 1995). Non-innundating forests or forests from higher Amazonian areas are forest with very high variation in vegetation, which depend especialy on the soil types (Egg and Vargas 2004). The other type of soils including alluvial, seasonally flooded, riverine systems are entisols (fluvisols according to the FAO/UNESCO classification system), with pH about 7 and 15 ppm available phosphorus (Fujisaka et al. 2000). Entisols occure in innundating forests along Ucayali, Mrañon, Amazon, Pastaza, Tigre and Napo rivers and affluent altitudinally lower then 200 m. In the period when rivers are rising, in case of river Amazon it is 13 m, we can see forests innundated for a few mounths. This forest has special adaptations on these types of innundation and great variation in distribution of nutrients, which are brought into the forest by floods (Egg and Vargas, 2004).

2.1.2 Prehistoric agriculture

In Americas, the Inca and pre-Inca civilizations are associated with the Peru/Bolivia center, and Maya and Aztec civilizations with the MesoAmerican center of Vavilov´s (1926) crop genetic diversity. In Amazonia, areas with higher population density in the pre-Colombian period should also exhibit a rich crop genetic heritage but the poor environment for archeological preservation and lack of research effort have not yilded much information to date (Clement 1999a).

The presence of large archeological sites in locations such as Marajo´ Island, in the mouth of the Amazon River and other locations along the main river course, were thought to be an anomaly, the result of populations that migrated down river from an Andean origin and occupied more fertile floodplain soils (Meggers 1971). Recent revisions of Amazonian prehistory, however, have discarded this conceptual model, which is reflected in the works of the French anthropologist Claude Levi-Strauss (Wiseman and Groves 2000). New studies are indicating that Amazonia supported much larger populations than previously imagined, and that cultural development was autocthonous, perhaps beginning with peoples who colonized the floodplain following migrations from the coastal regions north of Amazonia (Roosevelt 1994).

The reexamination of chronicles written by the first European explorers, as well as current archeological research, show that Amazonia was the location for significant cultural development, with large population complexes occupying the margins of the main rivers and developing an elaborate material culture and extensive trade networks. Archeological evidence indicates that pre-ceramic foraging populations were living at various sites in Amazonia between 11,000 and 10,000 years before present (b.p.). Initial occupation of the Pedra Pintada Cave near

Monte Alegre, Para´, Brazil, is estimated to be from 11,200 to 10,500 b.p., and excavations there have uncovered carbonized tree fruits, wood, and faunal remains, revealing a broad-spectrum economy of humid tropical forest and riverine foraging (Roosevelt et al. 1996).

In other parts of the lowland neotropics (Colombia, Ecuador, Panama and Peru), archeological and paleobotanical research indicates, that there was an intensification of practices surrounding plant exploitation and human interference with the environment between 10,000 and 8,600 b.p. These practices resulted in forms of horticulture emphasizing both native tubers and seed plants, and probably also involved the deliberate planting or management of various tree species. By 7,000 b.p. larger scale food production had emerged in Central America, with the cultivation of substantial areas or fields, away from houses (Piperno and Pearsall 1998).

In a site in Rondônia, Brazil, where human occupation by hunter-gatherers dates to 9,000 b.p., vestiges of agricultural activity, in the form of processing utensils, begin to appear around 4,500 b.p. (Miller 1992). At some time in the past, a number of native fruit trees were domesticated and incorporated into prehistoric agricultural systems. It is possible that this occurred initially through the 'dump heap'or incidental route to domestication, in which seeds of edible fruits collected in the forest were discarded near dwellings, as was observed at Pedra Pintada Cave. Notwithstanding, at some point keen observation and experimentation likely took the fore in tree domestication. While Lathrap (1977) believed the house garden of fruit trees and other useful plants to be the locus of agricultural experimentation, with root and tuber crops initially introduced as minor additions to the food supply.

Piperno and Pearsall (1998) suggested that the primary focus of early agricultural systems in the neotropics were on carbohydrate-rich root or tuber crops, with trees as secondary components. For horticultural crops, advances in archeological methods, such as the botanical identification of the sources of starch grains on stone artifacts (Piperno et al. 2000), and the combined use of genetics and archeology have allowed a relative dating of domestication, as in the case of the sequence of domestication of the corn (*Zea mays*) – beans (*Phaseolus spp.)* – squash (*Cucurbita spp*.) trinity reported by Smith (2001) for Mexico and North America. For neotropical tree crops, however, less information is available.

It is likely, though, that the domestication of trees went hand-in-hand with root-crop domestication, as the maintenance of gardens nearby dwellings would have provided an ideal location for the establishment of useful tree species from discarded seeds. By 3,000 – 2,000 b.p., large villages of many hectares existed on the middle and lower Orinoco River in Venezuela. By 2,000 years ago, large, socially stratified chiefdoms exhibiting elaborate ceremonial art and well-

established trading networks were thriving along the principal rivers of Amazonia (Roosevelt 1994).

2.1.3 Agriculture after European conquest

The pre-Colombian indigenous population possessed agricultural systems based on a great variety of cultivated plants, including fruit trees, and various food storage technologies. Although these native Amazonian populations were decimated by the combination of introduced diseases, missionization, warfare and slavery that accompanied European conquest, many elements of their agricultural and agroforestry systems have persisted, and continue to be a part of the agricultural practices of indigenous peoples until today (Hemming 1978; Smith et al. 1995; Schroth et al. 2004; Miller and Nair 2006). The lack of clear patterns suggests that the loss of the Amazonian indigenous population affected the crop genetic heritage severely. Dobyns (1966) estimated that 90-95 % of the Neotropical population was lost within 100-200 years after contact (Hemming 1978).

In Amazonia this loss meant a collapse from 3-5 millions to a low of about 200.000 people, often organized in small bands and restricted to „terra firme“, with relatively simple agricultural and subsistence technologies (Denevan 1992, Clement 1999a). Although individual farmers are responsible for selecting and propagating crops, the village is the unit of interest because identifies a domesticated plant population. Farmers within the village exchange germplasm and influence each others´ preferences and planting strategies. There is probably less exchange between villages than within, and less still between villages of different language groups, because there is simply less contact in general. Consequently, the fate of the village determined the fate of its crop genetic resources during the post-contact population collapse. It is possible that the 90-95 % population decline resulted in en equal loss of village units (Chernela 1986).

Loss in human numbers was quickly reflected in loss of crop diversity in the village site as the forest reclaimed the landscape. The major várzea societies, such as the Omagua on the Solimões River, disapeared almost completely (Roosevelt 1993). It is this low level of human survival in such important areas as the Omagua that is responsible for the tantalizing hints of a richer crop genetic past. Genetic erosion after contact depended not only upon population decline but upon the degree of domestication of each crop, its life history, the agroecosystem in which it was cultivated or managed, and the number of crops maintained by each human society.The Amazonian crop genetic heritage is poorly known for most crops, the partial exceptions being *Bactris gasipaes*, *Elaeis oleifera*, *Hevea brasiliensis* and *Theobroma cacao*, because they were

extensively prospected during the early 1980s by brazilian institutions. The historical accounts that record the impressions of the first Europeans to travel the Amazon provide only scant information on the nature of the indigenous agroforestry systems existing at that time (Clement 1999a).

Carvajal accompanied Francisco Orellana's expedition, which set out from the Napo River in Peru and traveled down the Amazon to its mouth in 1541 – 1542, encountering villages of very large size, with substantial stores of food, namely manioc (*Manihot esculenta*) bread, maize (*Zea mays*), dried fish and penned turtles (Carvajal 1970). Carvajal, for example, mentioned only that 'much fruits of all kinds' were found in one village, and that the road leading to another village was planted to fruit trees on one side and the other.

Acuña records that along with the staple crops of manioc, sweet potatoes (*Ipomea batata*), yams (*Dioscorea* sp.) and others, the Indians cultivated pineapples (*Ananas comosus*), guavas (*Psidium guajava*), abius (*Pouteria caimito*), Brazil nuts (*Bertholletia excelsa*), bananas (*Musa* sp., cotton (*Gossypium* sp.) and tobacco (*Nicotiana tabacum*), as well as numerous medicinal plants (Acuña 1994).

Father Joaõ Daniel's, between 1757 and 1776, described in detail indigenous methods of slash-and-burn agriculture with stone axes, and mentioned some of the fruit trees cultivated e.g. the abiu (*Pouteria caimito*), the custard apple (*Annona squamosa*) and the biriba (*Rollinia mucosa*) are listed as dooryard, and the cashew (*Anacardium occidentale*) as the fruit most cultivated by both 'wild' and 'tame' Indians, with several varieties, distinguished by color, size and acidity. Two types of genipap (*Genipa americana*) were distinguished, a large-fruited cultivated variety and a smaller wild one. The guava (*Psidium guajava*) was described as cultivated, but also growing spontaneously in savannas and open areas. Papaya (*Carica papaya*) was described as growing anywhere, without any special care. Indians grew passion fruit (*Passiflora edulis*) on trellises in their fields.

Some other species, such as the various species in the genus *Inga*, the pitomba (*Talisia esculenta*) and the ginja (identity unknown, perhaps *Eugenia uniflora*), are also mentioned, although no details are given about their cultivation. Nevertheless, it is probable that these were also cultivated, as other fruits such as the mangaba (*Parahancornia mangaba*), the cupui (*Theobroma subincanum*) and sorva (*Couma utilis*) are specifically described as wild fruits (Daniel 1976). The Brazilian naturalist Alexandre Rodrigues Ferreira, employed by the Portuguese crown to scientifically explore the Amazon from 1783 to1792, found that despite the colonial decimation of native peoples, a richness of indigenous material culture still existed, e.g. as shown in his portraits of members of the Yurupixuna and Mura tribes (Ferreira 1972).

2.1.4 Tree domestication

Poor historical records does not permit us to know the full extent to which trees were cultivated by pre-Columbian societies in Amazonia at the time of European conquest, the few documents that do exist, indicate that many forest fruits had been domesticated. It is also possible that many more species were in a state of incipient domestication (Clement 1999a). It means species that have previously been considered wild or non-domesticated begins to show signs of domestication.

A number of commonly cultivated Amazonian fruit trees have the characteristics of long periods of selection and genetic improvement. Clement (1989, 1999b) suggested a center of crop diversity in Western Amazonia, based on the genetic diversity of fruit tree domesticates, such as abiu, South American sapote (*Quararibea cordata*), peach palm (*Bactris gasipaes*), biribá (*Rolinia mucosa*), mapatí (*Pourouma cecropiifolia*) and araça-boi (*Eugenia stipitata*).

The domestication and genetic improvement of native fruit trees also occurred in other locations around Amazonia. Clement (1989) cited guaraná (*Paullinia cupana* var. *sorbilis*) from the Maue´s region in Amazonas, and murici (*Byrsonima crassifolia*), from the region of Santarém, Pará. In the Parakanã Indian Reserve (Pará), murici trees planted from seeds obtained from a regional market bear fruit in a variety of shapes, sizes and flavors, a possible indicators of a past selection process (Miller and Nair, 2006). Large pajurá fruits (*Couepia bracteosa*), sold in the Manaus market, much larger than those produced by wild forest trees, also come from the Santarém region (Ducke 1946; Miller and Nair 2006).

Inhabitants of Alter-do-Chão, a region of savanna near Santarém, are reported to cultivate a number of cashew varieties, each with a specific use. Populations of cashew growing in the savannas of Roraima are possibly of pre-Columbian origin (Miller and Nair 2006). The breeding of a ‘precocious’ variety of cashew in the Northeast of Brazil, the principal cashew-growing region of the country, was based on genetic material from Amazonas (Barros et al. 2002). The cashew traditionally cultivated by the Waimiri Atroari tribe, whose territory straddles parts of the states of Amazonas and Roraima, generally fruits in less than one year after planting from seed.

Early bearing may be an indicator of genetic selection and domestication and is observed in some varieties of abiu. Macambo fruit and trees (*Theobroma bicolor*) in the Peruvian Amazon, popular in agroforestry systems there and in urban markets also exhibit distinct characteristics of domestication. Wild varieties are softer-husked and smaller, while domesticated trees bear very quickly and in more abundance, with harder and larger fruits. Detribalized ribereños still practice selection today with this popular fruit. Similar variation occurs with the tucumã palm (*Astrocaryum tucuma*), whose fruits exhibit differences in pulp

thickness, presence of fibers, color and taste. A five-fold difference in wholesale prices between mediocre and superior tucumã fruits has been observed in the markets in Manaus, and farmers readily identify palms that consistently bear better fruit (Miller and Nair 2006).

2.1.5 Agroforestry

Given the lack of solid historical accounts about agroforestry practices in Amazonia at the time of European contact, i.e. how the trees were actually cultivated and managed, perhaps the next best possible source of information is the agricultural practices of modern day indigenous peoples. In light of historical and archeological evidence for the existence of complex and stratified societies in Amazonia, the use of the lifestyles of modern day indigenous groups as a model for either Paleolithic adaptations or for a pre-Columbian scenario must be used with caution. A more likely model is that modern tribal groups represent the fragments of populations and cultures that survived and regrouped following the colonial decimation (Roosevelt 1989; Roosevelt et al. 1996).

Agricultural systems relying on a number of domesticated plants also appear to have survived the social upheavals caused by European conquest. Clement (1999a) identified a list of at least 138 species of plants that were under cultivation or management at the time of European arrival in Amazonia of which 68 % are trees or woody perennials. An indicator of how complex the pre-Columbian agricultural systems may have been the number of varieties of crop plants observed in modern-day indigenous fields. Chernela (1986), for example recorded 137 cultivars of manioc in two Tukano villages in the Upper Rio Negro region of Amazonas, Brazil and described the social practices associated with their distribution and maintenance.

Whether the specific cultivation methods, employed by contemporary indigenous groups are the same as those of their ancestors, is a difficult question to answer. Nevertheless, it is probable that the complex indigenous agroforestry systems described in the ethnobiological literature of the past few decades are direct descendants of the systems in existence prior to European arrival. For example the Amuesha natives in eastern Peru use the following system: Old fruit trees are left in fields cleared from fallows. Cocoa is planted under plantains, then underplanted with ice-cream bean (*Inga. edulis*) (Salick and Lundberg 1990).

The Bará natives in Colombia/Brazil: Men plant fruit trees, both in fields and close to the longhouse. Species include peach palm, mango (*Mangifera indica*), papaya, lime (*Citrus sp*.), caimito, uvilla (*Physalis peruviana*), ice-cream bean. Other cultivated species include coca, calabash tree (*Crescentia cujete*), yagé (*Banisteriopsis caapi*) and tobacco (Jackson 1983). The transformation of indigenous agroforestry systems during the colonial period with the

establishment of the Portuguese in Eastern Amazonia, in the beginning of the 17th century (Bele´m was founded in 1616), a number of exotic species were introduced and incorporated into indigenous agroforestry systems (Miller and Nair, 2006).

By the mid 19th century, exotic fruit trees were fully incorporated into homegardens along the Amazon River. Traveling on the Amazon between Óbidos and Manaus in 1849, the British naturalist Bates (1910) described homegardens with banana, papaya, mango, orange (*Citrus* sp.), lemon, guava, avocado (*Persea americana*), abiu, genipap (*Genipa americana*) and biribá (*Rollinia deliciosa*), as well as coffee shrubs growing under the shade of the fruit trees (Miller and Nair, 2006).

Although the stratified societies or chiefdoms that existed along the Amazon and on Marajó Island disappeared in the early colonial period, it is likely that some of their subsistence practices lived on and were perpetuated by the population of mixed blood that eventually replaced them (Anderson-Gerfaud 1988; Strudwick and Sobel 1988). Father Samuel Fritz, on the Marañon River in 1686, noted that Portuguese entered what is now Peru to extract cocoa with Indians (Fritz 1922). In 1739 French traveler Charles de la Condamine found cocoa growing wild along both banks of the Marañon River in Peru, nevertheless, that the Indians gave it little attention (Condamine 1944). This distribution, nonetheless, coincides with the areas occupied by the great chiefdoms encountered by the first Europeans to travel the Amazon. Although the use of the cocoa bean was first brought to the attention of Europeans during the Spanish conquest of Mesoamerica, the place of origin of cocoa is Amazonia (Dias 2001).

In the Maue´s site the mix of indigenous domesticates and introduced species raise the interesting possibility that at least in certain locations of Amazonia, where Indian villages were replaced by towns, the cultivation of fruit trees has been continuous from prehistoric times to the present day. The most traditional agroforestry systems in Amazonia are forms of cultivation of fruit trees and other useful plants, with origins dating to the dawn of agriculture in the region, several thousands of years ago. Despite the undeniable importance of introduced species, the fact that a number of native species, mostly fruit trees, were domesticated in the pre-Columbian era is a lesson concerning the potential of Amazonian flora (Miller and Nair 2006).

2.2 THE GENUS *INGA*

2.2.1 Origin

The genus *Inga* Mill. (Fabaceae) comprises ca. 300 species of trees restricted to tropical America. Each region has its preferred species of edible *Inga* sold in fruiting season in large

quantities in markets. *I. edulis* is one of the most widely distributed and economically useful in the whole Amazon region, highly valued by the local farmers as a fruit tree species (Pennington 1997). Within the area of distribution of *Inga*, there are some regional concentrations of useful species (Pennington and Fernández 1998). One is in the Amazon, where some species have been improved by human selection (Ducke 1946). Another center is Mesoamerica: central Mexico to NW Costa Rica. In the registers of plants cultivated in Mesoamerica before the arrival of the Europeans, there is only a dubios reference to *Inga* (posibly *I. jinicuil* (Cabrera 1978) from Mexico Michoacan by Dr Francisco Hernandez who lived in Mexico from 1570 to 1577 (Hernández 1959).

2.2.2 Phylogeny

Richardson et al. (2001) have carried out a molecular phylogenetic investigation of *Inga*. As *Inga* species are clearly distinct by unique combinations of continuously varying characters, 45 *Inga* species were tested in total by using the methods of internal transcribed spacers (ITS) of nuclear ribosomal DNA and the plastid locus trnL-F. The substitution rate estimates of ITS and trnL-F were used for estimation of evolution.

Earlier study provided strong evidence, that the neotropics are an active laboratory of speciation in this genus and suggest that a substantial proportion of species diversity in Amazonia may have arisen recently during the past 2–10 million years (MY). In phylogeny species are separated by relatively short branches, and many of them have very large effective population sizes (Richardson et al. 2001; Lavin 2006; ter Steege et al. 2013). Diversification may have been promoted by the later phases of Andean orogeny 5 MY, the bridging of the Panama Isthmus 3.5 MY and Quaternary climatic fluctuations (Coates and Obando 1996; Richardson et al 2001). The bridging of the Isthmus of Panama, may have been the earliest time that *Inga* species, which are dispersed by primates, migrated between South and Central America. It is possible that *Inga* species dispersed across the Panama Isthmus before it closed, but only 3 % of *Inga* species diversity was found on islands in the Caribbean and so such over-water dispersal events must therefore be rare, a fact reinforced by a maximum seed viability of only 1-2 weeks, which reduces to a few days if seeds are removed from the pod (Pennington and Fernandes 1998; Richardson´s et al. 2001).

However, more recent discoveries has led to suggestions that the present species diversity of the rain forests could be more recent, resulting from speciation through allopatric differentiation of populations in separate refugia (e.g. Haffer 1982; Prance 1982; Whitmore and Prance 1987). The recent study on *Inga* phylogeny made by Nicholls et al. (2015) showed intra-

specific divergence between *Inga umbellifera* populations from Panama, French Guiana, Peru and Ecuador, when the maximal divergence set was Panama, French Guiana and Peru and a minimal divergence alternative showed the two western populations Ecuador and Peru, which are geographicaly closer to each other than the two eastern populations in French Guiana, and also that the two French Guianan *I. umbellifera* populations with distinct leaf chemistries, one with high levels of tyrosine, one without any tyrosine, form separate, robustly supported clades which are not sister to each other.

Further support for such recent speciation comes from the two accessions of the widespread *I. edulis* that were collected from different parts of its range. Its polymorphism in trnL-F and ITS sequences within this species is of a similar magnitude to that detected between species. In contrast, in other species where multiple accessions (e.g. *I. chocoensis*, *I. oerstediana* and *I. laurina*) were collected from closely spaced localities have identical sequences. The variability within *I. edulis* mimics patterns of variation between species is consistent with a lack of time since speciation for the accumulation of interspecific differences that are greater than intraspecific polymorphism (Nicholls et al. 2015).

2.2.3 *Inga edulis*

Scientific name: *Inga edulis* Mart.

Local names used in Amazonian Peru (Pers. observ.; Reynel and Pennington 1997):

(wild form):, guabilla del monte - Loreto region; guabilla – Huánuco, Loreto and Ucayali region.

(cultivated form): guaba – Loreto and Ucayali regions; pacae soga - Junín region.

Distribution and ecology

Colombia and tropical South America east of the Andes, extending south to northwestern Argentina is the range of *I. edulis* natural distribution. It is also present in Atlantic coastal Brazil. *I. edulis* is doubtfully native in Panama (Pinnington 1997). It grows in hot, humid climates between 26°S and 10°N. The species is widely cultivated throughout its range in South America and has been introduce throughout Central America (Figure 1). It is light-demanding gap species of lowland rain forest, also called pioneer species. Even though seedlings often establish themselves in the shade, it requires light for growing and flowering. In the forests it becomes a canopy tree, but it is also common in secondary forest. *I. edulis* tolerates various types of soil - from acidic (pH 4.0) to alkaline soils even with high saturation of aluminium, although prefers sandy soils along watersides. *I. edulis* can also withstand temporal floods and high rate of soil skeleton (Reynel et al. 2003). Its natural altitudinal range is mostly below 750 m, but it has been

occasionally recorded as high as 1,200 m in Roraima, Brazil. It is cultivated up to 1,600 m. In Peru up to 700 m, temeperature range between 17 – 26.5 °C and anual precipitation 1,370 – 3,000 mm (Pennington 1997; Reynel et al. 2003).

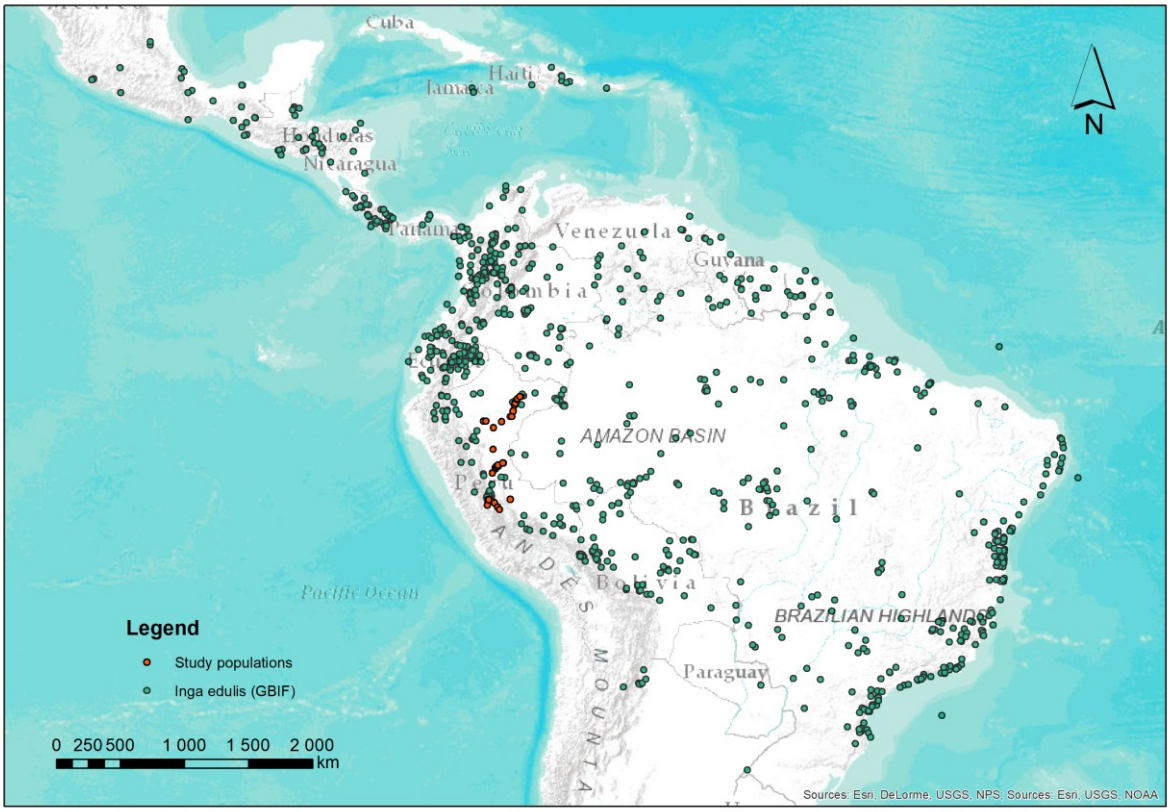

Figure 1. Distribution map of *Inga edulis* (1686 occurrences), including the sampled trees from the current study (259 occurrences). GBIF.org (10th October 2018) GIBF Occurrence Download https:doi.org/10.15468/dl.ik3uki.

Uses

I. edulis is one of the most widely distributed and economically useful species in the whole Amazon region, but the origin of the cultivated forms is uncertain, though probably Amazonian. More specifically other authors (León 1987; Clement 1999a) stated the probable origin of semi-domesticated forms of *I edulis* in West Amazonia. (León 1998; Pennington and Fernandes 1998). Historical records showed this species was cultivated in Peru for its edible fruit, by the pre-Colombian inhabitants and become one of the most used tree species in the Amazon region (Lawrence 1995; Pennington 1997; Kanmegne et al. 2000; Reynel et al. 2003). *I. edulis* has been introduced across most of tropical South and Central America (Pennington 1997). Although the history of cultivation of this species is not well-documented, morphological studies suggest that humans have semi-domesticated *I. edulis* over a considerable time period (Clement 1989, 1999a; Pennington 1997).

The useful part is the edible pulp (sarcotesta) that surrounds the seeds in the long pod. It is watery, soft, slightly sweet, generally white tissue, which is used for household consumption or sold on the local market. In Latin America this fast growing, acid-soil tolerant tree, which improves soil fertility through symbiotic nitrogen fixation, is traditionally used to shade perennial crops such as coffee and cocoa. *I. edulis* was found to have greater than 90% survival, relatively high growth rates and high nitrogenase activity (Tilki and Fisher 1998), indicating that it can actively fix nitrogen on soils as well with high acidity and aluminium toxicity (Nichols and Carpenter 2006). *I. edulis* also provides firewood and charcoal and produces sweet pulp suitable for human consumption (Pennington and Fernandes 1998). Its leaves are used in the folk medicine as anti-inflammatory and anti-diarrheic drugs. It has been also prove that *I. edulis* is useful as green manure and helps control weeds and erosion on degraded soils (Lawrence 1995). *I. edulis* has been used in these manners for hundreds years (Dawson et al. 2008).

Its high rate of growth, ability to improve degraded soil, the provision of economically useful products and usable biomass and its presence throughout the Amazon basin sparked interest of scientists in further exploring the potential of this species for agroforestry and other environmental improving systems (Weber 2001).

Field characters and botanical description

Tree to 40 m high and 65 cm in diameter, buttressed to 1 m high, and bole fluted to 2-3 m, cylindrical above. Bark smooth is pale greyish, sometimes lenticellate and with hoop marks. Slash 3-5 mm thick, pink or pale brown, with clear exudate. Flowers sweetly scented with greenish -yellow perianth, with filaments and pale yellow anthers. Mature fruit greyish-green. Young shoots angular, pale lenticellate, puberulous. Stipules are 2-6 mm long, oblong to lanceolate, appresed puberulous, usually terete, rarely winged in the upper half, puberulous; appendix absent. Foliar nectaries sessile, 2-3 mm diameter, aperture transversaly compressed to reniform. Petiolule is 1.5-3 mm long. Leaflets 4-6 pairs; terminal pair 10-19 x 3.8-8.9 cm, elliptic to obovate, apex acute, obtusely cuspidate to narrowly attenuate; base rounded, slightly asymmetrical; basal pair 3.8-7.5 x 1.9-4.3 cm, elliptic or ovate, apex obtusely cuspidate to narrowly attenuate, base rounded to truncate; upper midrib puberulous, lamina minutely scabrid, lower lamina scabrid to crisped puberulous; venation eucamptodromous to brochidodromous; secondary veins (12-) 15-20 pairs, parallel to slightly convergent, slightly arcuate; intersecondaries short to moderate; tertiaries oblique. Inflorescence (size 2-10 cm), axillary, sometimes clustered at the shoot apex in the axils of undeveloped leaves, up to 6 in each axil, a congested or less frequently lax spike; peduncle 1-5 cm long, puberulous; floral rhachis 1-4.5 cm

long; bracts (3-) 4-8 (-10) mm long, caducous; flowers sessile (size 2.5-5cm). Calyx open in bud; tube 4-9 mm long, tubular, lobes 1-2 mm long; puberulous. Corolla tube 0.9-1.9 cm long, lobes 2-4 mm long; sericeo-villose. Stamens 55-100, staminal tube 1-2 cm long, 1.5-2 mm diameter, included or slightly exserted, free filaments 1.5-3 cm long. Ovary of 1 carpel, glabrous, style slightly longer than stamens, style head cup-shaped, ovules 20-30. Legume 30-100(-200) x 2-5 cm, cylindrical, straight or spirally twisted, apex acute to rostrate, base tapered, faces completely covered by expanded margins, margins longitudinally ribbed; puberolous. The ripe legume has greyish-green colour. Seeds 2-3 x 1-1.5 cm, see botanical description in Figure 2. (Pennington 1997; Reynel et al. 2003).

The major flowering season throughout its range is from June to October, but in Amazonian Brazil and Peru there is a smaller peak in March and April. In Peru there are two flowering peaks from March to May and July to August. The fruiting season is difficult to assess as the majority of fruiting collections are immature fruit, but field observation throughout western Amazonia indicates that the major fruiting season is from October to January and in Peru between December and January. Today this species is widely cultivated for its edible fruit throughout its range in South America and Central America. Some selection for fruit size and quantity of edible flash has taken place over the years. The cultivated plants have larger flowers than wild populations and the legumes generally much longer than those on wild trees that rarely exceed 50 cm in length. Some of the best strains are found in western Amazonia (especially Amazonian Peru) where the fruit may exceed 2 m in length and 5-6 cm in diameter. The cultivated plants have larger flowers than wild populations, and the fruit of the latter rarely exceed 50 cm in length. The species is very commonly used as a shade tree in small gardens, due to its rapid growth and the broad spreading crown. It is sometimes found as a shade tree over coffee in Central America, but it is not favoured for this due to the damage caused to the coffee by children climbing for the edible fruit (Pennington 1997; Reynel et al. 2003).

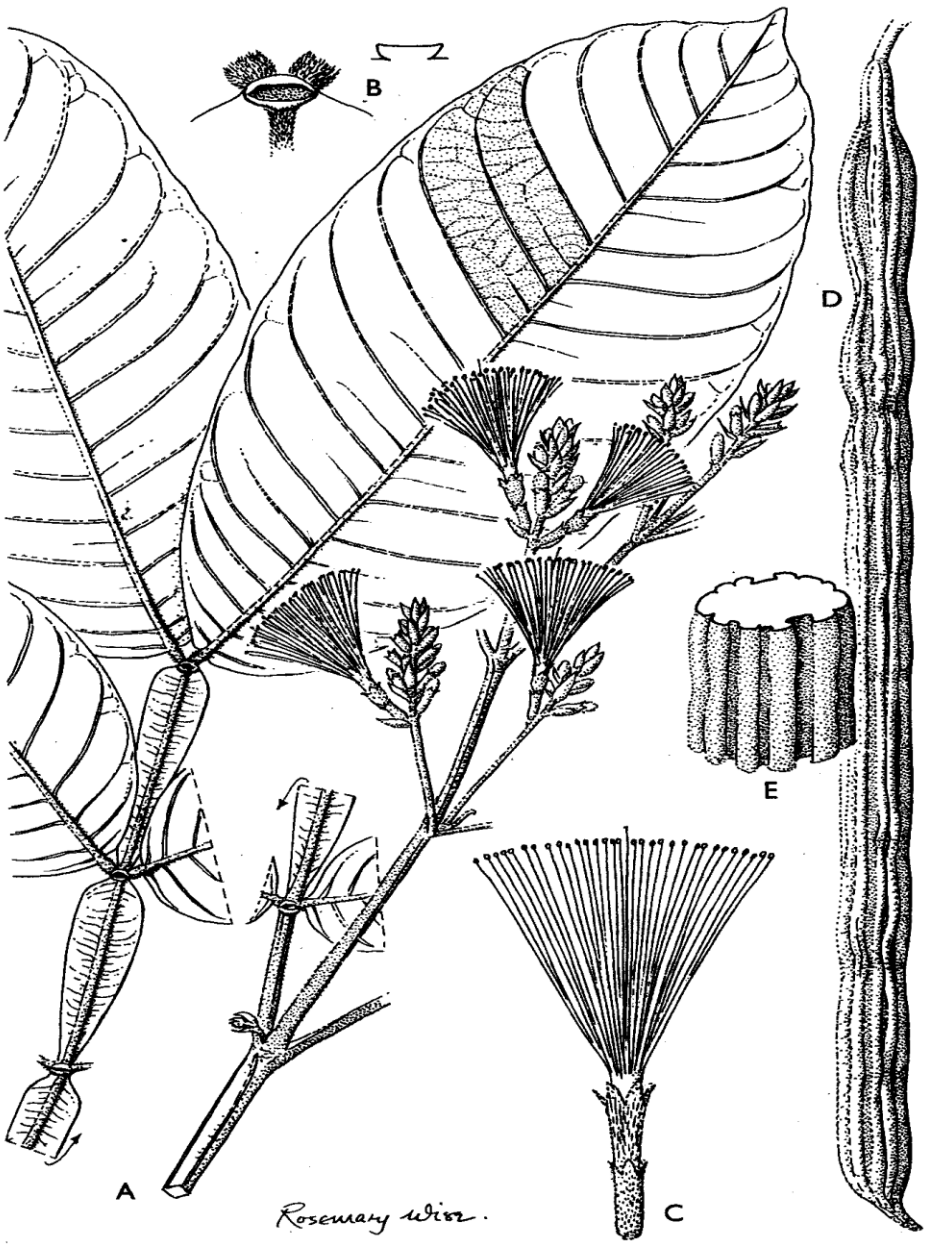

Figure 2. *Inga edulis* A habit and inflorescence x 2/3; B foliar nectary x 2.6; C flower x 1.3; D legume x 2/3; E legume section x 2/3 (Pennington 1997).

Mating system

Inga edulis is crosspolinated diploid 2n = 26 (Figueiredo et al. 2014), believed to be selfincompatible (Dawson et al. 2008). It has hermaphrodite brush type flowers with mainly nocturnal anthesis, which are pollinated by hawkmoths, bats or hummingbirds visits (Cruz-Neto et al. 2011). Fruiting occurs after three years producing a long pod containing recalcitrant seeds covered by a fleshy, slightly sweet, generally white edible sarcotesta (Pennington and Fernández 1998). The natural dispersal is provided usually by mammals and possibly birds, which consume sarcotesta and release the seed in a nearby area (Koptur 1984; Dawson et al. 2008). The

propagation method used by farmers is generally from seed, which is recalcitrant (Hollingsworth et al. 2005; Dawson et al. 2008; Jamnadass et al. 2009).

2.2.4 *Inga ingoides*

Scientific name: *Inga ingoides* (Rich.) Willd.

Local names used in Amazonian Peru (Pers. observ.; Reynel and Pennington 1997):

Local names: guabilla del monte, and shimbillo – Loreto region; guabilla – Loreto and Ucayali regions; pacay – Ucayali region.

Distribution and ecology

This species is distributed from Lesser Antilles and tropical South America to Bolivia, including coastal Brazil to southern Minas Gerais. Not yet recorded in Ecuador. A species comonly found in secondary forests and pasture, especially on periodicaly flooded and poorly drained sites and along riversides, on white sand in Amazonian Peru. Sometimes occures on non-innundating terraces. Altitudinal range in Peru is from sea level to 500 m. The temperature range for this species in Peru is between 15-26.5 °C. The average annual precipitation is between 1,600-3,000 mm (Pennington 1997; Reynel et al. 2003).

Uses

Inga ingoides is frequently used in gardens and pastures for its edible fruit and as a shade tree for cattle. Its ecological adaptability gives potential to this species to be used in a wide range of locations with limited conditions due to flood or poor soil drainage. The wood is also used for charcoal (Trinidad), as construction timber (Martinique) and for fuelwood. It is important to experience this species under cultivation, so that its growth characteristics can be compared with those of the species *I. edulis*. Its ecological adaptability suggests, that it could be used in agroforestry and in a wide range of locations with limited conditions (Pennington 1997; Reynel et al. 2003).

Field characters, botanical description and matting system

Tree to 30 m high and 45 cm diameter, larger specimens with short concave buttresses to 50 cm high. Bark smooth, greyish, lenticellate; slash up to 1 cm thick, reddish, without exudate. Leaflets 3-5 pairs, broadly eliptic, apex acute, sparsely crisped-pubescent above denser below, terminal pair 10-23 x 4.3-13.5 cm. Foliar nectaries stalked (1-2 mm long). Inflorescence congested or rarely lax raceme 3-12 cm. Flower size 3-7 cm, with greenish-yellow perianth,

white filaments and yellow anthers, with slight sweet scent. Legume cylindrical, straight, margins longitudinally ribed, size 25-30 x 1-1.8 cm, colour greenish-brown, see botanical description in Figure 3. (Pennington 1997; Reynel et al. 2003).

Inga ingoides is crosspolinated diploid 2n = 26 (Figueiredo et al. 2014). Flowering in Peru mostly August to November. Ripe fruit are rarely collected, and fruit at various stages of development are found troughout the year (Pennington 1997).

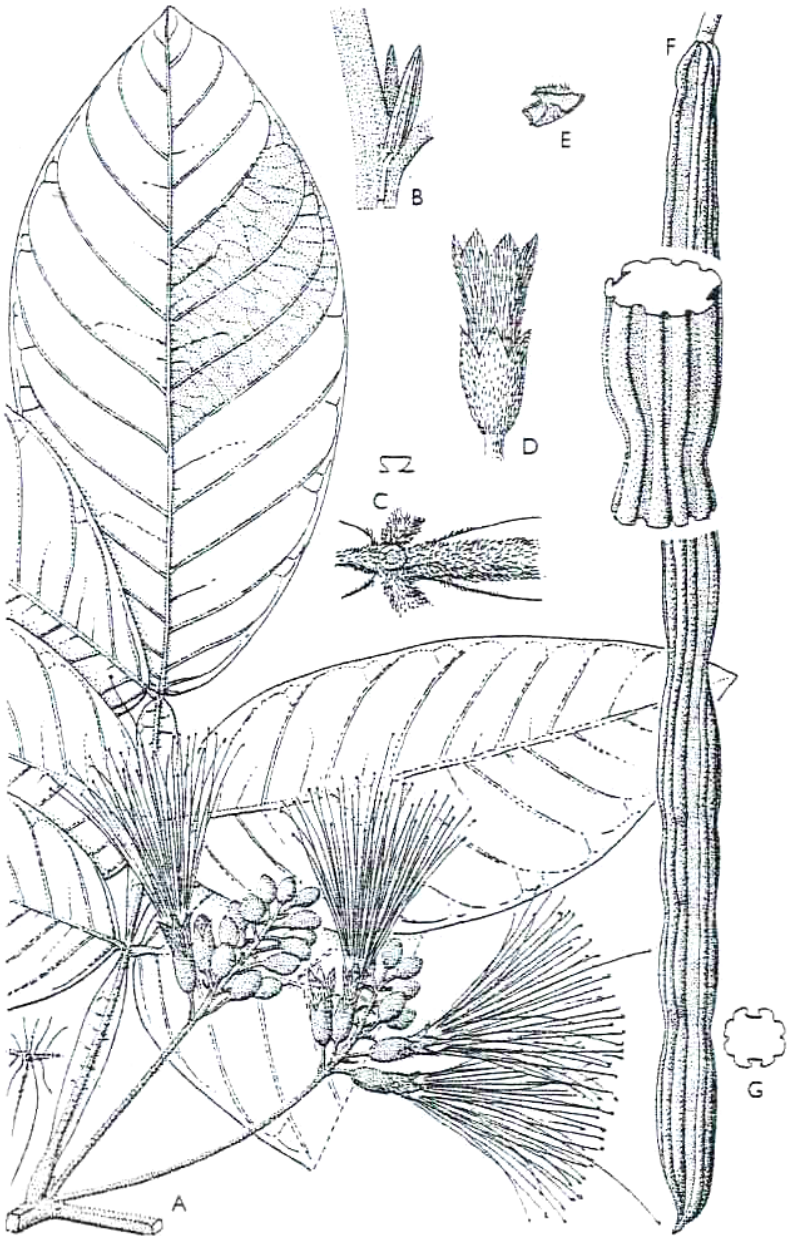

Figure 3. *Inga. ingoides* A habit x 2/3; B stipules x 1.3; C foliar nectary x 2; D flower x 2; E floral bract x 2.6; F legume section x 2/3 (Pennington 1997).

2.2.5 *I. edulis* and *I. ingoides* relationships and interspecific hybrids

Inga edulis and *I. ingoides* are sympatric species with overlapping distribution, but the former is more likely to be found in non-flooded sites since it can withstand only temporary floods. According to Pennington (1997), *I. ingoides* flowering season, from August to November, partially overlaps *I. edulis* June to October flowering season. *I. ingoides* is a close relative to *I. edulis*, sharing with it the cylindrical, ribbed legume. It differs from *I. edulis* in its flat, circular foliar nectaries (without the compressed aperture of *I. edulis*) and in the pedicellate flowers, often globose flower buds, the calyx which is closed in bud and in the generally broader calyx and much longer staminal filaments. *I. ingoides* shows a wide range of variation, as would be expected in such a widely distributed species, a few colections have a stalked foliar nectary. There is also variation in the quality of the calyx indumentum which is usually puberulous but sometimes longer and tomentose. Some specimens show some intergradation with *I. vera* especially in the shape of the calyx which is narrower than in typical *I. ingoides* (Pennington 1997; Reynel et al. 2003).

The interspecific hybrids between *I. edulis* and *I. ingoides* have not been described yet. However in another *Inga* species *I. vera* and *I. oerstediana* are some botanical specimens indistinguishable from each other and Sousa (1993) has suggested it might be a interspecific hybrids.

2.2.6 Future directions of research

Despite the importance of the issue, the history of native tree species planted on farms is largely unknown in the tropics, and there is a lack of research on geographically matched natural and planted stands to provide insights into their origins (Hollingsworth et al. 2005). Studies in *I. edulis* and *I. vera*, using microsatellite markers, compared natural vs. planted populations to understand habitat fragmentation and to clarify the impact of species domestication and possible diversity loss (Hollingsworth et al. 2005; Cruz-Neto et al. 2014). Research provided by Dawson et al. (2007) in Peruvian Amazon found that diversity was lower in planted compared to natural *I. edulis* populations but the values were still relatively high (~80 and 70 % of natural stands, respectively), indicating when farmers plant trees, good collection practice of seed from already cultivated *I. edulis* should be an effective means for ensuring long-term conservation on farms.

The genetic erosion of tree species due to anthropogenic influence has not been widely investigated in Peru; however, forest conversion to agricultural field has resulted in declining tree populations from which genetic erosion can be inferred (O´Neil et al. 2001). Because of the long history of cultivation and very extensive use in the Amazon, *I. edulis* has a great potential to

serve as a model species for further studies of other fruit tree species in this region. Morphological studies suggested that *I. edulis* has been partially domesticated in Latin America because the fruit found in farmers' fields are frequently observed to be twice as long as those on trees located in neighbouring natural forest (Pennington 1997).

In the only known genetic study made on *I. edulis*, Hollingsworth et al. (2005) employed nuclear SSRs to survey geographically proximate natural and planted stands collected from a number of locations in the Peruvian Amazon. They revealed that the high allelic richness of all planted stands was lower than that of any single natural population. The lower variation observed in planted material supported the presumption that human intervention in the Amazonian rain forest has had significant impact on the decresing level of population genetic diversity in *I. edulis*. In order to assess the process of domestication further, Dawson et al. (2008) employed maternally - inherited chloroplast polymorphisms on the same study material (conserved organellar sequences through denaturing high-performance liquid chromatography, DHPLC). Molecular markers indicated that cultivated *I. edulis* in smallholders' farms was genetically differentiated from local wild material and therefore it was unlikely to have originated from that. While unrelated planted populations had different genetic compositions from each other, suggesting multiple rather than single external sources for cultivated germplasm. Based on this data, it seems probable that *I. edulis* stands in agricultural landscapes in the region were not likely to suffer significantly from any inbreeding effects due to low genetic diversity (Jamnadass et al. 2009).

As shown previously, only few studies have been done on genetic matter of *I. edulis* and *I. ingoides.* Gaining in depth understanding of current levels of genetic variation and gene flow within tree populations in farmland is therefore essential, as it determines whether or not new infusions of diversity is sought to supplement existing variation and if additional interventions are needed to enhance the natural pollination. Through assessing genetic bottlenecks and by providing direct estimates of gene flow through paternity exclusion, molecular markers can provide insights into these issues (Jamnadass et al. 2009). To the best of our knowledge there has been no significant research using molecular markers to evaluate gene flow in *I. edulis* and *I. ingoides* species. The proved intraspecific genetic variation could not only create opportunities for selection, but also provide an adaptive buffering capacity to changing user requirements and environmental pressures (White et al. 2005).

Characterization descriptors for *I. edulis* species for evaluation of phenotypic variance are not available.

3 HYPOTHESES

Inga edulis is probably recently domesticated and one of the most used species in the Amazon region for fruit and shade tree. Chloroplast haplotype composition results in previous study displayed a completely different pattern between natural and cultivated populations of *I. edulis* in Amazonian Peru and exclude origin of domesticated forms in a local wild material. Usualy the crops in an initial process of domestication, show no clear genetic structuring between local wild and cultivated populations. Therefore we suppose wild and cultivated populations of *I. edulis* could show similar gentic structure and could have been domesticated from the local wild material. *Inga ingoides* is still underutilized and poorly known species, but unlike the previous species *I. ingoides* is more tolerant to floods. Various authors have found sympatry in severel *Inga* species. Due to *I. ingoides* is close relative to *I. edulis* and both species are found in the same phylogenetic node we hypothesized, introgression between these two species could exist.

4 OBJECTIVES

Objectives of the study are to observe genetic variation and structure of *Inga edulis* and *I. ingoides* in Amazonian Peru using microsatellite markers.

The specific aims of the study are as follow:

i) To design characterisation descriptors for *I. edulis*.

ii) To compare the wild and cultivated *I. edulis* legume lengths.

iii) To compare the wild and cultivated *I. edulis* populations from three different geographical regions.

iv) To compare *I. edulis* and *I. ingoides* wild populations genetic variation and structure and to asses the degree of their genetic divergence and introgression among populations from three geographically separated peruvian Amazon tributary rivers.

5 METHODOLOGY

5.1 PLANT MATERIAL

Inga edulis and *I. ingoides* were sampled in Amazonian Peru between 2009 and 2012. The two species were identified according to morphological aspects detailed in Table 1. The major differences between them were: presence of clear internal bark exudate (*I. edulis*) and absence of internal bark exudate (*I. ingoides*); foliar nectary aperture transversaly compressed to reniforme (*I. edulis*) and regular shaped (*I. ingoides*); inflorescence spike (*I. edulis*) and raceme (*I. ingoides*); and the habitat preference of the wild trees at non-innundating terraces (*I. edulis*) and periodicaly flooded and poorly drained sites (*I. ingoides*).

Table 1. Key to species identification according to ecological and morphological aspects (source: Pennington 1997; Reynel et al. 2003)

Identification clue	***Inga edulis***	***Inga ingoides***
Habitat	non-innundating terraces	periodicaly flooded and poorly drained sites
Soil	acid soil tolerant, poorly drained	poorly drained, white sand
External bark	pale greyish	greyish
Internal bark	pink or pale brown, with clear exudate	reddish, without exudate
Young shoots	palle lenticellate, puberulous	sparsely lenticelate, pubescent
Stipule	2-6 mm long, oblong to lanceolate, apperessed puberulous, caducous	0.6-1.5 cm long, oblong to eliptic, pubescent, caducous
Petiole	usually terete, rarely winged in the upper half, puberulous	winged or terete, pubescent
Rhachis wing	to 1.6 cm wide	to 2.7 cm wide
Foliar nectary	sessile 2-3 mm diameter, aperture transversaly compressed to reniforme	stalked 1-2 mm long;
Calyx	open in bud	closed in bud
Anthers	pale yellow	yellow
Inflorescence	sometimes clustered at the shoot apex in the axils of undeveloped leaves, up to 6 in each axil, a congested or rarely lax spike	often clustered near the apex in the axils of undeveloped leaves, up to 4 in each axil, a congested or less frequently lax raceme
Peduncle	1-5 cm long	1.5-8 cm long puberulose
Pedicel	missing	3-10 mm long
Fruit	cylindrical, straight or spirally twisted, apex acute to rostrate, faces completely covered by expanded margins	cylindrical, straight, apex tapered, faces almost comleted covered by the expanded margins
Legume colour	greyish-green	greenish-brown
Legume size	30-100 x 2-5 cm	25-30 x 1-1.7 cm

At least five sexually mature trees with fruits were randomly selected and sampled in each geographically distinct population. Each sampled tree's geographical coordinates were recorded, and 200 m was the minimum distance between any two studied trees. Voucher specimens were kept in the Regional Herbarium of Ucayali IVITA-Pucallpa, Perú, with the code AR1-384. Sampled individual's young leaves were collected and preserved in micro test-tubes with silica gel for further DNA extraction.

Throughout the development of this thesis the appropriate phenotype characterisation descriptors were created and designed (see Result section) according to several other fruit or *Fabaceae* tree descriptors (Batlle and Tous 1997; IBPGR 1980; IBPGR 1993). The scale of possible variance was designed using literature on *I. edulis* morphology (León 1966; Duke 1983; Lawrence 1993; Villachica 1996; Paytan 1997; Pennington 1997; Reynel et al. 2003; Lewis et al. 2005) and modified according to the authors own observation.

Morphological characteristics has been designed and observed to describe differences among cultivated populations from three diferent geographical regions and wild populations of *I. edulis* in Peruvian Amazon. Morphological characteristic measurements for the study purposes were focused on *I. edulis* legume length. Morphological characteristics of *I. ingoides* were not evaluated. Considering the fact that this study was not designed as a field experiment, in which the monitored individuals would grow under clearly defined and same ecological conditions, but terrain research with a considerably diverse set of trees, where not only man's activity, but also the different ecological conditions could significantly affect the growth of the studied individuals, the determined features would serve only as indicative. Emphasis was placed on the sexual maturity of the sampled trees and presence of mature fruits. One to ten mature legumes were sampled per tree, in opposite sides and different heights of the crown, according to the availability of mature fruits on the tree. The mature fruits had the following phenological characteristics: seeds from creamy white to purple black up to viviparic, sarcotesta membranous creamy to generally white flashy, watery, soft and slightly sweet (Pennington and Fernandes 1998). The morphological description was focused mainly on features whose phenotypic expression is genetically determined and the influence of man and ecological conditions is expected to be low, but still these characteristics were performed only to describe the variability, but not for classification purposes of the *I. edulis* population. Genetic variability was further described by molecular methods.

5.2 STUDY SITE AND SAMPLING

A total of 259 individual *I. edulis* trees from 27 geogpaphical populations were evaluated in Amazonian Peru. Each population was numbered from 1 to 27 and coded, e.g. 1SRc, 23RPw (hereafter, the first two capital letters of the population's name are the geographic origin's initials, e.g. San Ramón or River Pacaya, the third letter meaning either (c=cultivated - managed by human or w=wild - growing spontaneously).

Cultivated trees were sampled in 22 geographically different populations, in home gardens and other landscapes surrounding the urban areas. The study site was for the study purposes divided into three regions Selva Central, Ucayali and Loreto.

Selva Central - sampling started in central Amazonian highland region in Junin and Pasco departments, with Köppen-Geiger oceanic climate (Kottek et al. 2006), month average temperature 22°C and precipitation 2000 mm per year, tropical evergreen montane forest "selva alta" (Longman and Jeník 1987), south western Amazonian geological region (SWA) (Levis et al. 2017). Sampling proceeded on Andean hills in the Chanchamayo and Perene river watersheds. Trees were sampled in fields locally called "chacras". Usually those trees provided shade for coffee, cocoa and other crops, in home gardens or as components of other agroforestry systems for fruit consumption etc. In this region the trees were sampled in five different geographic populations: San Ramon (1SRc), Villa Rica (2VRc), Pichanaqui (3PIc), Satipo (4SAc) and San Martín de Pangoa (5SMc). Sampling proceeded in lowland Amazon regions locally called "selva baja", with characteristic tropical evergreen lowland forests vegetation and Köppen-Geiger tropical rainforest climate (Kottek et al. 2006), situated between SWA and north western Amazonian geological region (NWA) (Levis et al. 2017).

Ucayali - includes the departments of Ucayali and Huanuco, with month average temperature 24 - 26°C and precipitations between 1800 and 2000 mm per year (Egg and Vargas, 2004). The region stretches from Andean foothills to Amazon plain and belongs to the upper Ucayali river watershed. *I. edulis* trees are usually cultivated in home gardens and "chacras" for fruit consumption, fire wood etc, but also as components of agroforestry systems for their multipurpose uses, including e.g. shading cocoa and coca. Sampling proceeded in seven different geographic populations: Atalaya (6ATc), Von Humboldt (7VHc), private forest plantation "Bosques Amazónicos" on Campo Verde – Tournavista road km 12 (8CTc), Campo Verde (9CVc), Antonio Raimondi (10ARc), Yarinacocha (11YAc) and Santa Sofia (12SSc);

Loreto – in the so called Loreto department, with monthly average temperature higher than 24°C and precipitations higher than 2400 mm per year (Kalliola and Flores-Paitan 1998). This region

is sprawled on the Amazonian plain in lower Marañon and Ucayali and upper Amazon River watersheds. Sampled trees were cultivated mainly “in home” gardens and “chacras” for fruit consumption, fire wood, shade, fodder, medicinal use, etc. Trees were sampled in ten different geographic populations: Bretaña (13BRc), Jenaro Herrera (14JHc), Lagunas (15LAc), Nauta (16NAc), Ex Petroleros (17EPc), El Dorado (18EDc), Manacamiri (19MAc), Santa Clotylda (20SCc), Indiana (21INc) and Mazán (22MZc).

Special attention was paid to the wild populations of *I. edulis* were wild indeed, because Dawson et al. (2008) in similar study said that the pairs of wild *vs*. domesticated populations were distinct and in close geographical proximity, normally within few kilometres. Even they say that “Despite these precautions, the long history of the use of both *I. edulis* and slash and burn agriculture in primary forest in the region means that it is not always easy to distinguish between natural and planted categories. For example, it is conceivable that material sampled from farms may occasionally represent natural remnant or regenerant trees, and that in some cases apparent primary forest may in fact be old-growth fallow”.

A total of 62 wild *I. edulis* trees were sampled in five geographically different populations, in lowland forest, in natural sites covered by original forest vegetation or transformed into secondary forests. The first two populations, River Pacaya (23RPw) and River Samiria (24RSw), were gathered in original vegetation in the Pacaya Samiria National Reserve in the Loreto department. The River Utiquinia (25RUw) population was sampled in secondary vegetation along the Utiquinia river, from the San José village heading downstream to the Ucayali river in the Ucayali department. The Macuya (26MAw) population was sampled in the Forest Investigation and Capacitacion Center (CICFOR) – Macuya in Huanuco department, a ‘terra firme’ forest remnant managed and protected by the National University of Ucayali, surrounded by deforested logged areas close to the city of Von Humboldt. The last wild population Sierra del Divisor (27SDw) was sampled near the Contamana city in Loreto department, 16 km (west) inland in secondary vegetation, which begins in undulated terrain and continues to the original vegetation in the protected mountain range Sierra del Divisor National Park.

Latitudes and longitudes were approximated from geographical coordinates captured from all individuals in the particular population. Sampled populations information and sampling details are presented in Table 2. No *I. edulis* trees were found in the Selva Central region original vegetation.

Table 2. *Inga edulis* 27 populations (cultivated and wild) sampling region (Site), population code (Pop.), sample size (N), geographic location - GPS coordinates in WGS84 (Latitude S and Longitude W) and altitude in meters above sea level.

I. edulis	**Site**	**Pop.**	**N**	**Latitude S**	**Longitude W**	**Altitude (m)**
cultivated	San Ramon	1 SRc	10	11°08′	75°21′	828 - 1200
Selva Central	Villa Rica	2 VRc	5	10°44′	75°16′	1467 - 1494
	Pichanaqui	3 PIc	10	10°55′	74°52′	497 - 631
	Satipo	4 SAc	10	11°16′	74°38′	550 - 677
	San Martín de Pangoa	5 SMc	10	11°26′	74°30′	788 - 949
cultivated	Atalaya	6 ATc	10	10°43′	73°45′	223 - 244
Ucayali	Von Humboldt	7 VHc	8	8°51′	75°00′	210 - 243
	Campo Verde-Tournavista	8 CTc	18	8°35′	74°46′	180 - 207
	Campo Verde	9 CVc	12	8°31′	74°47′	198 - 210
	Antonio Raimondi	10 ARc	11	8°29′	74°49′	147 - 158
	Yarinacocha	11 YAc	5	8°20′	74°36′	144 - 154
	Santa Sofia	12 SSc	8	8°09′	74°15′	152 - 159
cultivated	Bretaña	13 BRc	16	5°15′	74°20′	103 - 108
Loreto	Jenaro Herrera	14 JHc	5	4°54′	73°40′	100 - 127
	Lagunas	15 LAc	10	5°14′	75°37′	108 - 135
	Nauta	16 NAc	5	4°30′	73°34′	106 - 139
	Ex Petroleros	17 EPc	5	4° 5′	73° 27′	97 - 108
	El Dorado	18 EDc	12	3° 57′	73° 25′	109 - 151
	Manacamiri	19 MAc	5	3° 43′	73° 17′	95 - 97
	Santa Clotylda	20 SCc	5	3° 40′	73° 15′	93 - 128
	Indiana	21 INc	7	3°29′	73°02′	92 - 108
	Mazán	22 MZc	10	3°30′	73°04′	93 - 122
wild	Pacaya river	23 RPw	12	5° 41′	74° 57′	110-131
	Samiria river	24 RSw	6	5° 14′	75° 28′	105-123
	Utiquinia river	25 RUw	12	8° 10′	74° 17′	150-160
	Macuya	26 MAw	27	8° 53′	75° 0′	216-233
	Sierra del Divisor	27 SDw	5	7° 13′	74° 57′	196-231

A total of 77 *I. ingoides* individual trees were collected in riparian situations along three Amazon River tributaries and in upland forests (Figure 4A and B). The RPI population (hereafter, the first two letters of the population name are the initials derived from the site name, the third letter means I=*I. ingoides*) was sampled from original vegetation along the river Pacaya. The RSI population was observed in original vegetation on the Samiria river springs. Both rivers belong to the protected area Pacaya Samiria National Reserve. The RUI population was sampled in secondary vegetation along the Utiquinia river from the San José village, situated on non-inundating terraces, to the periodically flooded and poorly drained sites heading downstream to the Ucayali river. Sampled populations are detailed in Table 3. Latitudes and longitudes were approximated from geographical coordinates captured from all individuals in the particular population.

Table 3. Geographic location, sample size and study site where the *I. ingoides* and *I. edulis* populations were sampled. Sample size (N).

Species	Site	Pop.	N	Latitude S	Longitude W	Altitude (m)
I. ingoides	Pacaya River	RPI	47	5° 25'	74° 34'	105 - 127
	Samiria River	RSI	16	5° 15'	75° 22'	91 - 131
	Utiquinia River	RUI	14	8° 12'	74° 19'	148 - 168

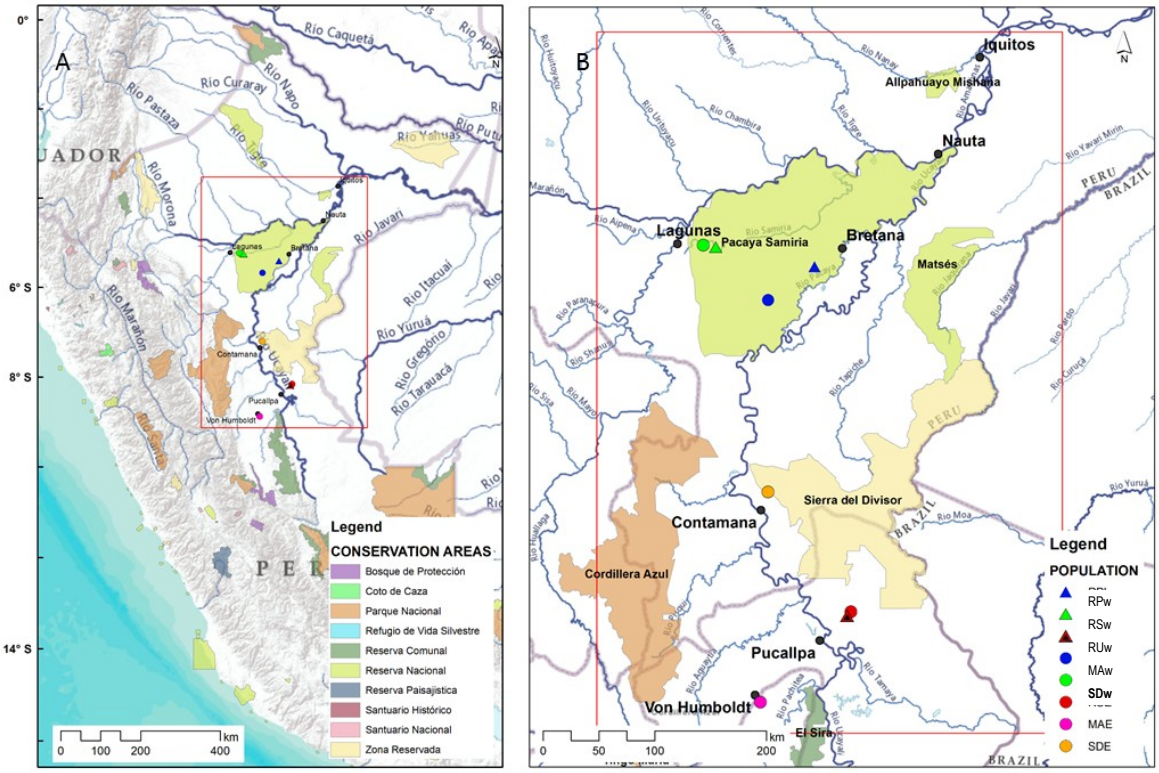

Figure 4. (A) Map of South America highlighting the study area; (B) Map with the rivers location, conservation areas and sampled populations located in the Samiria (RSw and RSI), Pacaya (RPw and RPI), and Utiquinia (RUw and RUI) rivers, and MAw and SDw populations.

5.3 GENETIC CHARACTERIZATION

To detect polymorphism in the studied *I. edulis* and *I. ingoides* populations it was used the simple sequence repeat (SSR) also called misrosatellite method, which is frequently used for plant genotyping.

5.3.1 DNA extraction

The total genomic DNA was extracted from dried young leaves, using the Invitek, Invisorb® Spin Plant Mini Kit following the manufacturer's instructions. In the final step of the protocol the extracted DNA was eluted with 25 µl of elution buffer and stored. Evaluation of the quantity and quality of isolated DNA was not performed, because our laboratory was not equipped with a measuring instrument. The extracted DNA of each sampled tree was visualised via gel

electrophoresis. To perform the DNA visualization it was used 2 µl of DNA and mixed with 1.2 µl of Loading dye then pipeted to the wells of fluorescently labeled (1.5 µl SYBR Green) 1% agarose gel (100 mg agarose in 100 ml 1xTBE buffer) (Figure 5).

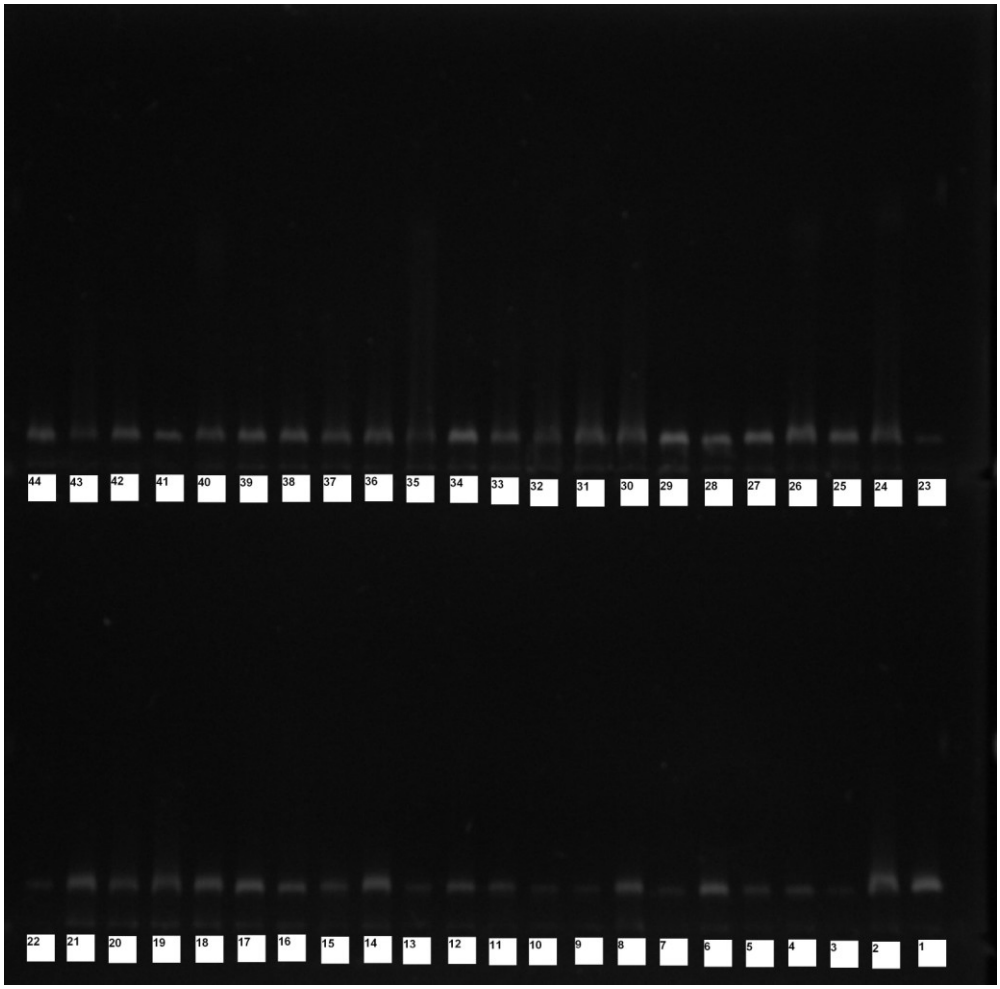

Figure 5. DNA visualization for samples 1 to 44.

5.3.2 PCR amplification

Four microsatellite markers were used to genotype all the individuals, *Pel*5 (Dayanandan et al. 1997) and *Inga*03, *Inga*08 and *Inga*33 (Hollingsworth et al. 2005). For following microsatellite detection, each forward primer was fluorescently labelled at 5´ end (6-FAM, NED or VIC) (Table 4).

Table 4. Primer sequences used in the study (source: author; Dayanandan et al. 1997; Hollingsworth et al. 2005).

Locus	Position	Fluorescent labeling	Primer sequences (from 5′)	Repeat motif	T_a (°C)	Range (bp)
*Inga*03	F	6-FAM	TTCCAAGCTTATACAAACCTCC	(CA)10	59	61-93
	R		AGATCCGTACGTGTGATGGT		59	
*Inga*08	F	NED	TTTGAGATGAATAGAGAAAGCC	(CT)4GT(CT)3	55	136-168
	R		GATTTAGCTTGTTGGTGTTTG		55	
*Inga*33	F	NED	TATAACCGATTCACCCCTTGATG	(CT)8(CA)8	55	214-240
	R		AAAGCACTATAAGATATGTGTGTC		55	
*Pel*5	F	VIC	TCTCTGCACACAGGAACCCTTTTGC	(AAAG)6	55	178-218
	R		CCCAGAAATAAGGCTCTTTTGCACA		55	

F: Forward molecular marker; R: Revesrse molecular marker; T_a: Annealing temperature

The singleplex PCR reaction was performed for each sample and primer pair in10 µl volume. The concentrations of the reaction components were as following: dNTPs 0.05 µM, $MgCl_2$ 2 mM, 5x GoTaq Flexi Buffer 1X, forward primer (0.1 µM) and reverse primer (0.1 µM), GoTaq® Flexi DNA Polymerase (Promega) 0.5 U (Table 5). Amplifications were undertaken in Biometra® T1 Thermocycler using the following profile: 95 °C for 2 min; 95 °C for 15 s, 55 °C (*Inga*03) and 59 °C (*Inga*08, *Inga*33 and *Pel*5) for 30 s, 72 °C for 30 s, 30 cycles; 72 °C for 15 min (Table 6).

Table 5. PCR composition (source: author).

Components	Final Volume	Final Conc.
Water, nuclease-free	6 µl	
dNTPs (1.25 mM)	0.4 µl	0.05 mM
$MgCl_2$ Solution (25 mM)	0.8 µl	2 mM
5x Green GoTaq® Flexi Buffer	2 µl	1 X
Primer F (10 µM, +dye)	0.1 µl	0.1 µM
Primer R (10 µM)	0.1 µl	0.1 µM
GoTaq® DNA Polymerase (5 u/µl)	0.1 µl	0.5 U
template DNA	0.5 µl	(10 -50 ng)
Total	10 µl	

Table 6. PCR profile (Hollingsworth et al. 2005).

Temperature (°C)	Time	Cycles
95	2 min	
95	15 s	
55/59	30 s	
72	30 s	30
72	15 min	

To verify the sucess of the PCR reaction the amplicons were visualized via gel electrophoresis. It was used 2 µl of the PCR product mixed with 1.2 µl of Loading dye and then pipeted to the wells of fluorescently labeled (2 µl SYBR Green) 2% agarose gel (200 mg agarose in 100 ml 1xTBE buffer). It was used 1.2 µl GeneRuler™ 100bp Plus DNA Ladder (Fermentas Life Science).

5.3.3 Fragmentation

The reaction mixture for the fragmentation analysis was performed from the following components: Hi-Di™ Formamide (Applied Biosystems) (24.8 µl), GeneScan 500 ROX Size Standard (Applied Biosystems) (0.2 µl) and from 0.5 to1 µl of the PCR product. The amplified products were separated on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and run according to the manufacturer´s protocol Due to the different fluorescently labeled primers and fragmenths lengths it was possible to run the folowing pairs of amplikons *Inga*03 (6-fam – blue peak, 61-93 bp) with *Inga*33 (NED – yellow, 214 – 240 bp) and *Inga*08 (NED – yellow, 136 – 168 bp) with *Pel*5 (VIC – green, 178 – 218 bp) together in one test tube.

The denaturation was maintained by 5 min heating 95 °C, than cooled for 3 min on ice. In the genetic anlyzer it was used polymer POP-4™ (Thermo Fisher Scientific) and Applied Biosystems ™ 310 Running Buffer. It was used 47 cm long capilary. Conditions of injection were: 5 seconds /15 kV. The running temperature 60 °C. The time of run was between 25 and 30 minutes according to the fragment lengths. Fragments sizes were determined manually using the ROX500 internal size standard and the global southern algorithm implemented by ABI PRISM GeneMapper software version 4.0 (Applied Biosystems).

5.4 DATA ANALYSIS

5.4.1 *Inga edulis* morphology

The morphological data obtained during sampling of *I. edulis* following the designed characterisation descriptors for *I. edulis* were analysed using standard statistical parameters. In case of tree height, branching height, crown diameter, trunk diameter, leaf petiole length, leaflet length, leaflet width, rhachis wing length, rhachis wing width, number of leaflets, legume length, legume diameter, number of seeds per legume, seed length, seed width and seed weight were the data described using: minimal value (MIN), maximal value (MAX), arithmetic mean (MEAN), standard deviation (SD), middle number in the group (MEDIAN) and the most frequently occurring number in the numeric data set (MODUS). The trunk shape, bark colour, bark texture, mature leaf colour, rachis wing shape, seed shape and seed colour were expressed as a percentage. Due to the missing data no more morphological features mentioned in the proposed characterisation descriptors for *I. edulis* were analysed.

The *I. edulis* legume length was emphasized in this study to draw a null hypothesis based on the domestication process, since the trees were selected for their legume length (H0: is the

legume length of the domesticated trees higher than the wild ones?). In no other characteristic we can not draw such a hypothesis.

The legume length was assessed with the values measured in 448 pods sampled from 259 trees of the *I. edulis* populations from the Peruvian Amazon, and all the individual values were used in the following analysis (not averaged per tree). The legume lengths in cultivated populations originated in Selva Central, Ucayali and Loreto regions and in wild populations normality was tested with the Kolmogorov-Smirnov test, and all the 4 groups displayed normal distribution but one, the wild trees group. A non-parametric Kruskal-Wallis test (k independent samples) was performed to check for significant differences in the groups' average followed by the non-parametric Mann-Whitney U post-hoc test (Sokal and Rohlf 1997). The statistical analyses were performed using the IBM SPSS Statistics software vs. 22.

5.4.2 Molecular data

The estimated genetic diversity parameters included the average number of alleles per locus (A), the effective number of alleles (N_e), the number of private alleles (P_a), the mean allelic richness (R_S) that uses a rarefaction index to consider differences in sample size (El Mousadik and Petit 1996), the observed (H_O) and expected (H_E) heterozygosities (Nei 1987), the fixation index F_{IS} and the among populations differentiation F_{ST} (Weir and Cockerham 1984). The analyses were performed using GenAlEx 6.5 (Peakall and Smouse 2012), except for the allelic richness (A_R) which was computed using FSTAT 2.9.3 (Goudet 1995). Using the Genepop 4.3 software (Rousset 2008), it was tested the heterozygote deficiency for each population.

Genetic variation at the level of populations and groups (i.e. cultivated and wild *I. edulis* populations) was investigated with a hierarchical analysis of molecular variance (AMOVA), which partitions the total variance into covariance components due to inter-group differences, inter-populations within groups differences, and inter-population differences, in the Arlequin 3.5 software (Excoffier and Lischer 2010). Similarly it was estimated the AMOVA for the grouping structure of *I. edulis* in comparison with *I. ingoides* using locus-by-locus analysis. The variance components and genetic variation was estimated using a non-hierarchical and hierarchical analysis considering all of the populations or the two groups (species), respectively. Levels of significance were determined by computing 1,000 random permutation replicates.

A Bayesian clustering method was performed in the STRUCTURE vs. 2.3.4. software (Pritchard et al. 2000) to infer population genetic structure. The number of genetic clusters (K) was estimated and the individuals sampled from cultivated and wild populations were fractionally assigned to the inferred groups. Afterwards, the allele frequencies were estimated in

each of the *K* groups and the proportion of genome derived from each group for each tree. Model allowing population admixture and correlated allele frequency (Pritchard et al. 2000) was applied. However, due to the weak population structure found in the *I. edulis* populations, a model that incorporated *a priori* sampling location information (Hubisz et al. 2009), i.e. a "locprior" model. This improved model has the advantage of allowing cryptic structures to be detected at a lower level of divergence and does not bias towards detecting structure spuriously when none is present, helpful in situations when the standard structure models do not provide a clear signal of structure (Hubisz et al. 2009). Two groups of populations were used as priors, i.e. cultivated and wild populations (see Table 2). The alternative ancestry prior $1/K$ was used due to unbalanced population sampling (Wang 2017). The number of clusters (*K*) was set from one through twenty-seven and the simulation was run ten times at each *K* value to confirm the repeatability of the results. Each run comprised a burn-in period of 25,000, followed by 100,000 Markov chain Monte Carlo (MCMC) steps. We used the ΔK distribution statistic of Evanno et al. (2005) to determine the most appropriate number of genetic clusters through the detection of the second rate of change in LnP(D). Hence, the STRUCTURE output data were parsed using the STRUCTURE HARVESTER (Earl and von Holdt 2012) to determine the optimal *K* value following the referred method. Alignment of cluster assignments across replicate analyses was then conducted in CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) and subsequently visualized using STRUCTURE PLOT (Ramasamy et al. 2014). The results of Bayesian clustering were further mapped in the ArcGIS® Desktop version 10.2 software (Ormsby 2010).

A principal coordinate analysis (PCoA) was computed based on the pairwise Nei's genetic distance matrix and the Hardy-Weinberg equilibrium (HWE) was tested for each wild *I. edulis* and *I. ingoides* population and locus (Markov-Chain method), the linkage disequilibria (LD) tests were done for all loci combinations, and the average frequency of null alleles were computed per population. The grouping structure was further explored using a locus-by-locus analysis of molecular variance (AMOVA), implemented with the Arlequin 3.5 software (Excoffier and Lischer 2010). The variance components and genetic variation was estimated using a non-hierarchical and hierarchical analysis considering all of the populations or the two groups (species), respectively. The significance values were computed by a permutation test from 1,000 permuted matrices.

A Bayesian clustering method was carried out using the STRUCTURE version 2.3.3 software (Pritchard et al. 2000) to estimate the number of genetic clusters (*K*) and to fractionally assign individuals of both *Inga* species to the inferred groups. It was applied the model which allows population admixture and correlated allele frequency. The *K* was set from one to eight,

and the simulation was run ten times at each *K* value to confirm the repeatability of the results. Each run comprised a burn-in period of 25,000, followed by 100,000 Markov chain Monte Carlo (MCMC) steps. Afterwards, the STRUCTURE output data were parsed using the program Structure-sum (running under the R platform) (Ehrich et al. 2007), mainly to determine the optimal *K* value following Nordborg et al. (2005) and Evanno et al. (2005) methods. Therefore, we used the ΔK distribution statistic of Evanno et al. (2005) to determine the most appropriate number of genetic clusters through the detection of the second rate of change in LnP(D). In addition, the similarity coefficient between 10 structure runs was computed, and for values higher than 0.9 we assumed that each run ended with a similar result. An alignment of cluster assignments across replicate analyses was then conducted in the CLUMPP 1.1.2 software (Jakobsson and Rosenberg 2007), and subsequently visualized using DISTRUCT 1.1 (Rosenberg 2004).

6 RESULTS

6.1 CHARACTERISATION DESCRIPTORS FOR *I. EDULIS*

This characterization descriptors list was intended to be comprehensive for the descriptors it contains. Furthermore, these may include a limited number of additional traits thought desirable by a consensus of user of the particular crop. We do not, however, assume that curators will characterize accessions of their collections using all descriptors given. Descriptors should be used when they are useful to users, either collection curators for the management and maintenance of their germplasm material or to all other users of plant genetic resources for promoting their sustainable use.

6.1.1 Tree traits

Tree shape (Figure 6)

1 Roundish; 2 Broad shape; 3 Shrubby; 4 Vase shape; 5 Other (specify in Note)

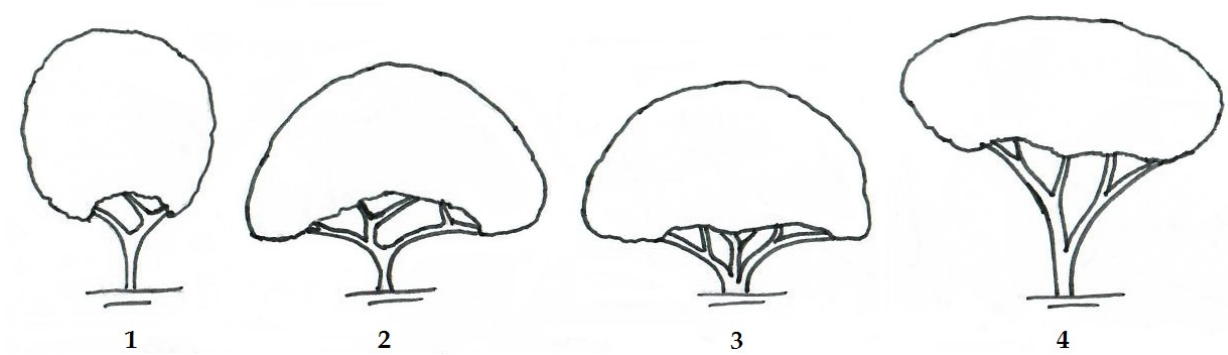

Figure 6. Tree shape (Alexandr Rollo 2018)

Tree height [m]

Measured height of mature trees from ground level to the top of the tree.

Branching height [m]

Measured height of mature tree from ground level to the branching node.

Crown diameter [m]

Measured as the mean diameter using two directions.

Trunk diameter [cm]

Record diameter at 10 cm height (If the tree shape is shrubby; If branching node is lower than 10 cm above ground measure all stems) or in 30 cm (in case of other tree shapes).

Trunk shape (Figure 7)

1 Cylindrical; 2 Buttresed; 3 Buttresed and bole fluted; 4 Other (specify in descriptor Note)

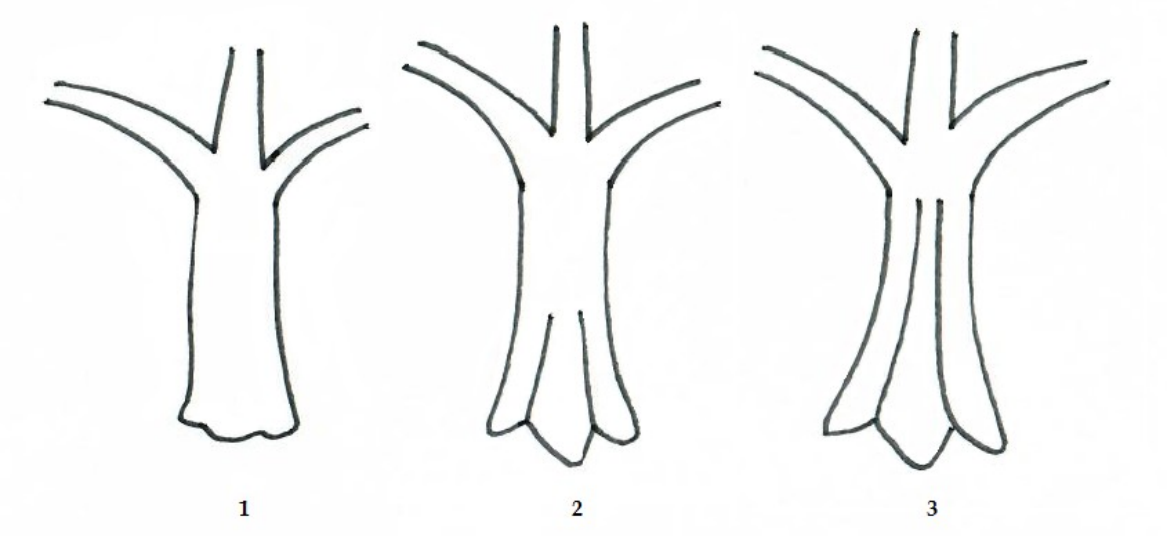

Figure 7. Trunk shape (Alexandr Rollo 2018)

Bark colour

If possible use colour codes from the Royal Horticulture Society. If these are not available, use the following colour codes.

1 Pale greyish; 2 Other (specify in descriptor Note)

Bark texture

1 Smooth; 2 Lenticelate with hoop marks; 3 Other (specify in descriptor Note)

Bark slash colour

If possible use colour codes from the Royal Horticulture Society. If these are not available, use the following colour codes.

1 Pink; 2 Pale brown; 3 Other (specify in descriptor Note)

Bark slash exudate

1 Clear; 3 Other (specify in descriptor Note)

6.1.2 Leaf traits

Randomly select 10 mature and healthy leaves and record the average.

Number of leaflets pairs of mature leaves (Figure 8)

1 Three leaflets pairs; 2 Four leaflets pairs; 3 Five leaflets pairs; 4 Six leaflets pairs; 5 Seven leaflets pairs; 6 Other (specify in descriptor Note)

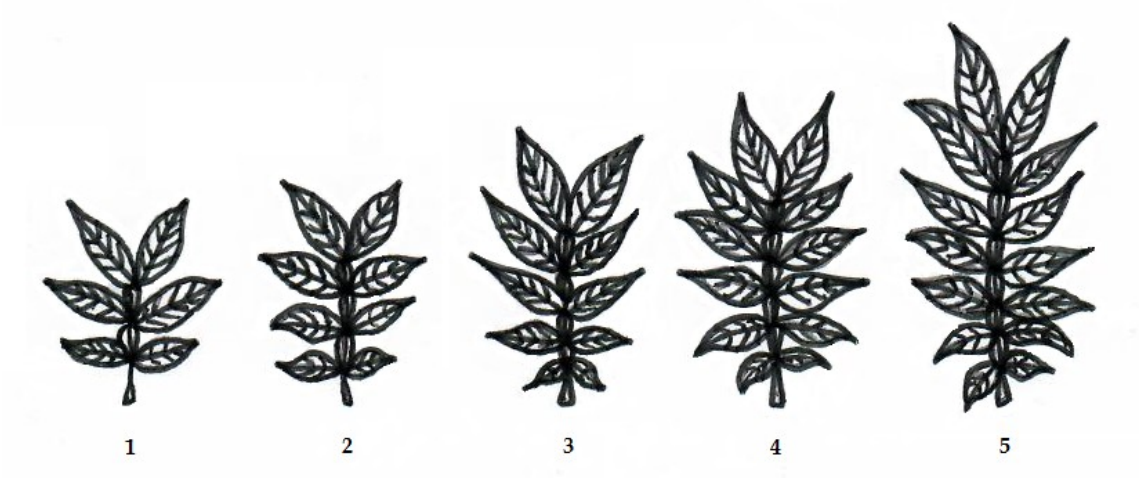

Figure 8. Number of leaflets pairs of mature leaves (Alexandr Rollo 2018)

Leaflet blade shape (Figure 9)

1 Elliptic; 2 Oblong; 3 Ovate; 4 Obovate; 5 Lanceolate; 6 Other (specify in descriptor Note)

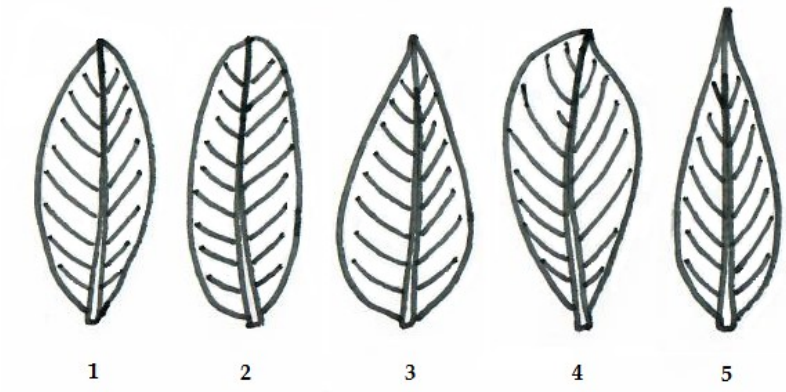

Figure 9. Leaflet blade shape (Alexandr Rollo 2018)

Lower lamina puberulance

1 Scabrid; 2 Scabrid to crisped; 3 Crisped; 4 Other (specify in descriptor 4.1.4.6 Note)

Leaflet venation (Figure 10)

1 Eucamptodromous; 2 Brochydodromous; 3 Other (specify in descriptor Note)

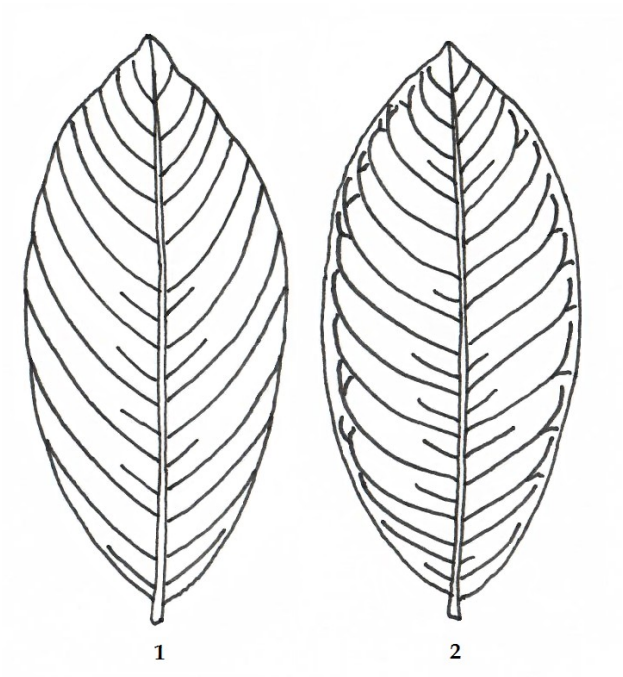

Figure 10. Leaflet venation (Alexandr Rollo 2018)

Number of secondary veins pairs

Take one leaflet from every second pair of leaflets and record the average number of secondary veins pairs of the 10 selected leaves.

Secondary veins pairs position (Figure 11)

1 Parallel; 2 Slightli convergent; 3 Other (specify in descriptor Note)

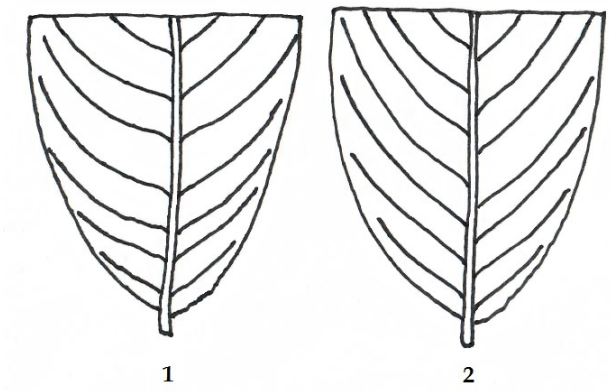

Figure 11. Secondary veins pairs position (Alexandr Rollo 2018)

Intersecondaries length (Figure 12)

1 Moderate; 2 Short; 3 Other (specify in descriptor Note)

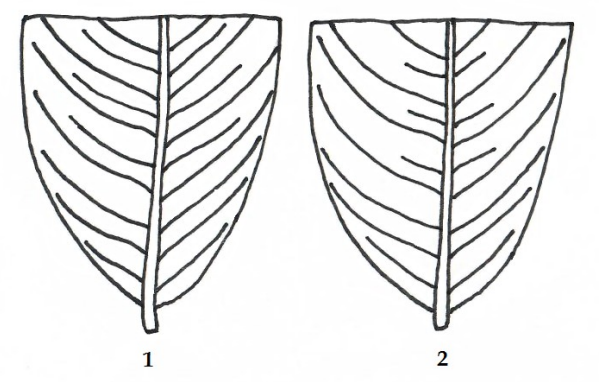

Figure 12. Secondary veins length (Alexandr Rollo 2018)

Leaflet apex shape (Figure 13)

1 Acute; 2 Mucronate; 3 Apiculate; 4 Cuspidate; 5 Acuminate; 6 Other (specify in descriptor Note)

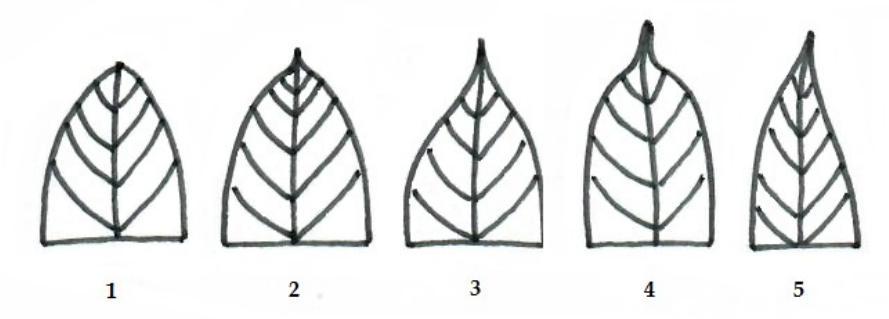

Figure 13. Leaflet apex shape (Alexandr Rollo 2018)

Leaflet base shape (Figure 14)

1 Acute; 2 Asymmetric; 3 Rounded; 4 Truncate; 5 Cordate; 6 Other (specify in descriptor Note)

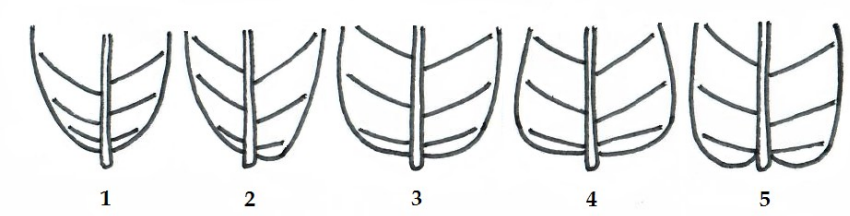

Figure 14. Leaflet base shape (Alexandr Rollo 2018)

Mature leaf colour

If possible use colour codes from the Royal Horticulture Society, or specify in desriptor Note)

Leaf rhachis length [cm]

Measure from the base of the petiole to the top foliar nectary.

Leaf petiole length [cm]

Measure from the base of the petiole to the base of the first foliar nectary.

Leaflet petiole length [cm]

Take one leaflet from every second pair of leaflets of the 10 selected leaves and record the average leaflet petiole length. Measure from the base of the leaflet petiole to the base of leaflet blade.

Leaflet length [cm]

Take one leaflet from every second pair of leaflets of the 10 selected leaves and record the average leaflet length. Measure in the longest part of the leaflet.

Leaflet width [cm]

Take one leaflet from every second pair of leaflets of the 10 selected leaves and record the average leaflet width. Measure in the widest part of the leaflet.

Petiole wing (Figure 15)

0 Absent; 1 Present

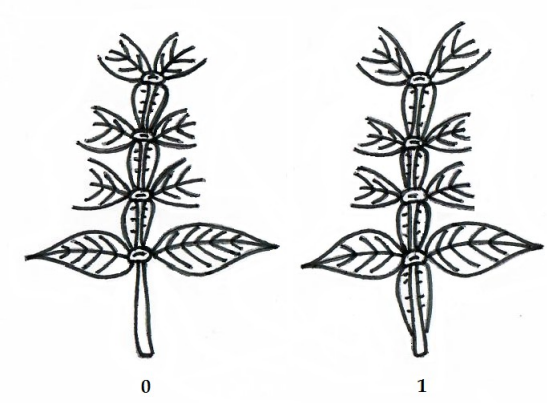

Figure 15. Petiole wing (Alexandr Rollo 2018)

Rhachi wing shape (Figure 16)

Observe the terminal rhachis wing of the 10 selected leaves.

1 Obovate; 2 Cuneate; 3 Linear; 4 Elliptic; 5 Triangular; 6 Obcordate; 7 Other (specify in descriptor Note)

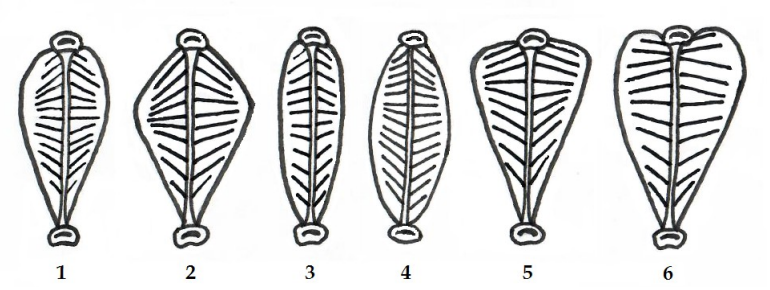

Figure 16. Rhachis wing shape (Alexandr Rollo 2018)

Rhachi wing defformation (Figure 17)

Observe second rhachis wing (between second and third foliar nectary) of the 10 selected leaves.

0 Absent; 1 One side of the wing base turned inward; 2 Wing base turned inward;

3 One side of the wing base is shortened; 4 Wing base does not reach to the second foliar nectary; 5 Other (specify in descriptor Note)

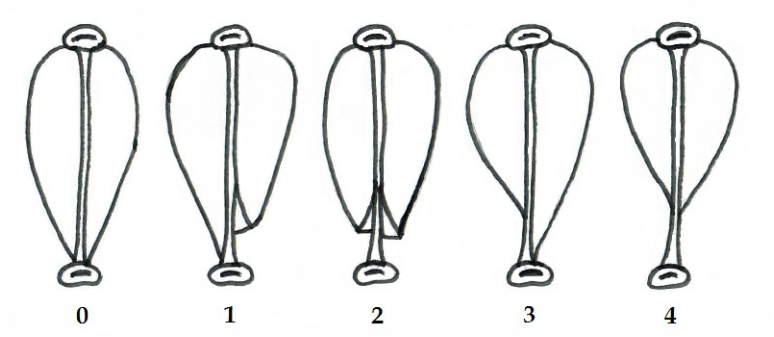

Figure 17. Rhachis wing deformation (Alexandr Rollo 2018)

Rhachis wing length [cm]

Take the second rhachis wing (between second and third foliar nectary) of the 10 selected leaves and record the average rhachis wing length. Measure in the longest part of the rhachis wing.

Rhachis wing width [cm]

Take the rhachis wing (between second and third foliar nectary) of the 10 selected leaves and record the average rhachis wing width. Measure in the widest part of the rhachis wing.

6.1.3 Legume traits

Randomly select 10 mature and healthy legumes at least with their pedicels per tree and record the average.

Legume shape (Figure 18)

Record the predominant shape using 10 fruits per tree.

1 Cylindrical; 2 Straight twisted; 3 Spirally twisted; 4 Other (specify in descriptor Note)

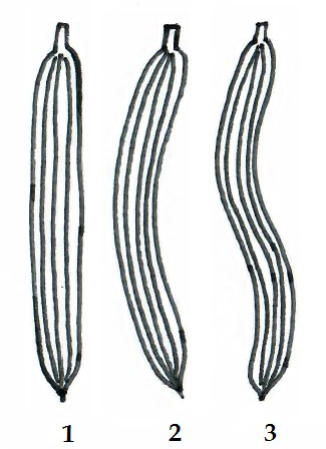

Figure 18. Legume shape (Alexandr Rollo 2018)

Legume surface undulation (Figure 19)

Record the predominant shape using 10 fruits per tree.

1 Plane; 2 Undulated; 3 Other (specify in descriptor Note)

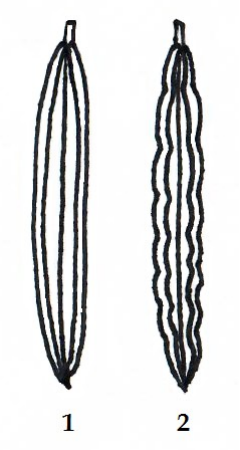

Figure 19. Legume surface undulation (Alexandr Rollo 2018)

Legume pedicel insertion (Figure 20)

1 Vertical; 2 Oblique; 3 Other (specify in descriptor Note)

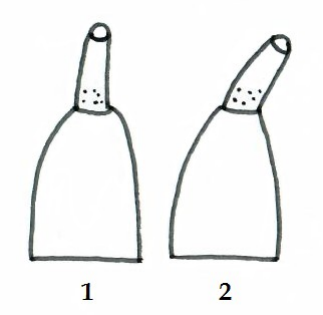

Figure 20. Legume pedicel insertion (Alexandr Rollo 2018)

Legume base (Figure 21)

1 Acute; 2 Obtuse; 3 Tapered; 4 Other (specify in descriptor Note)

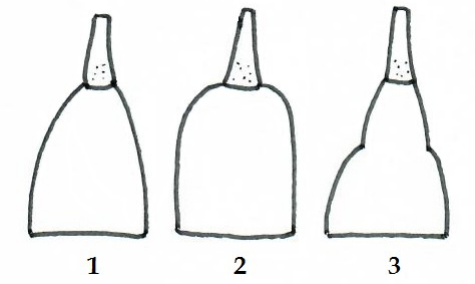

Figure 21. Legume base (Alexandr Rollo 2018)

Legume apex shape (Figure 22)

1 Acute; 2 Rostrate; 3 Other (specify in descriptor Note)

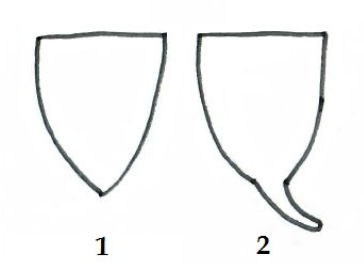

Figure 22. Legume apex shape (Alexandr Rollo 2018)

Legume faces (Figure 23)

1 Completely covered by expanded margins; 2 Largerly covered by expanded margins; 3 Exposed 2-3 mm; 4 Other (specify in descriptor Note)

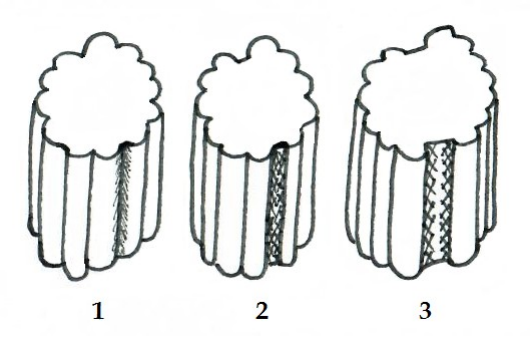

Figure 23. Legume faces (Alexandr Rollo 2018)

Legume section outline (Figure 24)

1 Circular; 2 Elliptical; 3 Other (specify in descriptor Note)

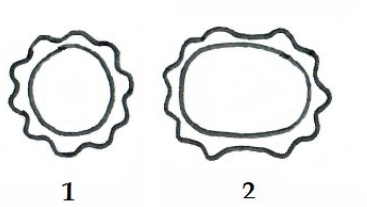

Figure 24. Legume section outline (Alexandr Rollo 2018)

Legume length [cm]

Measure from the base to the tip of the pod apex using 10 fruits per tree.

For the detail information on legume length see Results section (6.2).

Legume diameter [cm]

Measured at the widest point of the legume using 10 fruits per tree.

Number of seeds per legume

Calculate all seeds in the legume using 10 fruits per tree.

Legume shell hairiness

0 Not hairy; 1 Puberulose

Legume shell colour

If possible use colour codes from the Royal Horticulture Society, or specify in descriptor Note.

Pedicel length [cm]

Measure at the longest length, but only if the complete pedicel is available using 10 fruits per tree.

Total weight of 10 legumes [g]

Determine the weight of 10 fruits.

Pulp weight of 10 legumes [g]

Remove the pulp-covered seeds from the 10 opened fruit shells, separate the pulp from the seeds and determine its weight.

Seed weight of 10 legumes [g]

Remove 10 seeds from the pulp-cover and determine their weight using 10 fruits per tree.

Legume shell weight of 10 legumes [g]

Remove the seeds and the pulp-cover from 10 fruit shells and determine their weight.

Pulp colour of fresh fruit

If possible use colour codes from the Royal Horticulture Society, or specify in descriptor Note.

Adherence of fruit pulp to seed

(Scratch with your finger nails)

0 Absent; 1 Weak; 2 Intermediate; 3 Strong

Pulp texture of ripe fruit

1 Soft; 2 Intermediate; 3 Firm

Pulp consistency

1 Watery; 2 Fleshy; 3 Other (specify in desriptor Note)

Pulp sweetness

0 Absent; 1 Slightly sweet; 2 Sweet; 3 Very sweet

Pulp taste

0 Absent; 1 Reminiscent of cinnamon; 2 Reminiscent of blueberry; 3 Other (specify in descriptor Note).

6.1.4 Seed traits

Randomly select 10 healthy seeds out of the total seeds from the 10 collected fruits from one tree and record the average.

Seed shape (Figure 25)

1 Elliptical; 2 Tetragonal; 3 Triangular; 4 Other (specify in desriptor Note)

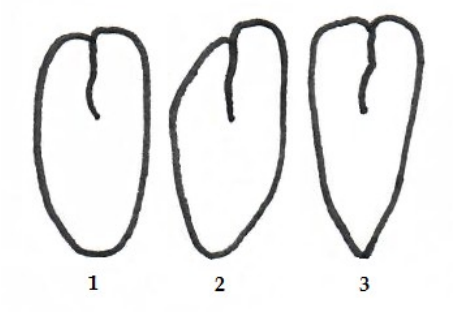

Figure 25. Seed shape (Alexandr Rollo 2018)

Seed colour

If possible use colour codes from the Royal Horticulture Society, or specify in descriptor Note.

Seed length [mm]

Measure in the longest part of the seed.

Seed width [mm]

Measured in the widest part of the seed.

Seed thickness [mm]

Measured at 90°from the measurement of the seed width.

Note

Specify any additional information here.

Any suggestions for improvement of the "Characterization descriptors for *I. edulis*" will be highly appreciated.

6.2 *INGA EDULIS* MORPHOLOGICAL VARIABILITY

The characterization descriptors have been designed during the field work according to the literature, but also to the authors own terrain observation during the field work. This was also the reason why some of the mentioned features from the characterization descriptors have been observed only on small sample size, due to the fact that the existence of the variability of the feature was discovered during the field work.

It was not possible to use the gathered morphological data to determine specific conclusions and comment them in the discusion section, instead of the legume length. Some parameters were not evaluated because of the scope of the thesis, or due to unsufficient sample size. Anyway, these data could open the discussion on phenotypic variance of *I. edulis* species in the Amazonian Peru and could be an interesting stimulus for further studies focused on this species.

The results obtained after evaluation following the designed characterization descriptors for the studied *I. edulis* populations are detailed in (Table 7, 8, 9, 10, 11 and 12).

Inga ingoides morphological characterisations were not evaluated.

Table 7. Morphological desription of *I. edulis* population and wild and cultivated *I. edulis* populations – quantitative characteristics. Number of samples (N), minimal value (MIN), maximal value (MAX), arithmetic mean (MEAN), standard deviation (SD), middle number in the group (MEDIAN) and the most frequently occurring number in the data set (MODUS).

Pop.	basic statistic	Tree height [m]	Branching height [m]	Crown diameter [m]	Trunk diameter [cm]	Leaf petiole length [cm]	Leaflet length [cm]	Leaflet width [cm]	Rhachis wing length [cm]	Rhachis wing width [cm]	No. of leaflets	Legume length [cm]	Legume diameter [cm]	No. of seeds per legume	Seed length [mm]	Seed width [mm]	10 seed weight [g]
I. edulis	**N**	258	250	178	251	2070	2070	2070	2070	2070	2070	448	319	210	1140	1140	1000
	MIN	1,5	0,1	1,5	1,6	2,2	6,6	2,3	2,2	0,7	4	13	0,5	8	1,4	0,8	4,3
	MAX	18	3,5	18	61,4	7,0	17	8,3	6,4	3,1	6	148	5,3	28	6,5	2,0	81
	MEAN	6,6	0,8	7,3	18,1	4,1	12,0	5,1	3,6	1,5	4,9	72,6	2,6	19,3	4,0	1,5	37,7
	SD	2,8	0,7	3,1	9,7	0,8	2,0	1,0	0,7	0,4	0,6	26,4	0,8	4,2	0,9	0,3	17,4
	MEDIAN	6	0,7	6,5	17,2	4,0	11,9	4,9	3,5	1,4	5	76	2,7	19	4,0	1,5	35,7
	MODUS	5	0,1	5	23,9	4,9	12,0	4,5	2,9	1,4	5	31	2,9	19	4,4		37,4
I. edulis cultivated	**N**	196	193	172	195	1730	1730	1730	1730	1730	1730	329	281	195	1090	1090	940
	MIN	1,5	0,1	1,5	4,5	2,2	6,6	2,3	2,2	0,7	4	13	1,0	8	1,8	0,9	8,5
	MAX	12	3,5	16	46	6,8	17,0	8,1	4,9	3,1	6	148	5,3	28	6,5	2,0	81
	MEAN	6,2	0,8	7,2	18,0	4,1	12,0	5,0	3,5	1,5	4,9	83	2,7	19,3	4,1	1,5	39,4
	SD	2,2	0,6	2,9	7,7	0,8	1,9	0,9	0,6	0,4	0,6	20,0	0,8	4,3	0,8	0,2	16,6
	MEDIAN	6	0,6	6,5	17,5	4,1	11,9	4,9	3,5	1,4	5	83	2,8	19	4,2	1,5	37,4
	MODUS	5	0,1	5	23,9	4,2	11,9	4,5	3,0	1,4	5	104	2,9	19	4,4	1,5	
I. edulis wild	**N**	62	57	6	56	340	340	340	340	340	340	119	38	15	60	60	60
	MIN	1,5	0,2	2,5	1,6	2,3	6,6	2,3	2,5	0,8	4	26,6	0,5	14	1,4	0,8	4,3
	MAX	18	2,5	18	61,4	7,0	17	8,3	6,4	2,6	6	59	2,1	25	3,3	1,4	24,8
	MEAN	7,8	1,5	9,6	18,6	3,9	12,0	5,1	3,8	1,4	4,9	39	1,7	19,4	2,4	1,0	12,1
	SD	3,8	1,1	7,2	14,7	0,9	2,0	1,0	0,9	0,4	0,4	8,7	0,5	4,2	0,7	0,2	6,8
	MEDIAN	7	1,7	8,8	14,2	3,7	11,9	4,9	3,7	1,4	5	36	1,9	20	2,5	0,9	10,8
	MODUS	7	2,5		6,4	3,7	12,0	4,5	2,9	1,4	5	31					

Table 8. Morphological desription of *I. edulis* populations cultivated in Selva Central, Ucayali and Loreto regions – quantitative characteristics. Number of samples (N), minimal value (MIN), maximal value (MAX), arithmetic mean (MEAN), standard deviation (SD), middle number in the group (MEDIAN) and the most frequently occurring number in the data set (MODUS).

Pop.	basic statistic	Tree height [m]	Branching height [m]	Crown diameter [m]	Trunk diameter [cm]	Leaf petiole length [cm]	Leaflet length [cm]	Leaflet width [cm]	Rhachis wing length [cm]	Rhachis wing width [cm]	No. of leaflets	Legume length [cm]	Legume diameter [cm]	No. of seeds per legume	Seed length [mm]	Seed width [mm]	10 seed weight [g]
I. edulis cultivated **Selva Central**	**N**	44	45	45	45	450	450	450	450	450	450	94	93	76	410	410	410
	MIN	2,5	0,1	3	4,8	2,8	6,6	2,3	2,5	0,9	4	38	1,4	8	2,2	1,0	10,1
	MAX	10	3,5	11	38,2	5,4	17	8,3	4,8	2,6	6	127	4,0	26	5,6	1,9	81
	MEAN	6,6	1,1	7,1	18,6	4,2	12,0	5,1	3,5	1,5	5,1	78	2,9	19,0	3,7	1,4	34,9
	SD	2,0	0,7	2,3	7,5	0,6	2,0	1,0	0,5	0,4	0,6	18,5	0,6	3,8	0,8	0,2	16,6
	MEDIAN	7	0,9	7	17,8	4,2	11,9	4,9	3,5	1,4	5	79	2,9	19	3,6	1,5	29,9
	MODUS	7	0,1	10	15,3	3,9	12,0	4,5	3,6	1,4	5	69	2,7	18		1,7	
I. edulis cultivated **Ucayali**	**N**	72	71	64	72	640	640	640	640	640	640	120	112	110	590	590	440
	MIN	1,5	0,1	1,5	4,5	2,2	6,6	2,3	2,4	0,8	4	33	1,8	11	1,8	0,9	8,5
	MAX	11	2,5	15,5	46	6,1	17	8,3	4,9	3,1	6	132	4,3	28	6,5	1,9	77
	MEAN	5,3	0,6	6,2	15,3	3,9	12,0	5,1	3,5	1,5	4,7	80	3,0	19,8	4,3	1,5	42,2
	SD	2,0	0,4	2,6	7,7	0,8	2,0	1,0	0,7	0,5	0,5	18,6	0,6	4,4	0,8	0,2	14,5
	MEDIAN	5	0,5	5,5	13,4	3,9	11,9	4,9	3,5	1,4	5	82,8	3,0	19,5	4,4	1,6	40,2
	MODUS	5	0,6	5	11,1	3,3	12,0	4,5	4,4	1,3	5	81	2,9	19	3,6	1,5	
I. edulis cultivated **Loreto**	**N**	80	77	63	78	640	640	640	640	640	640	115	76	9	90	90	90
	MIN	2,5	0,1	2,5	5,9	2,8	6,6	2,3	2,2	0,7	4	13	1,0	10	3,4	1,0	20,5
	MAX	12,0	2,5	16,0	43,9	6,8	17,0	8,3	4,9	2,9	6	148	5,3	22	5,2	2,0	73,2
	MEAN	6,7	0,8	8,3	20,1	4,3	12,0	5,1	3,6	1,5	4,7	90	2,0	16,5	4,1	1,5	46,1
	SD	2,3	0,7	3,1	7,1	0,8	2,0	1,0	0,7	0,4	0,5	21,2	0,9	4,2	0,5	0,3	23,3
	MEDIAN	6	0,7	8	20,2	4,2	11,9	4,9	3,5	1,4	5	88,2	1,7	18,5	4,0	1,4	39,8
	MODUS	5	0,3	6,5	23,9	3,9	12,0	4,5	3,4	1,4	5	110,0	1,3	19			

Table 9. Morphological desription of *I. edulis* populationion – qualitative characteristics. Number of samples (N).

Pop.	Trunk shape	Bark colour (RHS code)	Bark texture	Mature leaf colour	Rhachis wing shape	Rhachis wing deformation	Legume shape	Seed shape	Seed colour
	N 221	N 177	N 169	N 166	N 209	N 215	N 144	N 113	N 94
I. edulis	Cylindrical 182 (82%)	Grey green (133D) 2 (1%)	Smooth 81 (48%)	Dark green (136A) 19 (11%)	Obovate 119 (57%)	Absent 136 (63%)	Cylindrical 63 (44%)	Elliptical 92 (81%)	White (N155D) 1 (1%)
	Buttresed 39 (18%)	Dark green (137A) 1 (0.6%)	Lenticelate with hoop marks 88 (52%)	Dark green (137A) 106 (64%)	Cunate 37 (17.5%)	Present 79 (37%)	Straight twisted 8 (6%)	Tetragonal 18 (16%)	Grey brown (N199A) 1 (1%)
		Grey brown (199A) 74 (42%)		Dark green (141A) 36 (22%)	Linear 10 (5%)		Spirally twisted 22 (15%)	Triangular 3 (3%)	Dark brown (200B) 1 (1%)
		Grey brown (N199A) 12 (7%)		Brown green (137C) 5 (3%)	Elliptic 35 (16.5%)		Cylindrical and straight twisted 24 (16.5%)		Brown purple (N77A) 35 (37.5%)
		Grey brown (199C) 75 (42%)			Triangular 6 (3%)		Cylindrical and spirally twisted 8 (5.5%)		Brown purple (187A) 1 (1%)
		Grey brown (N199C) 2 (1%)			Obcordate 2 (1%)		Straight twisted and spirally twisted 9 (6%)		Brown purple (N186A) 3 (3%)
		Brown green (191A) 1 (0.6%)					Cylindrical, straight twisted and spirally twisted 10 (7%)		Black (N187A) 1 (1%)
		Brown green (194A) 7 (4%)							Black (202A) 51 (54.5%)
		Brown green (139C) 1 (0.6%)							
		Yellow brown (N167A) 1 (0.6%)							
		Yellow brown (168D) 1 (0.6%)							

Table 10. Morphological desription of cultivated and wild *I. edulis* populationions – qualitative characteristics. Number of samples (N).

Pop.	Trunk shape	Bark colour (RHS code)	Bark texture	Mature leaf colour	Rhachis wing shape	Rhachis wing deformation	Legume shape	Seed shape	Seed colour
I. edulis cultivated	N 177	N 170	N 163	N 166	N 173	N 175	N 142	N 110	N 92
	Cylindrical 138 (78%)	Grey green (133D) 2 (1%)	Smooth 78 (48%)	Dark green (136A) 19 (11%)	Obovate 101 (58.5%)	Absent 111 (63%)	Cylindrical 62 (44%)	Elliptical 89 (81%)	White (N155D) 1 (1%)
	Buttresed 39 (22%)	Dark green (137A) 1 (0.6%)	Lenticelate with hoop marks 85 (52%)	Dark green (137A) 106 (64%)	Cunate 29 (17%)	Present 64 (37%)	Straight twisted 8 (5.5%)	Tetragonal 18 (16%)	Grey brown (N199A) 1 (1%)
		Grey brown (199A) 68 (40%)		Dark green (141A) 36 (22%)	Linear 7 (4%)		Spirally twisted 22 (15.5%)	Triangular 3 (3%)	Brown purple (N77A) 34 (37%)
		Grey brown (N199A) 12 (7%)		Brown green (137C) 5 (3%)	Elliptic 28 (16%)		Cylindrical and straight twisted 24 (17%)		Brown purple (187A) 1 (1%)
		Grey brown (199C) 75 (44%)			Triangular 6 (3.5%)		Cylindrical and spirally twisted 8 (5.5%)		Brown purple (N186A) 3 (3%)
		Grey brown (N199C) 2 (1%)			Obcordate 2 (1%)		Straight twisted and spirally twisted 8 (5.5%)		Black (N187A) 1 (1%)
		Brown green (191A) 1 (0.6%)					Cylindrical, straight twisted and spirally twisted 10 (7%)		Black (202A)51 (56%)
		Brown green (194A) 6 (4%)							
		Brown green (139C) 1 (0.6%)							
		Yellow brown (N167A) 1 (0.6%)							
		Yellow brown (168D) 1 (0.6%)							
I. edulis wild	N 44	N 7	N 6	N 7	N 36	N 40	N 2	N 3	N 2
	Cylindrical 44 (**100%**)	Grey brown (199A) 6 (86%)	Smooth 3 (50%)	Dark green (137A) 3 (64%)	Obovate 18 (50%)	Absent 25 (62.5%)	Cylindrical 1 (50%)	Elliptical 3 (100%)	Dark brown (200B) 1 (50%)
	Buttresed 0 (0%)	Brown green (194A) 1 (14%)	Lenticelate with hoop marks 3 (50%)	Dark green (141A) 4 (22%)	Cunate 8 (22%)	Present 15 (37.5%)	Straight twisted and spirally twisted 1 (50%)		Brown purple (N77A) 1 (50%)
					Linear 3 (8.5%)				
					Elliptic 7 (19.5%)				

Table 11. Morphological desription of cultivated *I. edulis* populationion in Selva Central – qualitative characteristics. Number of samples (N).

pop	Trunk shape	Bark colour (RHS code)	Bark texture	Mature leaf colour	Rhachis wing shape	Rhachis wing deformation	Legume shape	Seed shape	Seed colour
	N 45	N 44	N 45	N 45	N 45	N 44	N 41	N 39	N 29
***I. edulis* cultivated Selva Central**	Cylindrical 32 (71%) Buttresed 13 (29%)	Grey green (133D) 2 (4.5%) Dark green (137A) 1 (2.3%) Grey brown (199A) 14 (31.8%) Grey brown (N199A) 1 (2.3%) Grey brown (199C) 17 (38.6%) Brown green (191A) 1 (2.3%) Brown green (194A) 5 (11.3%) Brown green (139C) 1 (2.3%) Yellow brown (N167A) 1 (2.3%) Yellow brown (168D) 1 (2.3%)	Smooth 25 (55%) Lenticelate with hoop marks 20 (50%)	Dark green (136A) 1 (2%) Dark green (137A) 34 (76%) Dark green (141A) 6 (13%) Brown green (137C) 4 (9%)	Obovate 22 (49%) Cunate 7 (15.5%) Linear 3 (7%) Elliptic 11 (24.5%) Triangular 2 (4%)	Absent 31 (70.5%) Present 13 (29.5%)	Cylindrical 6 (14.5%) Straight twisted 3 (7.5%) Spirally twisted 8 (19.5%) Cylindrical and straight twisted 11 (27%) Cylindrical and spirally twisted 1 (2.5%) Straight twisted and spirally twisted 7 (17%) Cylindrical, straight twisted and spirally twisted 5 (12%)	Elliptical 32 (82%) Tetragonal 7 (18%)	Grey brown (N199A) 1 (3.5%) Brown purple (187A) 1 (3.5%) Brown purple (N186A) 3 (10.5%) Black (N187A) 1 (3.5%) Black (202A) 23 (79%)

Table 12. Morphological desription of cultivated *I. edulis* populationions in Ucayali an Loreto – qualitative characteristics. Number of samples (N).

Pop.	Trunk shape	Bark colour (RHS code)	Bark texture	Mature leaf colour	Rhachis wing shape	Rhachis wing deformation	Legume shape	Seed shape	Seed colour
I. edulis cultivated **Ucayali**	N 70	N 63	N 59	N 60	N 64	N 66	N 59	N 61	N 55
	Cylindrical 51 (73%)	Grey brown (199A) 24 (38%)	Smooth 30 (51%)	Dark green (136A) 8 (13%)	Obovate 38 (60%)	Absent 57 (86%)	Cylindrical 34 (57.5%)	Elliptical 49 (80%)	Brown purple (N77A) 33 (60%)
	Buttresed 19 (27%)	Grey brown (N199A) 6 (9.5%)	Lenticelate with hoop marks 29 (49%)	Dark green (137A) 37 (62%)	Cunate 9 (14%)	Present 9 (14%)	Straight twisted 4 (7%)	Tetragonal 10 (16%)	Black (202A) 22 (40%)
		Grey brown (199C) 32 (51%)		Dark green (141A) 14 (23%)	Linear 2 (3%)		Spirally twisted 4 (7%)	Triangular 2 (4%)	
		Brown green (139C) 1 (1.5%)		Brown green (137C) 1 (2%)	Elliptic 11 (17%)		Cylindrical and straight twisted 9 (15%)		
					Triangular 2 (3%)		Cylindrical and spirally twisted 2 (3.5%)		
					Obcordate 2 (3%)		Straight twisted and spirally twisted 1 (1.5%)		
							Cylindrical, straight twisted and spirally twisted 5 (8.5%)		
I. edulis cultivated **Loreto**	N 62	N 63	N 59	N 61	N 64	N 65	N 42	N 10	N 8
	Cylindrical 55 (**89%**)	Grey brown (199A) 30 (48%)	Smooth 23 (39%)	Dark green (136A) 10 (16%)	Obovate 41 (65%)	Absent 23 (35%)	Cylindrical 22 (52%)	Elliptical 8 (80%)	White (N155D) 1 (12.5%)
	Buttresed 7 (11%)	Grey brown (N199A) 5 (8%)	Lenticelate with hoop marks 36 (61%)	Dark green (137A) 35 (58%)	Cunate 13 (20%)	Present 42 (65%)	Straight twisted 1 (2.5%)	Tetragonal 1 (10%)	Brown purple (N77A) 1 (12.5%)
		Grey brown (199C) 26 (41%)		Dark green (141A) 16 (26%)	Linear 2 (3%)		Spirally twisted 10 (24%)	Triangular 1 (10%)	Black (202A) 6 (75%)
		Grey brown (N199C) 2 (3%)			Elliptic 6 (9%)		Cylindrical and straight twisted 4 (9.5%)		
					Triangular 2 (3%)		Cylindrical and spirally twisted 5 (12%)		

6.2.1 Comparison of cultivated and wild *I. edulis* tree legume lengths

From a total of 259 individual trees, 448 legumes were collected and measured: 329 and 119 legumes from 197 cultivated trees and 62 wild trees, respectively. The cultivated populations originated in Selva Central (94 mature legumes collected in 45 trees from 5 populations), Ucayali (120 mature legumes collected in 72 trees from 7 populations) and Loreto (115 mature legumes collected in 80 trees from 10 populations) regions and wild populations (119 mature legumes collected in 62 trees from 5 populations). The longest legume (148 cm) was found in 18 EDc, a cultivated population in the Loreto region El Dorado village (Table 13).

Table 13. Average legume length per population (minimal and maximal legume lengths are in brackets), and number of legumes (N_l) per population in a total of 448 legumes.

I. edulis	Population	Average legume length (cm)	N_l
cultivated Selva Central	1 SRc	78 (27-117)	35
	2 VRc	77 (58-100)	8
	3 PIc	90 (65-127)	18
	4 SAc	71 (38-104)	14
	5 SMc	70 (40-116)	19
cultivated Ucayali	6 ATc	81 (43-132)	11
	7 VHc	69 (28-124)	16
	8 CTc	85 (61-114)	35
	9 CVc	75 (44-93)	18
	10 ARc	72 (33-99)	22
	11 YAc	83 (59-118)	10
	12 SSc	84 (56-111)	8
cultivated Loreto	13 BRc	81 (13-118)	16
	14 JHc	74 (62-89)	6
	15 LAc	84 (54-138)	14
	16 NAc	101 (58-133)	13
	17 EPc	107 (78-131)	13
	18 EDc	93 (64-148)	14
	19 MAc	79 (57-104)	8
	20 SCc	80 (51-103)	9
	21 INc	99 (75-126)	8
	22 MZc	94 (72-118)	14
wild	23 RPw	31 (20-45)	32
	24 RSw	42 (31-58.5)	15
	25 RUw	43 (24-62)	35
	26 MAw	42 (27-72.5)	32
	27 SDw	40 (32.5-50)	5

For the the legume length description see Tables 7, 8 and 9.

The Kruskal-Wallis non-parametric test, using the four groups of populations (regions), showed significant among groups' differences in legume length (Fig. 26). The Mann-Whitney U post-hoc test produced three homogeneous groups, which indicated that legume length average (78 cm) in the Selva Central region was not significantly different ($P<0.05$) from the Ucayali region's 80 cm long average, and, both values, were significantly different from the Loreto's value (90 cm) (Figure 26). The Selva Central and Ucayali regions could be one group, considering the cultivated trees average legume length. The Loreto region cultivated trees produced the highest legume length average. The average legume length 83 ± 1.17 cm (mean ± standard error SE) in cultivated trees was significantly higher than 39 ± 0.95 cm legume length average in wild trees.

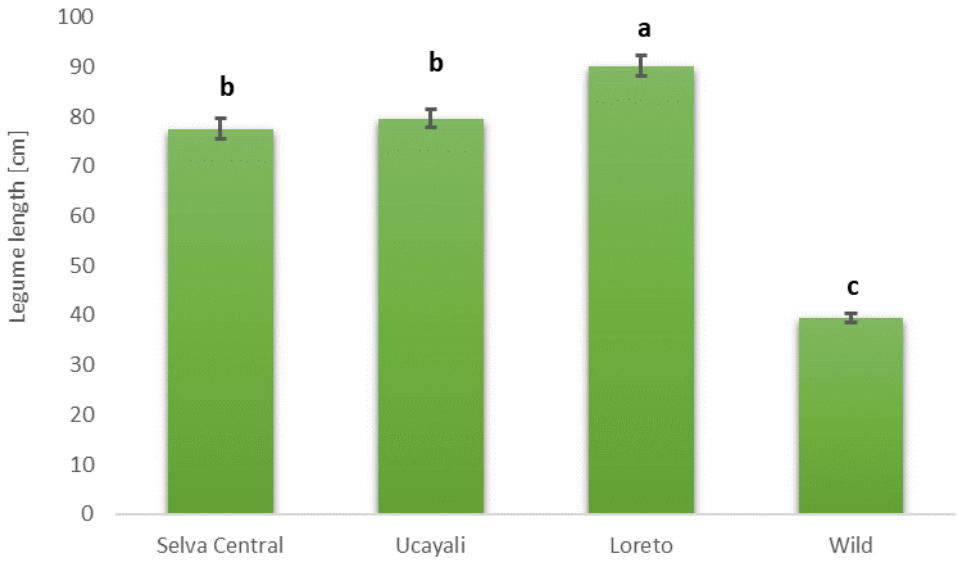

Figure 26. Legume length comparison among cultivated populations originated in Selva Central, Ucayali and Loreto regions and wild populations. Significantly different means showed with SE are followed by different letters ($p< 0.05$).

6.3 *INGA EDULIS* GENETIC DIVERSITY

It was identified a total of 71 alleles using the four microsatellite markers after genotyping all the individuals from the 27 populations. The average A was 5.7, the R_S 4.4, the H_O was 0.59 ± 0.03, and H_E was 0.69 ± 0.02. The overall inbreeding coefficient F_{IS} was 0.11 ± 0.03 (Table 14).

Table 14. Summary of genetic diversity of the *I. edulis* 27 populations. Sample size (N), average number of alleles per locus (A), allelic richness (Rs), effective number of alleles (N_e), expected heterozygosity (H_E), observed heterozygosity (H_O) and fixation index (F_{IS}) averaged overloci. Sig. refers to the significance resulting from the heterozygote deficiency test (a conservative α value for the test of at least P<0.01 was used, due to the low number of individuals per population: NS, not significant, ** P<0.01 and ***, significant P<0.001). Standard errors in brackets.

I. edulis	Population	N	A	R_S	N_e	H_O	H_E	F_{IS}	*Sig*
cultivated **Selva Central**	1 SRc	10	5.5	3.9	2.91 (0.42)	0.53 (0.13)	0.67 (0.06)	0.18 (0.16)	NS
	2 VRc	5	5.5	5.5	4.09 (0.52)	0.70 (0.17)	0.83 (0.03)	0.06 (0.22)	NS
	3 PIc	10	5.0	4.0	3.32 (0.91)	0.55 (0.13)	0.67 (0.08)	0.15 (0.16)	NS
	4 SAc	10	5.8	4.1	2.96 (0.72)	0.58 (0.15)	0.63 (0.09)	0.06 (0.19)	NS
	5 SMc	10	6.3	4.4	3.39 (0.84)	0.60 (0.12)	0.67 (0.10)	0.07 (0.08)	NS
cultivated **Ucayali**	6 ATc	10	4.8	3.7	2.72 (0.51)	0.53 (0.18)	0.61 (0.09)	0.16 (0.19)	NS
	7 VHc	8	5.3	4.5	3.86 (0.55)	0.63 (0.09)	0.77 (0.04)	0.15 (0.09)	**
	8 CTc	18	7.3	4.6	4.17 (0.78)	0.65 (0.08)	0.76 (0.05)	0.12 (0.07)	**
	9 CVc	12	5.8	4.5	4.15 (0.78)	0.67 (0.12)	0.76 (0.05)	0.11 (0.12)	NS
	10 ARc	11	5.5	4.3	3.48 (0.21)	0.61 (0.15)	0.74 (0.02)	0.14 (0.21)	***
	11 YAc	5	5.0	5.0	3.22 (0.77)	0.65 (0.15)	0.66 (0.16)	-0.10 (0.03)	NS
	12 SSc	8	5.5	4.6	4.03 (0.99)	0.50 (0.21)	0.75 (0.07)	0.36 (0.25)	***
cultivated **Loreto**	13 BRc	16	8.0	4.8	4.52 (1.45)	0.64 (0.12)	0.71 (0.10)	0.08 (0.09)	NS
	14 JHc	5	3.8	3.8	2.55 (0.26)	0.50 (0.19)	0.66 (0.05)	0.20 (0.31)	NS
	15 LAc	10	6.8	4.8	4.42 (1.26)	0.65 (0.12)	0.73 (0.11)	0.03 (0.14)	NS
	16 NAc	5	4.3	4.3	2.93 (0.64)	0.60 (0.18)	0.64 (0.15)	-0.04 (0.16)	NS
	17 EPc	5	3.8	3.8	2.79 (0.31)	0.55 (0.19)	0.69 (0.06)	0.18 (0.29)	NS
	18 EDc	12	5.0	3.8	3.03 (0.68)	0.46 (0.17)	0.62 (0.12)	0.33 (0.26)	**
	19 MAc	5	3.8	3.8	2.59 (0.61)	0.50 (0.13)	0.61 (0.11)	0.10 (0.15)	NS
	20 SCc	5	3.8	3.8	2.58 (0.33)	0.45 (0.13)	0.66 (0.07)	0.25 (0.20)	NS
	21 INc	7	3.8	3.3	2.60 (0.84)	0.43 (0.21)	0.50 (0.17)	0.20 (0.27)	NS
	22 MZc	10	5.5	4.3	3.76 (0.72)	0.60 (0.16)	0.73 (0.07)	0.18 (0.18)	**
MEAN (cultivated)		**197**	**5.3**	**4.2**	**3.37 (0.16)**	**0.57 (0.03)**	**0.69 (0.02)**	**0.14 (0.04)**	
wild	23 RPw	12	8.3	5.2	5.06 (1.17)	0.63 (0.17)	0.72 (0.13)	0.09 (0.18)	NS
	24 RSw	6	6.5	5.8	5.32 (1.37)	0.75 (0.08)	0.79 (0.13)	-0.08 (0.09)	NS
	25 RUw	12	7.3	5.2	4.58 (1.15)	0.67 (0.14)	0.76 (0.07)	0.11 (0.17)	NS
	26 MAw	27	11.0	5.4	5.98 (1.99)	0.66 (0.16)	0.75 (0.12)	0.12 (0.10)	NS
	27 SDw	5	4.0	4.0	2.77 (0.94)	0.70 (0.13)	0.60 (0.11)	-0.30 (0.07)	NS
MEAN (wild)		**62**	**7.4**	**5.1**	**4.74 (0.64)**	**0.68 (0.06)**	**0.72 (0.05)**	**-0.01 (0.06)**	
MEAN (cultivated **and** wild)		**259**	**5.7**	**4.4**	**3.62 (0.18)**	**0.59 (0.03)**	**0.69 (0.02)**	**0.11 (0.03)**	

From the results of the current study, the population with the highest and lowest expected heterozygosity possessed the highest and the lowest allelic richness values, in both cultivated and wild populations. The allelic richness parameter correlated well with the populations' genetic diversity parameters, which is not surprising since the number of individuals sampled per population was unevenly distributed. The population with the highest H_E was 2 VRc, 0.83, a cultivated population from Selva Central region, and the lowest was 21 INc (0.50), a cultivated population from the Loreto region, and they both possessed the highest and the lowest R_S 5.5 and 3.3, respectively. The wild populations 24 RSw and 27 SDw displayed the highest and lowest H_E values, 0.79 and 0.60, similarly with the highest and lowest R_S values, 5.8 and 4.0. Interestingly, the population with the highest A and N_e was 26 MAw, which could be partially explained by the highest sampled individuals number (27). The cultivated populations 8 CTc and 13 BRc also had a high number of sampled individuals, which was also reflected in the A and N_e parameters. The average expected genetic diversity is slightly higher in the wild compared to the cultivated populations, 0.72 and 0.69, respectively. The average number of alleles is much higher in the wild (7.4) than in the group of cultivated populations (5.3), but when we consider the allelic richness and effective number of alleles, those differences smoothed down (Table 14).

Six cultivated populations out of 22 had significant heterozygote deficiency, but this parameter was not significant in the wild populations (Table 14). The Selva Central group of cultivated populations lacked populations with significant inbreeding coefficient, and the Loreto and the Ucayalli regions had only two populations and more than half of the populations with heterozygote deficiency, respectively. Positive and significant F_{IS} values mirror differences between observed and expected heterozygosity, due to putative heterozygosity loss because of non-random mating of parents. Nevertheless, when we compared the overall inbreeding coefficient from the wild with the cultivated populations, no significant differences were found between them. Conversely, the allelic richness and the observed heterozygosity were significantly lower in the cultivated populations. The genetic differentiation (F_{ST}) was significantly higher in the cultivated than in the wild populations (Table 15).

Table 15 Diversity parameters comparison between cultivated and wild populations. Allelic richness (R_S), observed heterozygosity (H_O), inbreeding coefficient (F_{IS}) and among populations differentiation (F_{ST}). Probability values (P) for differences between groups for two-sided t-test after 1000 permutations. * = significant test (P < 0.05).

Diversity measures	Cultivated	Wild	P
R_S	4.24*	5.12*	0.006
H_O	0.58*	0.67*	0.031
F_{IS}	0.14	-0.01	0.050
F_{ST}	0.076*	0.017*	0.033

The private alleles (P_a) were identified and compared in *I. edulis* wild and cultivated populations. Seven private alleles were identified in three wild populations, the highest P_a per population was found in the 26 MAw population (3) and two in both 23 RPw and 25 RUw. Only one private allele was identified in four different *I. edulis* cultivated populations (2 VRc, 9 CVc, 14 JHc and 22 MZc). The locus *Inga*08 had the highest P_a (7 across all populations) and *Inga*33 and *Pel*5 only one (data not shown).

The cultivated populations possessed 13 exclusive alleles compared to the wild ones, and only two had a frequency lower than 5%. The regions with the higher number of cultivated populations with exclusive alleles was Selva Central, 80 %, followed by Ucayali, 60 %, and the Loreto region had the lower number of populations with those alleles (40 %) (data not shown).

6.4 *INGA EDULIS* POPULATION STRUCTURE

The results show low genetic structure between wild and cultivated stands. The population genetic structure was investigated by a hierarchical analysis of molecular variance (AMOVA), which revealed that most of the genetic diversity existed within populations (92%). The differentiation between cultivated and wild groups of populations (Φ_{CT} = 0.010) was not significant (P < 0.0958), and the variation among populations within species was appreciable, Φ_{SC} = 0.073 and significant (P < 0.0001) (Table 16).

Table 16. Hierarchical AMOVA between cultivated and wild population groups, among populations within cultivated and wild population groups and within *I. edulis* populations. Degrees of freedom (df), sum of squared deviation (SS), fixation indexes (Φ statistics), level of probability of obtaining a more extreme component estimate by chance alone (P). Significance of variance components were tested by a permutation test.

Source of variation	d.f.	SS	Variance componets	% of variation	Φ statistics	P
Between groups (cultivated *vs.* wild)	1	7.697	0.01486	0.97	Φ_{CT} = 0.010	<0.0958
Among populations within groups	25	86.53	0.11103	7.24	Φ_{SC} = 0.073	<0.0001
Within populations	491	690.914	1.40716	91.79	Φ_{ST} = 0.082	<0.0001
Total	517	785.141	1.53304			

The *I. edulis* genetic structure was further estimated using a Bayesian approach. Using the method of Evanno et al. (2005) the most appropriate number of genetic clusters (K) is 2 further referred to as red and green (Figure 27, 28 and 29). The red cluster was predominant in the wild populations and in the cultivated populations in the northernmost region (Loreto). Conversely, the green cluster was predominant in the southernmost region (Selva Central). The Ucayali region displayed a mixture of both types of cultivated populations, probably a mixture from the southern and the northern regions (Figure 28 and 29). For K = 2, the highest proportion of red cluster was observed in cultivated populations along the navigable river watersheds in

Loreto and Ucayali regions (e.g. 6 ATc, 11 YAc, 14 JHc, 15 LAc, 16 NAc, 19 MAc, 20 SCc, 21 INc and 22 MZc). Moreover, the green cluster was found to be prevalent in populations cultivated on the Andean foothills and 'terra firme' in Selva Central and Ucayali regions (e.g. 2 VRc, 3 PIc, 4 SAc, 7 VHc and 10 ARc) (Figure 29).

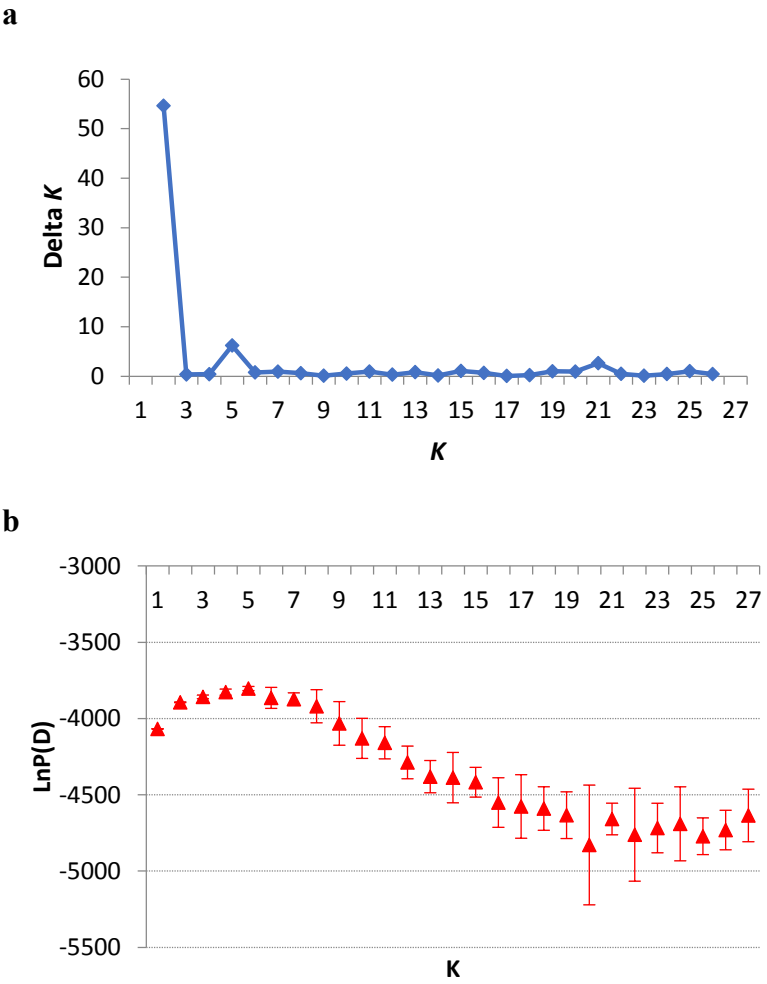

Figure 27. ΔK (a) and log-likelihood probability (LnP(D)) ± SD (b), after using the Evanno et al. (2005) to determine the number of inferred clusters (K).

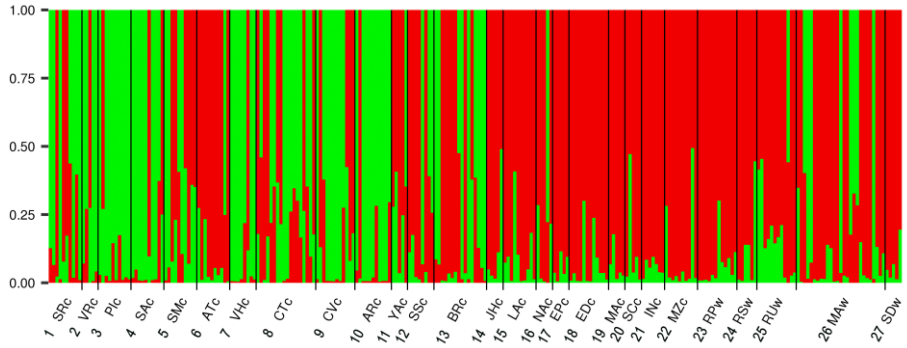

Figure 28. Proportion of genotype membership q (y-axis) based on the STRUCTURE cluster analysis. Plots of proportional group membership for the 259 trees for $K = 2$. Each tree is represented by a single vertical line, which is divided in different colors based on the genotype affinities to each K cluster (red and green). Divisions between populations are made with black lines.

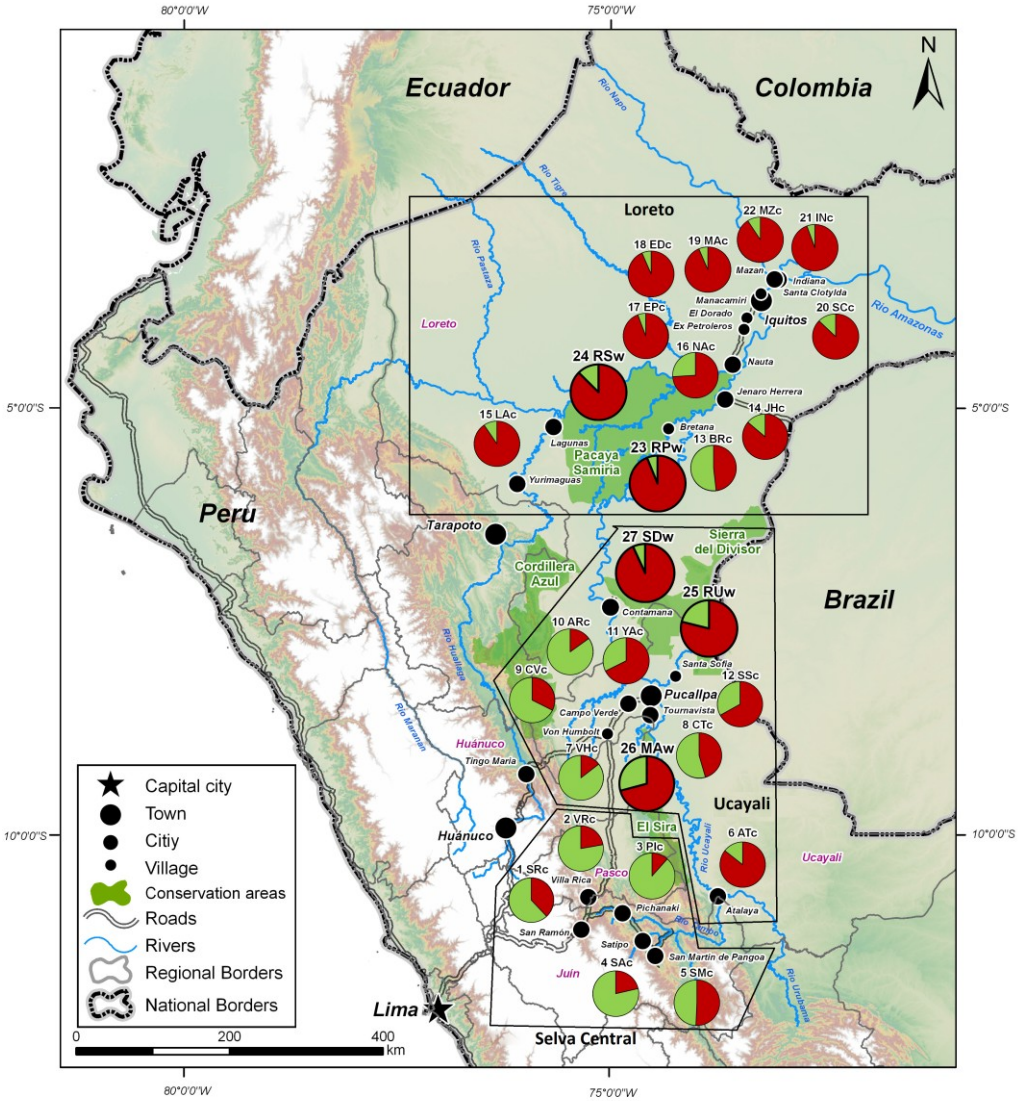

Figure 29. *Inga edulis* populations investigated in this study plotted in the map of Peru. Bayesian clustering for $K = 2$. Populations assigned to two clusters (red and green) corresponding to the *I. edulis* wild (bigger black-lined pie charts) and cultivated populations (smaller no-lined pie charts).

6.5 *INGA INGOIDES* AND *I. EDULIS* COMPARATION OF GENETIC DIVERSITY AND INBREEDING

The four SSR loci were polymorphic, with a total of 66 alleles in *I. ingoides* and 58 alleles in *I. edulis*. However, the higher number of alleles (N_a) could reflect the higher number of individuals (N) in some of the populations in both species: RPI (N=47; N_a=13.3) and MAw populations (N=27; N_a=11). The effective number of alleles (N_e) was higher in the *I. ingoides* southern population RUI (6.1), and lower in the northern one RSI (4.4). The *I. edulis* western population (MAw) held the highest N_e value (6), and the smallest value was found in the eastern SDw population (2.8). The rarefaction method displayed similar average allelic richness (A_R) values in both species (5.1), due to differences in sample size per population.

The expected heterozygosity (H_E) was also similar in both species (ca. 0.70), but the observed diversity (H_O) was lower for *I. ingoides* (0.54) compared with *I. edulis* (0.68), which leads to a positive inbreeding coefficient (F_{IS}) in the former. All the *I. edulis* populations are in HWE, but not the *I. ingoides* ones. High F_{IS} values, the loss of heterozygosity due to non-random mating of parents, reflected differences between observed and expected heterozygosity. *I. ingoides* populations (RPI, RSI and RUI) departures from HWE showed significant (P<0.001) heterozygote deficiency. On the contrary, the *I. edulis* populations F_{IS} values were not significant. The average frequency of null alleles was similar and low in both species. In addition, no LD was detected between different genotypes with the Fisher exact test among the different loci (P>0.05), indicating that all 4 loci segregate independently of each other in both studied species (Table 17.).

Table 17. Diversity parameters per population obtained with the 4 SSR polymorphic loci after genotyping the *I. ingoides* and *I. edulis* individuals. Sample size *(N)*, number of alleles per locus (N_a), effective number of alleles (N_e), allelic richness (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O) and fixation index (F_{IS}). Sig. refers to the significance resulting from the HWE test (after Bonferroni correction, NS, not significant and ***, significant P<0.001). Average estimate of null frequency (F-null). Standard errors in brackets.

Species	Pop.	N	N_a	A_R	N_e	H_O	H_E	F_{IS}	Sig.	*F-null*
I. ingoides	RPI	47	13.3	5.2	5.82 (1.61)	0.58 (0.14)	0.72 (0.15)	0.14 (0.15)	***	0.08
	RSI	16	7.5	4.5	4.39 (1.34)	0.47 (0.19)	0.66 (0.16)	0.27 (0.18)	***	0.10
	RUI	14	9.8	5.6	6.06 (1.94)	0.58 (0.13)	0.73 (0.16)	0.14 (0.11)	***	0.09
	Mean	**77***	**10.2**	**5.12**	**5.42 (1.63)**	**0.54 (0.16)**	**0.70 (0.16)**	**0.18 (0.15)**		0.09
I. edulis	RPw	12	8.3	5.2	5.06 (1.17)	0.63 (0.17)	0.72 (0.13)	0.09 (0.18)	NS	0.06
	RSw	6	6.5	5.8	5.32 (1.37)	0.75 (0.08)	0.79 (0.13)	-0.08 (0.09)	NS	0.00
	RUw	12	7.3	5.2	4.58 (1.15)	0.67 (0.14)	0.76 (0.07)	0.11 (0.17)	NS	0.06
	MAw	27	11.0	5.4	5.98 (1.99)	0.66 (0.16)	0.75 (0.12)	0.12 (0.10)	NS	0.06
	SDw	5	4.0	4.0	2.77 (0.94)	0.70 (0.13)	0.60 (0.11)	-0.30 (0.07)	NS	0.00
	Mean	**62***	**7.4**	**5.1**	**4.74 (0.64)**	**0.68 (0.06)**	**0.72 (0.05)**	**-0.01 (0.06)**		0.06

*sum

The loci with higher N_a (18) were different in both species: *Pel*5 in *I. edulis* and *Inga*03 and *Inga*33 in *I. ingoides*. The A_R per loci ranged from 4.2 (*Inga*08) to 11.5 (*Inga*33) based on the minimum sample size of 14 individuals in *I. ingoides* and from 3.3 (*Inga*08) to 7.1 (*Pel*5) based on the minimum sample size of 5 individuals in *I. edulis*. The *Inga*08 locus had the lowest H_e values in both species (0.24 and 0.47, in *I. ingoides* and *I. edulis*, respectively), and the *Pel*5 locus had the highest value (ca. 0.90) (Table 18.).

Table 18. Diversity parameters per locus obtained with the 4 SSR polymorphic loci after genotyping the *I. ingoides* and *I. edulis* individuals. Number of alleles per locus (N_a), effective number of alleles (N_e), allelic richness (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O) and fixation index (F_{IS}).

Species	Loci	N_a	A_R	N_e	H_O	H_E	F_{IS}
I. ingoides	*Inga*03	18	8.6	5.31 (1.20)	0.63 (0.09)	0.81 (0.06)	0.21 (0.10)
	*Inga*08	13	4.2	1.31 (0.05)	0.24 (0.06)	0.24 (0.03)	0.03 (0.13)
	*Inga*33	18	11.5	6.60 (0.93)	0.39 (0.08)	0.87 (0.02)	0.54 (0.11)
	*Pel*5	17	11.3	8.47 (1.13)	0.92 (0.05)	0.90 (0.02)	-0.05 (0.07)
	Mean	**17**	**8.9**	**4.77 (0.79)**	**0.48 (0.08)**	**0.67 (0.07)**	**0.26 (0.08)**
I. edulis	*Inga*03	16	6.3	5.56 (0.91)	0.86 (0.03)	0.83 (0.06)	-0.13 (0.09)
	*Inga*08	11	3.3	1.86 (0.21)	0.51 (0.08)	0.47 (0.05)	-0.15 (0.10)
	*Inga*33	13	4.9	3.58 (0.88)	0.46 (0.11)	0.68 (0.09)	0.28 (0.16)
	*Pel*5	18	7.1	7.97 (0.98)	0.90 (0.03)	0.92 (0.01)	-0.04 (0.05)
	Mean	**16**	**5.4**	**4.74 (0.64)**	**0.68 (0.56)**	**0.72 (0.05)**	**-0.01 (0.06)**

The private alleles were identified and compared in *I. edulis* wild and *I. ingoides* populations. For each *I. ingoides* population, the highest P_a per population was found in the RPI population (3.5 across loci) and the lowest value in the RSI (0.75). The locus *Inga*03 had the highest P_a (2.7 across all populations) and *Inga*33 had the lowest (1.33) in this species. Private alleles were identified in four *I. edulis* populations, the RPE had the highest P_a (1.25 across loci). The SDw population had no private allele, probably due to the low sample size. Only two alleles are common to the RPI/RPw pair, in the other pairs there are no common private alleles. The populations RUI and RSE hold the highest N/NP_a ratio, i.e., they have the highest number of private alleles compared to the population size (Table 19.).

Table 19. Number of private alleles (P_a) per population and locus of *I. edulis* and *I. ingoides* and number of private alleles per population (NP_a=sum of P_a)- N/NP_a represent the ratio of the number of individuals from a population (N) by the number of private alleles of that population.

Species	**Pop**	**Inga03**	**Inga08**	**Inga33**	**Pel5**	**NP_a**	**N/NP_a**	**Mean/pop**
I. ingoides	RPI	4	4	3	3	14	0.3	3.5
	RSI	1	1	0	1	3	0.2	0.8
	RUI	3	2	1	2	8	0.6	2.0
	Mean/loci	**2.67**	**2.33**	**1.33**	**2**	**8**	**0.4**	**3.3**
I. edulis	RPw	3	1	1	0	5	0.4	1.3
	RSw	0	0	2	2	4	0.7	1.0
	RUw	0	3	1	0	4	0.3	1.0
	MAw	4	2	1	1	8	0.3	2.0
	SDw	0	0	0	0	0	0.0	0.0
	Mean/loci	**1.4**	**1.2**	**1**	**0.6**	**4**	**0.2**	**1.1**

6.5.1 *Inga edulis* and *I. ingoides* population differentiation and structure

The PCoA analysis reveals populations' weak grouping, with the first and the second factor explaining 68 % and 15 % of the total variation, respectively. The AMOVA revealed an overall low among population variation (F_{ST} = 0.05; P < 0.0001), and the highest variation of the data set was found within populations (94 %). Undoubtedly, the group (a) including all the *I. edulis* populations clustered separately from group (b) the three *I. ingoides* populations. Furthermore, the analysis of molecular variance (AMOVA) confirmed a low, yet significant (P < 0.02) differentiation between the two *Inga* species F_{CT} = 0.036 (Figure 30.).

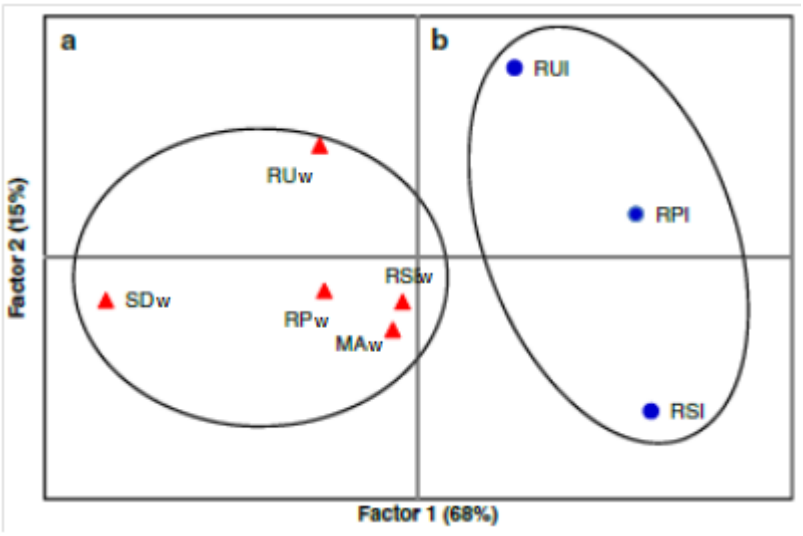

Figure 30. Principal coordinates analysis based on the Nei's pairwise genetic distances of *I. edulis* (filled triangles) and of *I. ingoides* populations (filled circles). Group (a) and group (b), included populations from both species along the Pacaya, Samiria and Utiquinia rivers, respectively. The population SDw is an outlier.

The *I. ingoides* populations at the three different rivers were clearly separated, as observed in Figure 30, widely separated along the second axis, although only explaining a small part of the variation. Indeed, the variation among populations within species was weak, $F_{SC} = 0.027$ (Table 20).

Table 20. AMOVA of the *Inga* populations, considering the whole data set and clustered in the two species (*I. edulis* and *I. ingoides*) according to the PCoA analysis (Figure 30). Sum of squared deviation (SS), degrees of freedom (df), level of probability of obtaining a more extreme component estimate by chance alone (P).

Source of variation	df	SS	Variance components	% of total variance	*F* statistics	*P*
All populations						
Among populations	7	25.996	0.07204	4.87	F_{ST} =0.05	<0.0001
Within populations	270	379.763	1.40653	95.13		
Total	277	405.759	1.47856			
I. edulis* vs. *I. ingoides						
Between species	1	10.84	0.05	3.64	$F_{CT} = 0.036$	<0.02
Among populations within species	6	15.15	0.04	2.57	$F_{SC} = 0.027$	<0.0001
Within populations	270	379.76	1.41	93.79	$F_{ST} = 0.062$	<0.0001
Total	277	405.76	1.50			

The STRUCTURE distinguished clusters and the mean likelihood indicated two peaks at $K = 2$ and $K = 4$. Methods for estimating the most appropriate K testing $K = 2$ to 8 for 139 individuals from 8 populations of the two species *I. edulis* and *I. ingoides* are shown (Figure 31, 32 and 33). Using the delta K criterion, the Bayesian clustering suggests the most probable presence of four groups (Figure 32.), yet all individuals with mixed ancestry. Thus, the genetic clusters uncover extensive gene flow among populations. The mixed ancestry was particularly evident in the close population pairs along the rivers, with the more isolated *I. edulis* MAw and SDw populations clearly less mixed (Figure 34a and 34b.).

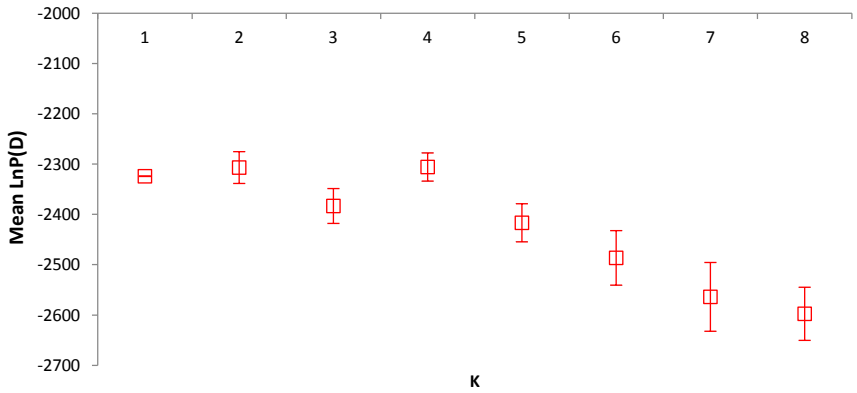

Figure 31. The mean Log-likelihood of K LnP(D) ± SD given K clusters, obtained through 10 runs with the STRUCTRURE algorithm, showing peaks at $K = 2$ and $K = 4$.

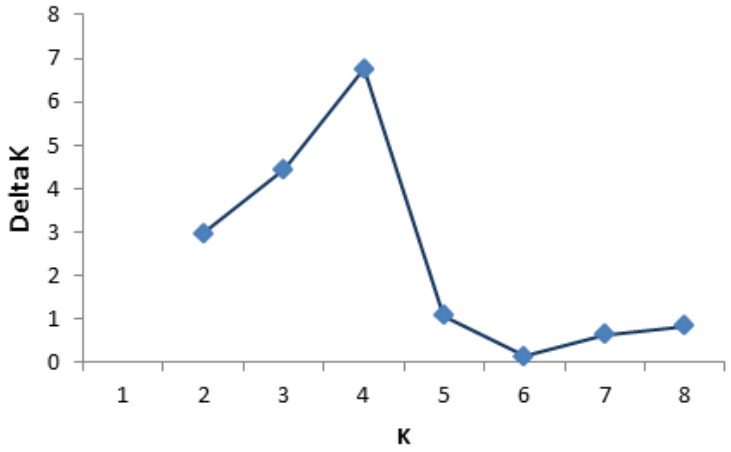

Figure 32. Delta statistics showing a clear peak for $K = 4$, indicating that this is the most appropriate number of genetic clusters.

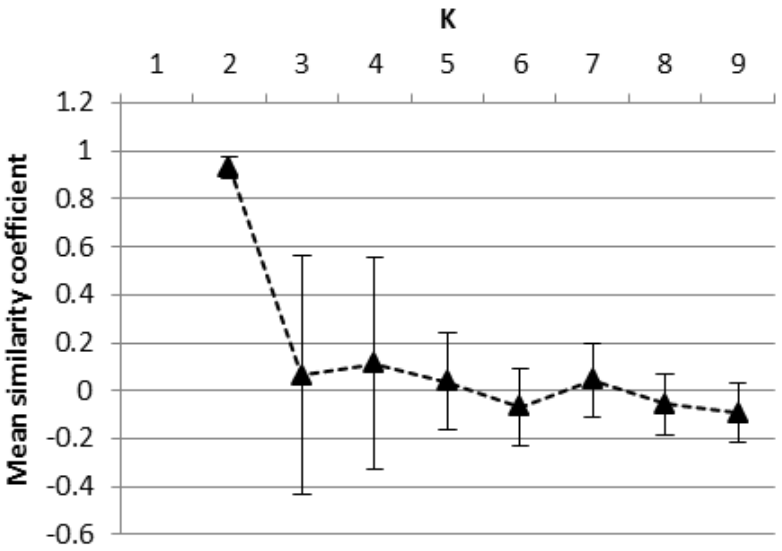

Figure 33. Similarity coefficient indicates the similarity between 10 runs ± SD. For $K = 2$ the similarity is higher than 0.9, meaning that each run ended with a similar result.

Additionally, it was found that the mean similarity coefficient, the similarity between the 10 runs, was consistently higher for $K = 2$ (Figure 33.). Considering $K = 2$, the clusters corresponded to the two species groups, which had a biologically meaningful result: a clear introgression between species (Figure 34a.).

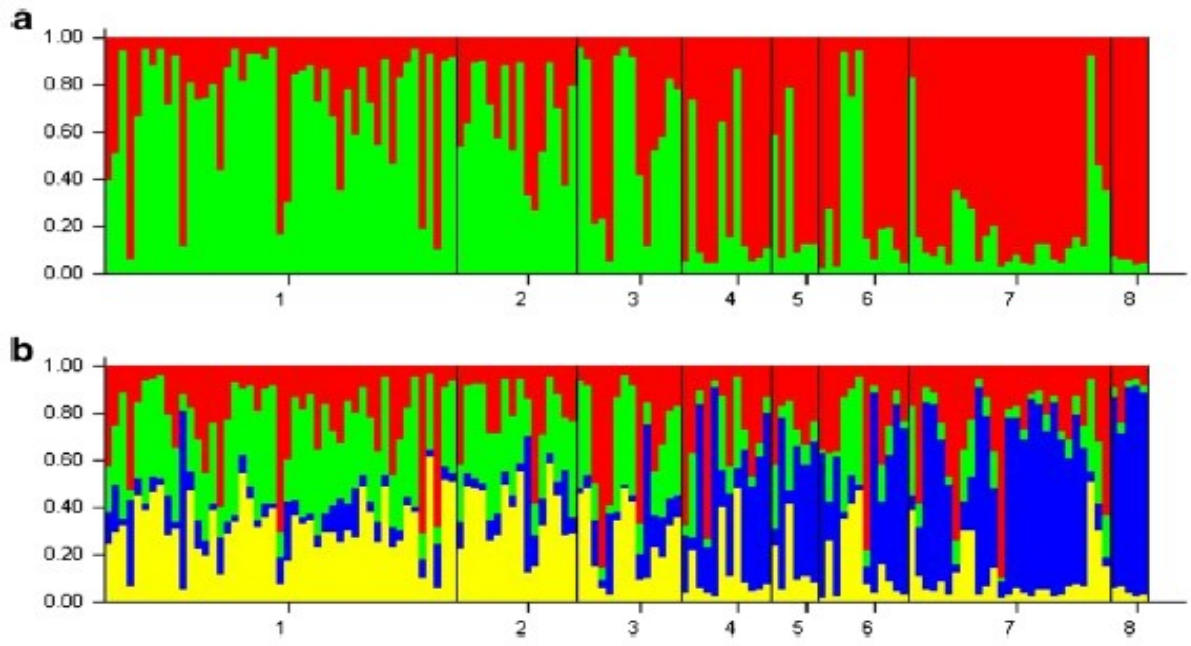

Figure 34. (a) Proportion of *I. edulis* and *I. ingoides* genotype membership q (y-axis) based on Bayesian cluster analysis. Each individual is represented by a single vertical line that is partitioned in different colors based on its genotype affinities to each cluster (K). Grey lines indicate the division between populations. Populations: 1 (RPI), 2 (RSI), 3 (RUI), 4 (RPw), 5 (RSw), 6 (RUw), 7 (MAw), 8 (SDw). (b) Plots of proportional group membership for the 139 trees for K = 4, yellow cluster 1, blue cluster 2, green cluster 3, red cluster 4.

The RUI/RUw populations seem to be the most mixed pair. The genetic clusters did not closely correspond to the morphological species, which suggested that gene flow has occurred between the species. The three *I. ingoides* populations seem to have the highest proportion of genotype affinities (or proportion of genotype membership) to both cluster 1 and 3, whereas *I. edulis* predominant proportion of genotype membership arised from cluster 2, in particular for the MAw and SDw populations (Figure 34b). For $K = 2$, the mean introgression was higher for *I. ingoides* (25 %) than for *I. edulis* (18 %), considering the number of individuals with more than 50% probability as belonging to the other species ($q > 50$ %), however the species introgression appears to be bidirectional. Nevertheless, if only the populations along the rivers are considered (RPw, RSw and RUw) the average introgression sums up to 28 % in *I. edulis* and the MAw and SDw populations have negligible values. The RUI population has the highest introgression degree (36 %), almost twice the other *I. ingoides* populations (Figure 34a).

7 DISCUSSION

7.1 *INGA EDULIS* MORPHOLOGICAL VARIABILITY

The morphological data obtained after evaluation following the designed characterization descriptors were shown to characterize the studied *I. edulis* population. These data do not correspond to the possibilities of their use in this study to determine specific conclusions and instead of the legume length they are not further discussed. Some parameters were not evaluated because of the scope of the thesis or due to unsufficient sample size. Despite their limited amount, caused by small sample size, or quality caused by different age, development and phenological stage of the studied trees in the moment of evaluation, these data represent the sampled *I. edulis* population morphological variability. There is neither macro- nor micro-phenological scale available at the moment for the *I. edulis* species. Due to the limited time and difficulty of fieldwork during sampling, it was not always possible to obtain material in the same stage of ripening. The number of measured legumes was reduced to at least one fruit per sampled tree and the range of fruit phenological stages has been extended for fruit rippening and senescance. The following range of legume phenological stages in which the legume could have been measured was as following: Fruit ripening - legumes and seeds reached their final length, seeds are creamy white with membraneous sarcotesta; “maturity” - seeds are changing its colour into purple black, sarcotesta is flashy and according to Pennington and Fernandes (1998) watery, soft, slightly sweet, generally white tissue; “senescence” - seeds reached typical purple black colour, sometimes viviparic, sarcotesta rotted or being eaten by worms. This was also the reason why some of the mentioned features from the characterization descriptors have been observed only on the small sample size. Anyway, these data could open the discussion on phenotypic variance of *I. edulis* species in the Amazonian Peru and could be an interesting stimulus for further studies focused on this species.

7.2 INFLUENCE OF DOMESTICATION ON *I. EDULIS* LEGUME LENGTH

Although the history of cultivation of *I. edulis* is not well-documented, a crop domestication study suggested that humans have domesticated this species over a considerable period of time (Clement 1999b). Indeed, Amazonia is a major world centre of plant domestication, where selection began in the Late Pleistocene to Early Holocene in peripheral parts of the basin (Clement et al. 2015). The origin of cultivated *I. edulis* trees is uncertain, though probably

Amazonian (Pennington 1997), nevertheless some authors referred it was started by European settlers in west Amazonia (León 1987; Clement 1999b). Since this tree was cultivated mainly for fruit production, domestication is expected to increase legume length (Clement 1999b; Reynel and Pennington 1997; Pennington 1997). To the best of our knowledge no study was made comparing both types of population, cultivated vs. wild considering this morphologic characteristic (legume length). Certainly, the longer legumes were found in cultivated trees compared to the wild trees clearly support the *I. edulis* domestication for food supply. Plant domestication is a long-term process in which natural selection interacts with human selection driving changes that improve usefulness to humans and adaptations to domesticated landscapes (Clement et al. 2015).

In the current study, maximum legume length in the wild and the cultivated populations was 73 and 148 cm, respectively, in agreement with Pennington (1997). This author reported that wild trees legumes rarely exceed 50 cm and cultivated trees could, exceptionally, produce legumes exceeding 2 m. The average legume length was higher in the Loreto region's cultivated trees, compared to Ucayali and Selva Central regions, the smallest fruits were observed in Selva Central. The species' different cultivation and uses, and differences in ecological conditions, could explain these results. Indeed, in Selva Central, the species was mainly used to shade coffee or cocoa prevalent to produce large fruits (León 1998; Miller and Nair 2006). Farmers were focused mainly on cash crop yield, instead of fruit yield provided by shade trees. The soil nutrients should be allocated in the cash crop fruits instead of shading trees. Additionally, large fruits could be more attractive for uninvited guests, which could cause damage in the crop due to *Inga* fruit collection. Another supporting argument is that the wild *I. edulis* local name inexistence among the Selva Central region inhabitants (A. Rollo pers. communication). Locals informed me that *I. edulis* is hard to find in the surrounding wild vegetation, indeed the species is rarely seen above 750 m (Pennington, 1997). I was also unable to find and sample wild trees in the Selva Central region. The local name for the cultivated *I. edulis* in Selva Central is "pacay soga", the name "Guaba" is used for cultivated type and "guabilla" or "guabilla del monte" in Ucayali and Loreto regions (A. Rollo pers. communication). The local names diversity in those regions might be related to the species abundance, both in the wild and cultivated form.

7.3 GENETIC DIVERSITY OF WILD AND CULTIVATED POPULATIONS OF *I. EDULIS* IN PERUVIAN AMAZON

The overall H_E (0.69) was slightly higher than the H_O (0.59), inducing an overall inbreeding coefficient index of 11%. In a meta-study for microsatellites and outcrossing species, the author showed similar value of H_E (0.65), but slightly higher H_O (0.63) (Nybom 2004). The results in this thesis further indicate that all the genetic diversity estimates were lower in the case of the cultivated populations compared to the wild ones, as well as the average inbreeding coefficient. These results confirm a loss of genetic diversity in the cultivated populations, in agreement with the studies done by Hollingsworth et al. (2005) and Dawson et al. (2008) with the same species. Those authors concluded that cultivated stands possessed lower total allelic richness than neighbouring wild populations, but the expected genetic diversity remained unchanged, indicating that the process of domestication reduced the number of alleles. Both authors stated that the wild plant material they studied were collected from nearby cultivated populations, in old-growth, primary forest, but due to i) the long history of the species use, ii) the habits of slash-and-burn in primary forest, and iii) gene flow among nearby stands, the wildness of the trees could be at stake (Clement et al. 2015; Levis et al. 2018). Nevertheless, Dawson et al. (2008) observed marked differences in the haplotype composition between natural and cultivated stands. In this thesis the wild material was sampled in natural vegetation in protected areas and secondary forest, and unless extensive long-distance gene flow existed, no ambiguities to distinguish both types existed.

In addition, the results from the legume length clearly distinguish the wild from cultivated material. It was visibly found an effect of the domestication on the natural resources of a species, which is an expected phenomenon when a species is used by humans (e.g., Ribeiro et al. 2001; Cruz-Neto et al. 2014). In some cases, the expected heterozygosity might be higher or similar in the cultivated population than that displayed in the wild population, due to a putative 'melting pot' phenomenon in the former populations (introduced alleles from different origins). Nevertheless, the allelic richness and observed heterozygosity found, in this thesis, in the cultivated populations was lower than in the wild ones, indicating the loss of rare alleles during selection as observed by other authors (e. g., Cruz-Neto et al. 2014; Jones et al. 2006).

Some cultivated populations from the current study had significant heterozygote deficit, particularly in the Ucayali region, the consequences of inbreeding effect in fruit trees, such as *I. edulis*, might impact fruit production due to inbreeding depression, with direct impacts on farmers' yield (Koptur 1984; Cruz-Neto et al. 2014). The fact that the species is self-

incompatible (Dawson et al. 2008) excludes the possibility of heterozygote deficiency due to self-pollination, probably related trees were introduced in those populations and the value reflected biparental inbreeding.

7.4 *INGA EDULIS* POPULATION STRUCTURE

The genetic variance partition in our study (92% of the variance was observed within populations and a low genetic structure, 7%, was detected among populations) is usual in outcrossing tropical forest tree species with high levels of gene flow (Finkeldey and Hattemer 2007). The hierarchical AMOVA showed that the *F*ct between wild and cultivated populations was 1%, yet not significant.

Dawson et al. (2008) found low genetic structure similarly to the results of this thesis in *I. edulis* natural and cultivated stands, with nuclear, but not with chloroplast microsatellite data. Nevertheless, the authors used only two chloroplast loci, which might have biased the results, since the smaller effective population size of the chloroplast genome makes it more susceptible to genetic drift and species differentiation (Petit et al. 2005). Conversely, a high genetic structure was found between natural and cultivated stands of *I. vera*, and the authors reasoned that the cultivated populations were derived from seeds coming from different mother-trees, but a different geographic origin was also possible (Cruz-Neto et al. 2014).

In the current study, for $K = 2$ (Figure 28 and 29), the wild populations displayed identical composition, with predominant red cluster. The uniform composition of the studied wild material could be due to the genus relative recent speciation (Richardson et al. 2001) and, also, to regional wild populations sampling (Pennington 1997). The red cluster prevailed in the northern cultivated populations: the Loreto region and along the Ucayali river in the Ucayali region could express large population centres occupying the main rivers' margins with extensive trade networks (Miller and Nair 2006). A tiny green genetic cluster is present in the wild populations and in the Loreto region cultivated populations. Conversely, the green cluster is relevant in Selva Central and Ucayali cultivated populations. The green cluster increases in the sub-Andean Selva Central region and in the higher elevated sites from the Ucayali region. In the Loreto region, the cultivated population 13 BRc possessed higher proportion of green cluster than others from this region. This population is nearby the Bretaña village, which was named after the Europeans, who arrived from the Andes and the coastal regions of Peru, during the rubber boom at the end of the XIX century (Eidt 1962).

Iquitos, in the Loreto region, is referred to as a crop domestication centre in Amazonia, created as populations expanded, and providing strong evidence that pre-conquest human populations had intensively transformed their plant resources (Clement 1999b; Clement et al. 2015). Indeed, the *I. edulis* domestication was, probably, made from local wild population and possibly started in the Loreto region, since the genetic structure of the cultivated populations from this region do not differ much from the wild ones. Moreover, they have bigger legumes and low allelic richness than the other cultivated populations, which could indicate that the selection intensity was higher here. Indeed, some authors claim that the possible origin of *I. edulis* domestication was in this region, which was also a spot of domestication for other species (Clement et al. 2015; Clement et al. 2010). Additionally, the crop is probably recently domesticated, since when the crop is an initial process of domestication no clear genetic structuring occurs, as in Brazil nut (Clement et al. 2010). The genetic differentiation between wild and cultivated populations is low and with admixture, the cultivated populations seem to have origin in the wild ones. Conversely, Dawson's et al. (2008) chloroplast haplotype composition results displayed a completely different pattern between natural and cultivated populations. The authors explained those results by a non-local origin of the *I. edulis* cultivated material. Results in this thesis do not support this theory, instead it was inferred that the cultivated populations had local germplasm origin, yet without representative sampling, which is expected, since few trees were probably selected in nearby wild populations. Indeed, a possible genetic drift effect (change in the frequency of the allele in a population due to random sampling of organisms) in the cultivated populations is expected.

7.5 PRACTICAL MEASURES TO MAINTAIN *I. EDULIS* GENETIC RESOURCES

The *I. edulis* germplasm management should focus both on the wild and the cultivated stands. In case of wild material, the protection of the original Amazonian vegetation remnants is key to maintaining the species' genetic resources in the region. In modern-day Amazonia, increasing deforestation for establishment of pastures has become a global concern due to its impacts on biodiversity (Miller and Nair 2006). Considering the cultivated stands, the villages and indigenous settlements are the units of interest because they are the domesticated plant population keepers. Consequently, the fate of the village will determine the maintenance of the crop genetic resources. For example, the post-Colombian population collapse resulted in an equal loss of village units and the loss in human numbers (ca 90-95% population decline)

quickly reflected in loss of crop diversity (Clement 1999; Miller and Nair 2006). The cultivated populations with low genetic diversity and/or high inbreeding estimates (e. g., 7 VHc, 10 ARc, 12 SSc, 14 JHc, 17 EPc, 18 EDc, 19 MAc and 21 INc) should be fuelled with new germplasm sources (from wild populations) to eliminate the risk of inbreeding and diversity loss, which might be reflected in the future crop value (inbreeding depression, flower abortion, and crop yield failure).

The results of the current study on *I. edulis* show significantly higher value of legume length average in cultivated than in the wild trees. The wide scale infusion from wild stands into farms could negatively affect fruit size and weaken domestication efforts over time. Additionally, the Loreto region displayed the highest average legume length and the populations with lower allelic richness, compared to the other regions' cultivated populations. This observation is supported by crop domestication in the Amazonian region studies (Clement et al. 2015 and references therein). Therefore, the cultivated stands in Selva Central and Ucayali region could, additionally, be a germplasm material source, and safeguard to long-term on-farm conservation, since the Loreto region possesses the populations with the lowest values of allelic richness. Hybridization programs using such germplasm source and local wild material with backward selection, could help increase the crop yield and genetic diversity in the cultivated populations. Additionally, new selection should consider the ongoing global change.

7.6 COMPARISON OF *INGA EDULIS* AND *INGA INGOIDES* GENETIC DIVERSITY

All studied populations of both species displayed high values of expected heterozygosity (mean $H_E \sim 0.70$, A_R=5.1). Those estimates were slightly lower than estimates in natural populations of tropical trees *I. vera* (H_E=0.87; A_R=7.7) (Cruz-Neto et al. 2014), *Symphonia globulifera* L. (H_E=0.89) (Dick and Heuertz 2008) and *Swietenia macrophylla* King (H_E=0.78) (Lemes et al. 2003), but were very similar to the expected heterozygosity estimated for *I. edulis* by Hollingsworth et al. (2005) in the same region (Peruvian Amazon) (H_E=66%). Normally, high levels of genetic diversity are maintained by high levels of gene flow facilitated by efficient pollen movement and the wide-spread occurrence of efficient self-incompatibility mechanisms (Dick et al. 2008). Some studies demonstrated that some *Inga* species are obligate outcrossers, dependent on cross pollination to set fruits and seeds (Koptur 1984; Cruz-Neto et al. 2014) (see following section).

Inbreeding values differed in both species. Whereas *I. edulis* fits the low inbreeding values found in the *I. vera* natural populations' study using the same set of molecular markers (Cruz-Neto et al. 2014), in this thesis the analyses revealed that the heterozygote frequencies in *I. ingoides* depart from the HWE, indicating either the existence of population substructure (due to the presence of genetically isolated groups, inbreeding, and/or spatial genetic structure) or null alleles. Since the estimated average frequency of null alleles is similar in both *I. edulis* and *I. ingoides*, it was hypothesized that these differences could be explained by demography characteristics, due to habitat preferences. The observed results may reflect the *I. ingoides's* pioneer ability. This species rapidly colonizes the forest gaps opened by the seasonal river fluctuation, which results in populations being formed by patches of related individuals with a highly significant deficiency in heterozygotes due to recurrent biparental inbreeding. Thus, the heterozygotes deficiency could lead to lower competition ability, possibly explaining why this species is rarely found outside the riparian zone.

In *Acacia senegal* (L.) Willd., Omondi et al. (2010) found that the only population with positive F_{IS} was even-sized, suggesting the existence of one or few cohorts, possibly established together as a result of some disturbance event, and they argued that the area was prone to flooding, which could provide a mechanism for non-random seed dispersal. Indeed, seeds dispersed downstream could help to explain the departure from HWE in *I. ingoides*, though this hypothesis ought to be tested using a similar approach found in the study made with *Calycophyllum spruceanum* in the Peruvian Amazon (Russell et al. 1999).

The differences found in *I. ingoides* N_e, a slightly higher value in the southern (RUI) population compared to the lower value in the northern population (RSI), may reflect altitudinal and flood pulse intensity differences, but may also reflect the high inbreeding value in RSI (the later reason, if it is the cause or the consequence, is difficult to disentangle). Indeed, *I. ingoides* tend to have a higher effective population size in less flooded southern areas than in those with higher river seasonal fluctuation, despite the species' tolerance to flooding, possibly due to lower biparental inbreeding. In the case of *I. edulis*, the highest N_e value was found in the western MAE population and the lower in the estern SDE population. The former population, situated closer to the Andean slopes, has a more favourable location than lesser elevated estern sites prone to flooding, but a lower value in the latter population is probably due to differences in the number of sampled individuals.

The number of private alleles in *I. ingoides* across loci was almost twice as high as in *I. edulis* for a similar number of sampled individuals (N), which may indicate a presence of more intense gene flow in the latter species, in agreement with negligible inbreeding values. Within

species, the number of private alleles seems to reflect N to a certain extent. Yet again, RUI has more than twice the P_a than RSI, for comparable N, this might be the result of a higher inbreeding value due to putative higher parental inbreeding and consanguinity in the RSI population.

7.7 *INGA EDULIS* AND *INGA INGOIDES* GENETIC STRUCTURE AND PUTATIVE SPECIES INTROGRESSION

The partition of genetic variance in the studied species (94% of the variance observed within populations and a low genetic structure 2.6% detected among populations), is very common in tropical forest tree species with high outcrossing rates and among populations with high levels of gene flow (Finkeldey and Hattemer 2007). In a previous study, similar results were found with individuals showing mixed ancestry and low differentiation among populations, reflecting strong gene flow of Kenyan populations of *Acacia senegal* (Omondi et al. 2010). Within genus *Inga*, Cruz-Neto et al. (2014) uncovered a similar pattern in the *I. vera* species.

Weak population genetic structure may be a consequence of the pollination system and also outcrossing in the populations under study. The majority of *Inga* species can be considered hawkmoth pollinated despite occasional visitation by bats and hummingbirds (Cruz-Neto et al. 2014 and references therein). Hawkmoths, bats and hummingbirds can fly across large areas, ca. 15 km, during their foraging routes carrying pollen grains to distant individuals (Koptur 1984). Pollen flow between distant individuals in different populations, due to pollinator behaviour, contributed to high outcrossing rate and weak population substructure found in, e.g., *I. vera* natural populations (Cruz-Neto et al. 2014). Additionally, natural seed dispersal is performed by mammals and possibly birds that eat the sarcotesta and drop seeds elsewhere (Koptur 1984). Indeed, in a broad study with tropical tree species with abiotic seed dispersal (gravity dispersed and wind dispersed) showed, on average, much higher differentiation among population (G_{ST}=0.138) than animal dispersed species (G_{ST}=0.050) (Loveless 1992).

The weak population's genetic structure together with the lack of isolation-by-distance (data not shown) suggests that species ecology, such as pollen and seed dispersal, and demographic history (impacted by flood) is a strong driver of population structure in the studied *I. edulis* and *I. ingoides* populations, as in the case of *Acacia senegal* (Omondi et al. 2010).

The Bayesian approach identified two to four clusters of genetically mixed individuals in both species, with higher admixture in those places where the two species were sympatric. Thus, we could assume that the populations were not reproductively isolated, and, probably, not well

separated taxonomically. Nevertheless, some authors claim that some species of the *Inga* genus are cross-incompatible (e.g., Koptur 1984), but the data they presented does not support that conclusion, since the fruit set from hand cross-pollinated trees is clearly superior to the control. Petit et al. (2004) reviewed the hybridization between two widespread and largely sympatric European oak species (*Quercus petraea* (Matt.) Liebl. and *Q. robur* L.). They indicate that the parental taxa remain distinct, despite regular levels of gene flow between them, and emphasize the low differentiation found between both species. Yet, nuclear markers show more or less important differences in allelic frequencies between species. In another study, Moran et al. (2012) indicate that hybridization is pervasive in many plant taxa, with consequences for species taxonomy and local adaptation. They also indicate that oaks (*Quercus* spp.) are a paradigmatic case, since they are thought to hybridize readily yet retain distinct traits, drawing into question the biological species concept for such taxa, but the true extent of gene flow is controversial. Such reasoning could be extended to the *Inga* genus.

I should clarify that the morphological identification of all the individuals of the current study were rechecked with the key species identification clues according to morphology and no ambiguities were found. Selection against hybrids could hamper speciation in the *Inga* genus, but at least the past gene flow should be present in contact areas, which is the case of populations' species pairs: RUI-RUw, RPI-RPw and RSI-RSw, except in the more isolated *I. edulis* MAw and SDw populations. Introgression may be facilitated when species co-occur in areas where no intermediate habitats exist between the species ranges (Moran et al. 2012 and references therein). In our studied species, it seems that the opportunity for introgression should be close to the riverside, since *I. edulis* is relatively flood tolerant, and *I. ingoides* is probably more shade intolerant or at least less competitive in this very harsh and competitive environment. Clearly the populations of *I. edulis* close to the rivers, where the two species overlap, suffer higher introgression, which is predictable due to the fact that the *I. ingoides* habitat is mainly found there. Endara and Jaramillo (2011) developed a study on the influence of microtopography on the distribution of *Inga* species. These authors indicate that one of the main factors explaining the distribution of the *Inga* species is the soil water content. Nine out the 16 more frequent *Inga* sympatric species they analysed had a significant preference for one type of microtopography: "slope" and "ridge" (well drained) or "valley" (poorly drained soils). This fact indicates the importance of the microhabitat to the sympatric species coexistence in the *Inga* species, and that edaphic specialization among species may create more available niches. Similarly, also in oaks, *Q. robur* appears to be more tolerant to soil anoxia than *Q. petraea*, and in mixed stands succession towards the later would be the rule, except under permanently humid conditions (Petit

et al. 2004). Indeed, dynamic speciation through disruptive selection is also a hypothesis to be considered for the two studied *Inga* species.

In summary, there are no studies available on *I. edulis* and *I. ingoides* hybridization. In this study in the two studied *Inga* species it was hypothesized, that the opportunity for hybridisation exists. Firstly, the natural distribution of the two species overlaps, although in our study the differences in habitat reflected the location of the sampled individuals of both species, with *I. edulis* found mainly in non-flooded terraces or temporarily flooded sites, and with *I. ingoides* found predominantly/largely in periodically flooded areas (Pennington 1997). Secondly, in some studies based on *I. ingoides* and *I. edulis* flowering phenology observations indicate synchronous flowering, which is also common in other *Inga* species (Cruz-Neto et al. 2011; Pennington 1997; Koptur 1984) . Thirdly, the putative introgression between both species is also supported by low differentiation in microsatellite allele frequencies between the two co-occurring species (3.6%), suggesting at least past gene flow (Moran et al. 2012). Lastly, both species are closely related from the genotypic point of view, which is also supported by the phylogenetic study done by Dexter et al. (2010), where they are found in the same node with 99% support. In addition, speciation in the *Inga* genus is recent, and it is considered a classic example of a recent radiation with evidence for many species arising within the last 10 million years, some of them as recently as 2 million years ago (Richardson et al. 2001). Actually, due to a rapid and recent burst of diversification from the most recent common ancestor of the extant species, they found a poorly resolved phylogeny.

8 CONCLUSION

8.1 PRACTICAL MEASURES TO MAINTAIN *I. EDULIS* GENETIC RESOURCES

The *I. edulis* germplasm management should focus both on the wild and the cultivated stands. In case of wild material, the protection of the original Amazonian vegetation remnants is key to maintaining the species' genetic resources in the region. In modern-day Amazonia, increasing deforestation for establishment of pastures has become a global concern due to its impacts on biodiversity (Miller and Nair 2006). Considering the cultivated stands, the villages and indigenous settlements are the units of interest because they are the domesticated plant population keepers.

The results of the current study on *I. edulis* show significantly higher value of legume length average in cultivated than in the wild trees. The wide scale infusion from wild stands into farms could negatively affect fruit size and weaken domestication efforts over time. Additionally, the Loreto region displayed the highest average legume length and the populations with lower allelic richness, compared to the other regions' cultivated populations. This observation is supported by crop domestication in the Amazonian region studies (Clement et al. 2015). Therefore, the cultivated stands in Selva Central and Ucayali region could, additionally, be a germplasm material source, and safeguard to long-term on-farm conservation, since the Loreto region possesses the populations with the lowest values of allelic richness. Hybridization programs using such germplasm source and local wild material with backward selection, could help increase the crop yield and genetic diversity in the cultivated populations. Additionally, new selection should consider the ongoing global change.

8.2 SUITABILITY OF A HYBRIDIZATION PROGRAM FOR *I. EDULIS* AND *I. INGOIDES*

The use of wild hybrids and the establishment of a breeding program making use of the two species and their incorporation into agroforestry systems, could bring important economical income to the periodically flooded arable lands with limited commercial use in the Amazon basin. The ability of "pioneer" light-demanding species to grow in open spaces and inhospitable lands, could bring those species into the forefront of our concerns, by making flooded sites usable by flood-resistant and performing hybrids. Natural hybrids occur and are common in the species contact areas, according to the rusults reached in this thesis, which are also indicative that artificial hybrids are possible and doable. Thus, natural hybrids' selection and/or artificial hybridization between *I. edulis* and *I. ingoides* could be applied to improve legume size and yield in the latter species, while maintaining tolerance to flooding.

The success of the hybrids is dependent on two very important aspects to the development of these hybrids for commercial deployment. Firstly, hybrid variation and therefore selection within hybrids is dependent on the diversity of the parent species involved. Secondly, successful hybrid utilization is largely dependent on the vegetative propagation ability of the species (Potts and Dungey 2004).

Our study revealed relatively high genetic diversity in both species, but care should be taken in avoiding related trees, particularly in the case of *I. ingoides*. We advise that future studies about hybridization and introgression in both species should be done together with flooding tolerance ability and legume and yield in hybrids testing, and wild hybrids could be procured by making use of today's available approaches. Also, vegetative propagation could be used to propagate hybrids, since *Inga* species can be easily propagated from semi-ripe branch cuttings, and, for example, *I. edulis* is considered an easy-to-root species.

9 REFERENCES

Acuña C.d. 1994. Novo descobrimento do Grande Rio das Amazonas. Agir, Rio de Janeiro.

Anderson-Gerfaud PC. 1988. Using prehistonc stone tools to harvest cultivated Wild cereals preliminary observations of traces and Impact Pages 175-195 *In* Industries Ithiques, tracéologie et technologie (S Beynes, ed) BAR international Senes 411, Oxford, UK.

Barros LM, Paiva JR, Cavalcanti JJ, Araujo JPP. 2002. Cajueiro. In Melhoramento de Fruteiras Tropicais. University Federal de Vicosa, pp.159 – 176.

Bates HW. 1910. The Naturalist on the River Amazon, London: J.M. Dent.

Batlle I, Tous J. 1997. *Carob tree [Ceratonia siliqua* L.*]*. Promoting the conservation and use of underutilized and neglected crops. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy.

Brooks TM, Mittermeier RA, da Fonseca GAB, Gerlach J, Hoffmann M, Lamoreux JF, Mittermeier CG, Pilgrim JD, Rodrigues ASL. 2006. Global biodiversity conservation priorities. *Science* **313**:58-61. doi:10.1126/science.1127609.

Cabrera L. 1978. Diccionario de azýtequismos. D.F.Oasis, Mexico.

Carvajal G.d. 1970. The Discovery of the Amazon. AMS Press, New York.

Chernela JM. 1986. Os cultivares de mandioca na área de Uaupés (Tukano). In Suma etnológica brasileira. Vozes/FINEP, Petrópolis, pp. 151 – 158.

Clement CR. 1989. A center of crop genetic diversity in Western Amazonia: a new hypothesis of indigenous fruit-crop distribution. *Bioscience* **39**: 624 – 631.

Clement CR. 1999a. 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. *Economic Botany* **53**: 188 – 202.

Clement CR. 1999b. 1492 and the loss of Amazonian crop genetic resources. II. Crop biogeography at contact. *Economic Botany* **53**: 203 – 216.

Clement C, De Cristo-Araújo M, Coppens D'Eeckenbrugge G, Alves Pereira A, Picanço-Rodrigues D. 2010. Origin and domestication of native Amazonian crops. Diversity **2**:72

Clement CR, Denevan WM, Heckenberger MJ, Junqueira AB, Neves EG, Teixeira WG, Woods WI. 2015. The domestication of Amazonia before European conquest. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20150813. doi:10.1098/rspb.2015.0813.

Condamine ChM. 1944. Relato abreviado de una viagem pelo interior da América meridional. Trad. de A. Avila. Sao Paulo, Edicóes Cultura, pp. 214.

Coates AG, Obando JA. 1996. The geological evolution of the Central American isthmus. In: Jackson JBC et al. (Editors), Evolution and Environment in Tropical America. University of Chicago Press, Chicago.

Cruz-Neto O, Machado I, Duarte JJr, Lopes A. 2011. Synchronous phenology of hawkmoths (Sphingidae) and *Inga* species (Fabaceae–Mimosoideae): implications for the restoration of the Atlantic forest of northeastern Brazil. *Biodiversity Conservation* **20**:751-765. doi:10.1007/s10531-010-9975-x.

Cruz-Neto O, Aguiar AV, Twyford AD, Neaves LE, Pennington RT, Lopes AV. 2014. Genetic and ecological outcomes of *Inga vera* subsp *affinis* (Leguminosae) tree plantations in a fragmented tropical landscape. PLoS One **9**. doi:10.1371/journal.pone.0099903.

Daniel J. 1976. Tesouro descoberto no Rio Amazonas. Biblioteca Nacional, Rio de Janeiro.

Dawson I, Hollingsworth P, Doyle J, Kresovich S, Weber J, Sotelo Montes C, Pennington T, Pennington R. 2008. Origins and genetic conservation of tropical trees in agroforestry systems: a case study from the Peruvian Amazon. *Conservation Genetics* **9**:361-372. doi:10.1007/s10592-007-9348-5.

Dawson IK, Lengkeek A, Weber JC, Jamnadass R. 2009. Managing genetic variation in tropical trees: linking knowledge with action in agroforestry ecosystems for improved conservation and enhanced livelihoods. *Biodiversity Conservation* **18**:969-986. doi:10.1007/s10531-008-9516-z.

Dawson IK, Leakey R, Clement CR, Weber JC, Cornelius JP, Roshetko JM, Vinceti B, Kalinganire A, Tchoundjeu Z, Masters E, Jamnadass R. 2014. The management of tree genetic resources and the livelihoods of rural communities in the tropics: Non-timber forest products, smallholder agroforestry practices and tree commodity crops. *Forest Ecology and Management* **333**:9-21. doi:10.1016/j.foreco.2014.01.021.

Dayanandan S, Bawa K, Kesseli R. 1997. Conservation of microsatellites among tropical trees (Leguminosae). *American Journal of Botany* **84**:1658-1663.

de Jong W. 1995. *Diversity, Variation and Change in Ribereno Agriculture and Agroforestry*. Doctoral thesis, Wageningen University for Life Sciences, The Netherlands.

Denevan WM. 1992. Native American populations in 1492: recent research and a revised hemispheric estimate. Pages xvii-xxxvii *in* W. M. Denevan, ed., The native population of the Americas in 1492. University of Wisconsin Press, Madison, Wisconsin.

Dexter KG, Pennington TD, Cunningham CW. 2010. Using DNA to assess errors in tropical tree identifications: How often are ecologists wrong and when does it matter? *Ecology Monographs* **80**:267-286. doi:10.1890/09-0267.1.

Dias LAS. 2001. *Melharamento Genético do Cacaueiro*. Funape, Viçosa, Brazil, 578p.

Dobyns H. 1966. An appraisal of techniques for estimating aboriginal population with a new hemispheric estimate, *Current Anthropology* **7**: 395–416.

Ducke A. 1946. Plantas de cultura precolombiana na Amazonia Brasileira. Notas sobre as espécies ou formas espontaneas que supostamente lhas terian dado origem. Boletin Técnico Inst. Agron. Norte, Belem N°8, 24p.

Duke JA. 1983. *Handbook of Energy Crops*, unpublished, [online]. In: http://www.hort.purdue.edu/newcrop/ [cit. 2008-05-02].

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**:359-361. doi:10.1007/s12686-011-9548-7.

Egg AB, Vargas CM. 2004. Ecología del Perú. Bruño, Lima, Peru, pp. 495.

Ehrich D, Gaudeul M, Assefa A, Koch MA, Mummenhoff K, Nemomissa S, Consortium I, Brochmann C. 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Molecular Ecology* **16**:2542-2559. doi:10.1111/j.1365-294X.2007.03299.x.

Eidt RC. 1962. Pioneer settlement in Eastern Peru. *Annals of the Association of American Geographers* **52**:255-278.

El Mousadik A, Petit RJ. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92**:832-839.

Endara MJ, Jaramillo JL. 2011. The influence of microtopography and soil properties on the distribution of the speciose genus of trees, *Inga* (Fabaceae:Mimosoidea), in Ecuadorian Amazonia. *Biotropica* **43**:157-164. doi:10.1111/j.1744-7429.2010.00668.x.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**:2611-2620.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**:564-567. doi:10.1111/j.1755-0998.2010.02847.x.

Ferreira AR. 1972. Viagem filoso´ fica pelas Capitanias do Grão-Pará, Rio Negro, Mato Grosso e Cuiabá. Conselho Federal de Cultura, Rio de Janeiro.

Figueiredo MF, Bruno RLA, Barros e Silva AE, Nascimento S, Oliveira IG, Felix LP. 2014. Intraspecific and interspecific polyploidy of Brazilian species of the genus *Inga* (Leguminosae: Mimosoideae). *Genetics and Molecular Research* **13 (2)**: 3395-3403 DOI http://dx.doi.org/10.4238/2014.April.29.18.

Finkeldey R, Hattemer HH. 2007. Tropical forest genetics. Springer-Verlag, Berlin, Heidelberg. doi:10.1007/978-3-540-37398-8.

Fritz S. 1922. *The Journal of the travels of Father Samuel Fritz*. The Haklupt Society.

Fujisaka S, Escobar G, Veneklaas EJ. 2000. Weedy fields and forest: interactions between land use and the composition of plant communities in the Peruvian Amazon. *Agriculture Ecosystems & Environment* **78**:175-186.

Goudet J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**:485-486.

Haffer J. 1982. "General aspects of the refuge theory", in Prance GT. Ed. Biological Diversification in the Tropics, New York: Columbia University Press. Pp 6-24.

Hamrick J, Godt M, Sherman-Broyles S. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forest* **6**:95-124.

Hemming J. 1978. Red gold: The conquest of the Brazilian indians. London, pp. 677.

Hernández F. 1959. Historia natural de Nueva España. Universidad Nacional Autónoma de México. 554p.

Hollingsworth PM, Dawson IK, Goodall-Copestake WP, Richardson JE, Weber JC, Sotelo Montes C, Pennington RT. 2005. Do farmers reduce genetic diversity when they domesticate tropical trees? A case study from Amazonia. *Molecular Ecology* **14**:497-501.

Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecological Resources* **9**:1322–1332.

International Board for Plant Genetic Resources. 1980. Tropical Fruit Descriptor. IBPGR Secretariat, Rome, Italy.

International Board for Plant Genetic Resources. 1993. Descriptors for Pigeonpea [*Cajanus cajan* (L.)Millsp.]. ICRISAT, Patancheru, India.

Jackson JE. 1983. The fish people: linguistic exogamy and Tukanoan identity in Northwest Amazonia. Cambridge University, Cambridge.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**:1801-1806. doi:10.1093/bioinformatics/btm233.

Jamnadass, R., Lowe, A., Dawson, I.K. 2009. Molecular markers and the management of tropical trees: the Case of Indigenous Fruits. *Tropical Plant Biology* **2**: 1-12.

Jones TH, Steane DA, Jones RC, Pilbeam D, Vaillancourt RE, Potts BM. 2006. Effects of domestication on genetic diversity in *Eucalyptus globulus*. *Forest Ecology and Management* **234**:78-84.

Junk W, Piedade MF. 2011. An introduction to South American wetland forests: distribution, definitions and general characterization. In: Junk WJ, Piedade MTF, Wittmann F, Schöngart J, Parolin P (eds) Amazonian floodplain forests. Ecophysiology, biodiversity and sustainable management, vol 210. Ecological Studies. Springer Netherlands, Heidelberg, pp 3-25. doi:10.1007/978-90-481-8725-6_1.

Kalliola R, Paytan SF. 1998. Geoecologia y Desarrollo Amazonico: Estudio Integrado en la Zona de Iquitos, Peru. Sulkava: Finnreklama Oy; p. 544.

Kanmegne J, Bayomock LA, Duguma B, Ladipo DO. 2000. Screening of 18 agroforestry species for highly acid and aluminium toxic soils of the humic tropics. *Agroforestry Systems* **49** (1), 31–39.

Koptur S. 1984. Outcrossing and pollinator limitation of fruit-set: Breeding systems of Neotropical Inga trees (Fabaceae: Mimosoideae). *Evolution* **38**: 1130-1143.

Köppen W. 1936. Das geographisca System der Klimate, in: Handbuch der Klimatologie, edited by: Köppen, W. and Geiger, G., 1. C. Gebr, Borntraeger, pp. 1–44.

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. 2006. World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* **15**:259-263. DOI: 10.1127/0941-2948/2006/0130.

Lathrap D. 1977. Our father the cayman, our mother the gourd: Spinden revisited, or a unitary model for the emergence of agriculture in the New World. In *Origins of Agriculture*; Reed, C.A., Ed.; Mouton: The Hague, Holand, pp. 713-751.

Lavin M. 2006. "Floristic and geographical stability of discontinuous seasonally dry tropical forests explains patterns of plant phylogeny and endemism,"in *Neotropical Savannas and Seasonally Dry Forests:Plant Diversity, Biogeography and Conservation*, eds R.T.Pennington, G.P.Lewis, and J.A.Ratter (Boca Raton, FL: CRC Press),433–447.

Lawrence A. 1993. *Inga edulis*: a tree for acid soils in the humid tropics. *NFT Highlights* 93-04, Nitrogen Fixing Tree Association, [online].
In: http://www.winrock.org/fnrm/factnet/factpub/FACTSH/inga.htm.

Lawrence A. 1995. Inga: Farmers knowledge and use of Inga species. In: Evans DO, Szott LT. ed. Nitrogen Fixing Trees for Acids Soils:pp. 142-151.

Lemes MR, Gribel R, Proctor J, Grattapaglia D. 2003. Population genetic structure of mahogany (*Swietenia macrophylla* King, Meliaceae) across the Brazilian Amazon, based on variation at microsatellite loci: implications for conservation. *Molecular Ecology* **12**:2875-2883.

León J. 1966. Central American and West Indian Species of *Inga* (*Leguminosae*). *Annals of the Missouri Botanical Garden* **53(3)**: 365-459.

León J. 1987. Botánica de los Cultivos Tropicales. San José, Costa Rica: Instituto Interamericano parala Cooperación Agrícola (IICA).

León J. 1998. History of the utilization of *Inga* as fruit trees in Mesoamerica and Peru. In. Pennington TD, Fernandes ECM. The genus *Inga*: Utilization. The Royal Botanic Gardens, Kew, London, United Kingdom.

Levis C,Costa FRC,Bongers F, Peña-Claros M,Clement CR, Junqueira AB, Neves EG, Tamanaha EK, Figueiredo FOG, Salomão RP, Castilho CV, Magnusson WE, Phillips OL, Guevara JE, Sabatier D, Molino JF, Cárdenas López D,Mendoza AM, Pitman NCA, Duque A, Núñez Vargas P, Zartman CE, Vasquez R, Andrade A, Camargo JL, Feldpausch TR, Laurance SGW, Laurance WF, Killeen TJ, Mendonça Nascimento HE, Montero JC, Mostacedo B, Amaral IL, Guimarães Vieira IC, Brienen R, Castellanos H, Terborgh J, Jesus Veiga Carim M de, Silva Guimarães JR da, Souza Coelho L de, Almeida Matos FD de, Wittmann F, Mogollón HF, Damasco G, Dávila N, García-Villacorta R, Coronado ENH, Emilio T, Andrade Lima Filho D de, Schietti J, Souza P, Targhetta N, Comiskey JA, Marimon BS, Marimon Jr. BH, Neill D, Alonso A, Arroyo L, Carvalho FA, Souza FC de, Dallmeier F, Pansonato MP, Duivenvoorden JF, Fine PVA, Stevenson PR, Araujo-Murakami A, Aymard GA, Baraloto C, Amaral DD do, Engel J, Henkel TW, Maas P, Petronelli P, Cardenas Revilla JD, Stropp J, Daly D, Gribel R, Ríos Paredes M, Silveira M, Thomas-Caesar R, Baker TR, Silva NF da, Ferreira LV, Peres CA, Silman MR, Cerón C, Valverde FC, Di Fiore A, Jimenez EM, PeñuelaMora MC, Toledo M, Barbosa EM, Matos Bonates LC de, Arboleda NC, Sousa Farias E de, Fuentes A, Guillaumet JL, Møller Jørgensen P, Malhi Y, Andrade Miranda IP de, Phillips JF, Prieto A, Rudas A, Ruschel AR, Silva N, Hildebrand P von, Vos VA, Zent EL, Zent S, Cintra BBL, Nascimento MT, Oliveira AA, Ramirez-Angulo H, Ramos JF, Rivas G, Schöngart J, Sierra R, Tirado M, Heijden J van der, Torre EV, Wang O, Young KR, Baider C, Cano A, Farfan-Rios W, Ferreira C, Hoffman B, Mendoza C, Mesones I, Torres-Lezama A, Medina MNU, Andel TR van, Villarroel D, Zagt R, Alexiades MN, Balslev H, Garcia-Cabrera K, Gonzales T, Hernandez L, Huamantupa-Chuquimaco I, Manzatto AG, Milliken W, Cuenca WP, Pansini S, Pauletto D, Arevalo FR, Costa Reis NF, Sampaio AF, Urrego Giraldo LE, Valderrama Sandoval EH, Valenzuela Gamarra L, Vela CIA, Steege H ter. 2017. Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science* **355**: 925–931.

Levis C, Flores BM, Moreira PA, Luize BG, Alves RP, Franco-Moraes J, Lins J, Konings E, Peña-Claros M, Bongers F, Costa FRC, Clement CR. 2018. How People Domesticated Amazonian Forests. *Frontiers in Ecology and Evolution* **5**. doi:10.3389/fevo.2017.00171.

Lewis G, Schrire B, Mackinder B, Lock M. 2005. *Legumes of the World*: The Royal Botanic Gardens, Kew, the United Kingdom. pp. 200.

Longman KA, Jeník J. 1987. *Tropical forest and its environment*. John Wiley & Sons, Inc., New York, USA. ISBN 0-582-44678-3.

Loveless MD. 1992. Isozyme variation in tropical trees: patterns of genetic organization. *New Forest* **6**:67-94. doi:10.1007/bf00120640.

Meggers BJ. 1971. Amazonia. Man and Culture in a Counterfeit Paradise. Aldine, Chicago, pp. 182.

Miller ET. 1992. *Arqueologia nos Empreendimentos Hidroeléctricos da Eletronorte: Resultados Preliminares*. Eletronorte: Brasília, DF, Brazil.

Miller RP, Nair PKR. 2006. Indigenous agroforestry systems in Amazonia: from prehistory to today. *Agroforestry Systems* **66**:151–164.

MINAG 2002. *Datos Climatologicos de la Región Ucayali*. Dirección del Información Agraria, Pucallpa, Peru.

Moran EV, Willis J, Clark JS. 2012. Genetic evidence for hybridization in red oaks (*Quercus* sect. Lobatae, Fagaceae). *American Journal of Botany* **99**:92-100. doi:10.3732/ajb.1100023.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nichols JD, Carpenter FL. 2006. Interplanting Inga edulis yields nitrogen benefits to *Terminalia amazonia. Forest Ecology and Management* **233**: 344–351. doi:10.1016/j.foreco.2006.05.031.

Nicholls JA, Pennington RT, Koenen EJM, Hughes CE, Hearn J, Bunnefeld L, Dexter KG, Stone GN, Kidner CA. 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus Inga (Leguminosae: Mimosoideae). *Frontiers in Plant Science* **6**: 710. doi: 10.3389/fpls.2015.00710.

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao K, Kalbfleisch T, Schulz V, Kreitman M, Bergelson J. 2005. The pattern of polymorphism in *Arabidopsis thaliana*. *PLOS Biology* **3**:e196. doi:10.1371/journal.pbio.0030196.

Nybom H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**:1143-1155.

Odar CAP, Rodríguez CV. 2004. *Boletin Meteorologico*:. UNU, Pucallpa, Peru, pp. 37.

Oliveira PJC, Asner GP, Knapp DE, Almeyda A, Galván-Gildemeister R, Keene S, Raybin RF, Smith RC. 2007. Land-use allocation protects the Peruvian Amazon. *Science* 317:1233-1236. doi:10.1126/science.1146324.

Omondi S, Kireger E, Dangasuk O, Chikamai B, Odee D, Cavers S, Khasa D. 2010. Genetic diversity and population structure of *Acacia senegal* (L) Willd. in Kenya. *Tropical Plant Biology* **3**:59-70. doi:10.1007/s12042-009-9037-2.

O´Neil GA, Dawson I, Sotelo-Montes C, Guarino L, Guariguata M, Current D, Weber JC. 2001. Strategies for genetic conservation of trees in the Peruvian Amazon. *Biodiversity and Conservation* **10**:837-850.

Ormsby T, Napoleon, E., Burke, R., Groessl, C., Bowden, L. 2010. Getting to know ArcGIS desktop. ESRI Press, Redlands, California.

Paytan SF. 1997. *Cultivo de Frutales Nativos Amazonicos*: 115-119. Tratado de Cooperacion Amazonica, Lima, Peru.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**:2537-2539. doi:10.1093/bioinformatics/bts460.

Pennington TD. 1997. The genus *Inga*: Botany. The Royal Botanic Gardens, Kew, London, United Kingdom.

Pennington TD; Fernandes ECM. 1998. The genus *Inga*: Utilization. The Royal Botanic Gardens, Kew, London, United Kingdom. pp. 167.

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**:689-701.

Piperno DR; Pearsall DM.1998. *The Origins of Agriculture in the Lowland Neotropics*. Academic Press: San Diego, CA, USA, p. 400.

Piperno DR, Ranere AJ, Holst I, Hansell P. 2000. Starch grains reveal early root crop horticulture in the Panamanian tropical forest. *Nature* **407**: 894 – 897.

Potts BM, Dungey HS. 2004. Interspecific hybridization of Eucalyptus: key issues for breeders and geneticists. *New Forest* **27**:115-138. doi:10.1023/a:1025021324564.

Prance GT. 1982. Biological diversification in the Tropics. New Yourk: Columbia University Press.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959.

Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG. 2014. STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus* **3**:431. doi:10.1186/2193-1801-3-431.

Reynel C, Pennington TD. 1997. El género Inga en el Perú Morfología, Distribución y usos. The Royal Botanic Gardens, Kew, London, United Kingdom.

Reynel C, Pennington RT, Pennington TD, Flores C, Daza A. 2003. Árboles útiles de la Amazonía Peruana. Tarea Asociación Gráfica Educativa, Lima, Peru, pp. 261-270.

Ribeiro MM, Plomion C, Petit R, Vendramin GG, Szmidt AE. 2001. Variation in chloroplast single-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theorrtical and Applied Genetics* **102**:97-103. doi:doi:10.1007/s001220051623.

Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* **293**: 2242-2245. doi:10.1126/science.1061421.

Roosevelt AC. 1989. Resource management in Amazonia before the conquest: beyond ethnographic projection. Pages 30-62 *in* D. A. Posey, and W. Bal6e, eds., Resource management in Amazonia: indigenous and folk strategies. Advances in Economic Botany 7. New York Botanical Garden, New York.

Roosevelt AC. 1993. The rise and fall of Amazonian chiefdoms. *L´Homme* **33 (2-4)**: 255 – 282.

Roosevelt AC. 1994. Amazonian anthropology: Strategy for a new synthesis. In: Amazonian Indians from Prehistory to Present: Anthropological Perspectives. Univ. of Arizona Press, Tucson, pp. 1 – 29.

Roosevelt A; Costa ML; Machado CL; Michab M; Mericer N; Valladas H; Feathers J; Barnett W; Silveira MI; Henderson A; Silva J; Chernoff B; Reese DS; Holman JA; Toth N; Shick K. 1996. Paleoindian cave dwellers in the Amazon: the peopling of the Americas. *Science* **272**:373-384.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**:137-138. doi:10.1046/j.1471-8286.2003.00566.x.

Rousset F. 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* **8**:103-106. doi:10.1111/j.1471-8286.2007.01931.x.

Russell JR, Weber JC, Booth A, Powell W, Sotelo-Montes C, Dawson IK. 1999. Genetic variation of *Calycophyllum spruceanum* in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. *Molecular Ecology* **8**:199-204. doi:10.1046/j.1365-294X.1999.00551.x.

Salick J, Lundberg M. 1990. Variation and Change in Amuesha Agriculture, Peruvian Upper Amazon. In: New Directions in the Study of Plants and People. NY Botanical Garden, New York, pp. 199 – 223.

Schroth G, Fonseca GAB, Harvey C, Gascon C, Vasconcelos HL, Izac A-MN. 2004. Agroforestry and Biodiversity Conservation in Tropical Landscapes. Island Press, Washington, DC.

Smith NJH, Fik TJ, Alvim PdT, Falesi IC, Serraõ EAS. 1995. Agroforestry developments and potential in the Brazilian Amazon. *Land Degradation & Rehabilitation* **6**:251 – 263.

Smith BD. 2001. Documenting plant domestication: the consilience of biological and archeological approaches. *Proceedings of the National Academy of Sciences* **98**:1324 – 1326.

Sokal RR, Rohlf FJ. 1997. Biometry. 3rd ed. edn. W.H. Freeman and Co., San Francisco.

Sousa SM. 1993. El género *Inga* (Leguminosae: Mimosoideae) del sur de México y Centroamérica, estudio previo para la Flora Mesoamericana. *Annals of the Missouri Botanical Garden* **80**: 223-269.

Strudwick J, Sobel GL. 1988. Uses of *Euterpe oleracea* Mart. In the Amazon estuary, Brazil. *Advances in Economic Botany* **6:** 225-253.

ter Steege H, Pitman NCA, Sabatier D, Baraloto C, Salomão RP, Guevara JE. 2013. Hyper dominance in the Amazonian tree flora. *Science* **342**:1243092. doi:10.1126/science.1243092.

Tilki F, Fisher RF. 1998. Tropical leguminous species for acid soils: studies on plant form and growth in Costa Rica. *Forest Ecological Management* **108:**175–192.

Vavilov NI. 1992. On the origin of cultivated plants (First published in 1926 in Novoye v Agronomii [News in Agronomy]). Pages 14-21 *in* V. E Dorofeyev, ed., Origin and geography of cultivated plants. Cambridge University Press, Cambridge.

Villachica H. 1996. Frutales y Hortalizas Promisorios de la Amazonia: 124-127. Tratado de Cooperacion Amazonica, Lima, Peru.

Wang J. 2017. The computer program structure for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources*. **17**: 981-990

Weber JC. 2001. Strategies for genetic conservation of trees in the Peruvian Amazon. *Biodiversity and Conservation* **10**: 837-850.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358-1370.

White D, Velarde SJ, Alegre J, Tomich TP. 2005. Alternatives to slash-and-burn (ASB) in Peru. *Summary Report and Synthesis of Phase II.* Alternatives to Slash-and-Burn Programme, Nairobi, Kenya.

Whitmore TC, Prance GT. Eds. 1987. *Biogeography and Quaternary History in Tropical America* (Oxford Science Publications, Oxford, 1987).

Wiseman B, Groves J. 2000. Introducing Le´vi-Strauss and Structural Anthropology. Icon Books, Cambridge.

APPENDIX

Inga edulis – habitus (cultivated/domestic form)

(wild form)

Inga edulis – legumes (left side: domestic; right side bottom: wild; right side top: another *Inga* sp.)

Inga edulis - edible pulp (domestic)

(wild)

Inga edulis – seeds (domestic)

(wild)

Inga edulis - inflorescence

I. edulis - legume and rhachis wing shape variability

Inga ingoides in its typical habitat (periodicaly flooded and poorly drained site).

Inga ingoides – leaf and legume

Detail on *I. ingoides* regular aperture of foliar nectary (one of the key species identification aspect).

All photos by Alexandr Rollo