The phylogeography of *Cheracebus* species located in the Rio Negro basin and its implications for our understanding of the historical biogeography of Amazonia.

By Amy Green

Contents

List of tables

List of figures

Acknowledgments

I would like to extend my thanks to the University of Salford for accepting my application to conduct a masters by research, and providing the resources needed to conduct my research to the best of my ability.

I would firstly like to thank and acknowledge my supervisor, Professor Jean Boubli and my co-supervisor Dr Robert Jehle. Secondly, I would like to thank Dr Chrysoula Gubili. Without their help my research would not have been possible. I would like to thank INPA museum, Manaus for supplying me with tissue samples for my research. Lastly, I would like to thank my family for funding my research, and my fiancée for supporting me throughout.

Declaration

Name of candidate: AMY GREEN

Roll number/Student number: @00252473

School: Environment and life sciences

Degree: MSc by research in biological sciences

This is to certify that the copy of my thesis, which I have presented for consideration for my postgraduate degree: -

1. embodies the results of my own course of study and research
2. has been composed by myself
3. has been seen by my supervisor before presentation

Signature of candidate: A.Green   Date: 11/12/16

Address (to which information concerning examiners' decision and final binding can be sent):

20 Florin Lane
Salford
M6 5TF

The candidate's supervisor is asked to declare here that s/he has approved the submission of the thesis. If the supervisor decides to withhold approval, the candidate shall have the right of appeal to the Associate Head/Dean of Research. A candidate may be permitted to submit a thesis despite the Supervisor withholding approval, providing the Associate Head/Dean of Research approves submission.

Signature of Supervisor: Date: 13/12/2016

TITLE: **The phylogeography of *Cheracebus* species located in the Rio Negro basin and its implications for our understanding of the historical biogeography of Amazonia.**

Abstract

The Amazon holds the greatest primate diversity in the world. To date, little is known about how tis diversity originated. Recent evidence points to the role of rivers as important geographical barriers to primates and potential engines of speciation. Titi monkeys are the most diverse group of primates in the Amazon. Few studies have addressed the origins of such high diversity. Although it has been shown that large Amazonian rivers divide many titi monkey species, it is not known if smaller tributaries also play a role in separating sister species. In this study, we investigate if populations of widow titi monkeys *Cheracebus lugens* inhabiting the Rio Negro basin in Brazil present any genetic differentiation across this river's left bank tributaries namely, Cauaburi, Marauia, Padauari, Araca and Demeni. We also examine for the first time using molecular phylogenetics, the phylogenetic relationships of *C. lugens* with other two species of this genus *C. lucifer* and *C. purinus* and the role of Japura and Solimoes rivers in the separation of these different species. We use three mitochondrial markers (COI, Cytochrome B, D-Loop) and one nuclear marker (RAG1). We constructed haplotype networks and preferred phylogenetic analysis to analyse their genetic differences. Our results show genetic differences for these markers between the three species of titi monkey but little variation between the samples of *C. lugens* collected on the left bank of Rio Negro. Suprisingly our samples of *C. lugens* from the right bank of Rio Negro were sister to the clade including *C. purinus* and *C. lucifer*, making *C. lugens* paraphyletic. However, the sample of size for this study of very small, and we should interpret our findings with a grain of salt. More samples will be required to fully understand the relationships amongst all species of widow titi monkeys and to better understand population level variability in *C. lugens*.

Introduction

1.0 Statement of the Problem

The Amazon rainforest is home to the world's greatest diversity of primates. Currently, there are 19 genera, 7 subfamilies, and 199 recognized species and subspecies of new world monkeys. Recent boom in new species descriptions reflects how little is yet known about Amazon biodiversity. We also know very little about the origins of such great species diversity. Why is the Amazon forest so species rich? This is an old question posed by early 19th century naturalists that still inspire scientists today. Many theories have been proposed, to explain tropical species diversity including hypotheses concerning the origins (historical biogeography, e.g., Haffer 1969) and maintenance (ecological, e.g., Connell 1976) of such diversity.

Table 1.0: The number and species and subspecies in a certain part of the world, in order starting with the highest number to lowest (Rylands & Mittermeier, 2009).

Area	Number of species and subspecies
Neotropics	199
Asia	186
Africa	169
Madagascar	100

For primates, the riverine hypothesis of diversification through vicariance have been often cited as an important mechanism promoting speciation (e.g., Boubli et al 2015, Wallace 1852). The English naturalist, Alfred R. Wallace was the first to propose that large Amazonian rivers were barriers to primates (Wallace, 1852). Wallace's riverine hypothesis, as it later became known, implies that major Amazonian rivers significantly reduce or prevent gene flow between populations inhabiting opposite banks of rivers (Ayres and Clutton-Brock, 1992). Rivers as geographical barriers for primates have been a debated and controversial

topic in Madagascar, South East Asia, Africa and South America so far. Although several studies point to the importance of rivers as vicariant agents, others show no effect. It is becoming obvious then, that some rivers are barriers to some species whereas others are not (e.g. Ayres and Clutton-Brock, 1992, Boubli et al. 2015, Buckner et al. 2015). Boubli et al in 2015, showed that the Rio Negro is an important driver of diversification in three genera of primate, *Cacajao, Callicebus (*now *Cheracebus,* Byrne et al 2016) and *Cebus* but not for *Saimiri, Chiropotes, Alouatta* or *Ateles*. These results add to the complexity of this issue. The fact is that Amazonian rivers do not only act as genetic barriers, as Amazonian rivers differ in their history, volume, speed of flow, and chemistry among other factors and these differences may significantly affect the way in which rivers sever species distributions and constitute a physical barrier to species dispersal (Boubli et al, 2015).

One topic that remains overlooked is the role of smaller tributaries in the origins of speciation events. With such high primate diversity but with only a handful of large rivers, we would expect that even small tributaries should contribute to primate diversification. In this study, we investigate this possibility for the first time for widow titi monkeys (genus *Cheracebus*) inhabiting the Rio Negro basin in Brazil. Although Boubli et al (2015) show that the Rio Negro is responsible for the vicariant event that gave rise to two cryptic taxa of *Cheracebus lugens*, we still don't know if the Negro tributaries further split this widely distributed species. Given what we currently know about the effects of rivers as barriers to the dispersal of Amazonian primates (see above), I expect that Rio Negro tributaries will also play a role as a barrier, subdividing what we now recognise as one species of *Cheracebus lugens* into genetically structured population or distinct taxa. Even though we know rivers may act as barriers to primates, to what extent they act as genetic barriers is less well known. The topic of rivers acting as genetic barriers is an important question that remains unresolved, and studies addressing this question in a systematic way across the Amazon could potentially

trigger a new wave of taxonomic revisions on a large number of different taxa across the region. Taxonomic revisions are important overall because it is important to state if they are in fact the same species, especially if they are so genetically and morphologically similar. Studies such as this are used to discover whether a species may or may not qualify to be classed as the same species, no matter how small are large the genetic difference may be. Finally, my study will shed light on our understanding of historical biogeography of Amazonia. Given how little is known in terms of the geological history of the basin, biological data can offer valuable information. For Amazonian rivers, concordance in the divergence time for multiple species pairs along the two banks of a river would provide evidence for the timing of river formation (Boubli et al, 2015, Ribas et al 2012).

1.1 Primate molecular phylogenetics

The last two decades have seen an unprecedented boom in new primate species discoveries in the Amazon. The reason behind such new boom in species discoveries stems from renewed interest in Biodiversity studies and collecting expeditions to remote areas of the Amazon. Such expeditions are bringing back great numbers of new species of different taxonomical groups including primates. More important however, is the increased use of modern molecular methods in phylogenetic studies leading to a re-evaluation of previous taxonomical hypothesis often including previously overlooked diversity. The application of molecular genetics methods in primate taxonomy is now widespread in particular since protocols are becoming more streamlined, costs are lowering and high computer power is being used to analyse the increasing amount of data generated is becoming more accessible.

Phylogenetic analysis using DNA is now an essential tool for studying the phylogenetic relationships of organisms from bacteria to humans. It is now possible to study the evolutionary relationships of virtually all levels of classifications of organisms, and understand the process of adaptive evolution at a molecular level (Nei & Kumar, 2000). From the time of Charles Darwin, many biologists have tried to reconstruct the evolutionary history of all forms of life on earth. However, the evolutionary change of morphological and physiological characters is so complex, even the use of phylogenetic trees has been questioned in terms of how accurate they really are. Recent advances in molecular biology have changed this dramatically. As every organism genetic code is written in DNA, it is possible to study the phylogenetic relationships of organisms by comparing their DNA (Nei & Kumar, 2000). Phylogenetics is concerned with reconstructing the evolutionary history for a set of organisms, depicting relationships through a network of tree like structures (Borodovsky et al, 2010). Based upon similarities and differences in their physical and genetic characteristics, the taxa joined together in the tree are implied to have descended from a common ancestor (Baum, 2016).

There have been various phylogenetic studies carried out recently based on taxonomic relationships in several groups of primates, titi monkeys being one of the groups. A combination of diversity and divergence times has subsequently caused the genus *Callicebus* to be divided into three *Cheracebus, Callicebus* and *Plecturocebus* (Byrne et al, 2016).

Phylogenetic analysis of DNA sequences is increasingly helping in our understanding of historical biogeographical and evolutionary processes in Neotropical primates. Accurate estimates of phylogenetic relationships of species have allowed the reconstruction of historical dispersal events that led to present day distributions and the evaluation of historical processes in shaping present day patterns of biological diversity (Hoyos, 2016). These phylogenies reveal surprisingly deep divergence dates for major *Callicebus* clades. To date

there has been no explicit molecular investigation of phylogenetic relationships of *Callicebus,* and consequently the evolutionary history of titi monkeys remains poorly studied (Byrne et al, 2016). Studies have been carried out regarding *C.lugens* on the upper Rio Negro in Brazil by Casado, Bonvicino and Seuanez (2007) They discovered the karyoptypes were the same on either side of the Rio Negro, but had distinct haplotypes, suggesting different evolutionary lineages.

1.2 Phylogenetic trees

In the origin of species, Darwin founded evolutionary biology on the idea that organisms share a common origin and subsequently diverged through time. Phylogenies represent attempts to reconstruct the evolutionary history of life and our ability to infer phylogeny has increased dramatically. Not only has it become relatively easy to determine the DNA sequence of a gene, but computers have significantly sped up the process (huelsenbeck & Ronquist, 2001).

The species are connected via a set of lines called branches or edges which represent the evolutionary relationships between them (Figure 2). The species at the external nodes or leaves of the tree are either existing species that have not yet evolved in to new species, or are species that are now extinct (Zvelebil & Baum, 2008).

1.3 Types of phylogenetic tree

There are different types of phylogenetic trees. In some trees, the length of a branch reflects the number of genetic changes that have taken place in a particular DNA sequence in that lineage. In other trees, branch length can represent chronological time, and branching points can be determined from fossil record (Cockell, 2015). Phylogenetic trees may either be presented with a root (rooted) or without a root (un-rooted) and the tree branching pattern is known as the tree topology. A rooted tree contains a specific node, known as the root, and

the distance from the root to any other node is the same. Rooted trees tend only to be used when it is implied that evolution occurs at a constant rate among lineages within a phylogeny; the root is the common ancestor for all taxa in the phylogeny (Pometto et al, 2005).

An unrooted tree specifies relationships, however does not show the evolutionary path completely or assume common ancestors (Pevsner. J, 2008).

1.4 Phylogeography

Current patterns of gene flow may show little resemblance to the historical connections among populations, but both are relevant to the contemporary distributions of species and their genes (Freeland et al, 2011). Patterns of animal and plant distribution across planet earth have a source of fascination to naturalists. Explaining why species occupy the ranges that they do has been a problem for many scientists over the years, mostly based on ecological factors and environmental conditions. What we do know is that historical events have played a very important part in species distribution between different eras (Beebee & Rowe, 2005). Understanding how historical events have helped shape the current geographic distribution of alleles, populations and species is the major goal of phylogeography.

Phylogeography is a field of study concerned with the principles and processes governing the geographic distribution of genealogical lineages, especially those among and within closely related species (Avise, 2000). A relatively young discipline, the phylogeographic revolution inspired by mitochondrial (mt) DNA analysis that were produced nearly three decades ago, has transformed the study of population genetics and speciation in several ways. In particular, it has drawn closer empirical and conceptual connections between micro-evolutionary genetics and phylogenetic biology (Weiss & Ferrand, 2007).

Casado et al (2006) carried out a phylogeographic study of *Cheracebus lugens* based on cytochrome b DNA sequence data. They discovered that *C. lugens* at either side of the Rio

Negro were genetically distinct in term of mitDNA and consisted of a cryptic species. They were also able to time this divergence using a standard rate of sequence divergence/MA, retrieving 2 million years since separation. In terms of this widely distributed species, this is all it is known. Nothing is known to date on the genetic differences and times of divergences of populations of this species that are separated by the numerous tributaries found in this basin.

1.5 Amazon river barriers

The riverine hypothesis was first proposed by Alfred Russel Wallace in 1849, when he argued that primate distributions were affected by river barriers and showed that the river basin was divisible into 4 major geographic areas bounded by the Amazon, Negro and Madeira rivers. His theory has attracted the attention of scientists more recently such as Ayres & Clutton-Brock (1992) who observed a negative correlation between primate similarity on opposite banks and rate of river discharge in Amazonian rivers. Capparella (1987) observed genetic divergence in understory birds in relation to river width.

The riverine hypothesis postulates that the development of the Amazon river drainage due to the Miocene uplift of the Andes fragmented the ranges of once widespread species, resulting in a relevant casual process of speciation across riverbanks. It has been observed that geographic boundaries of primate species, or taxa, are coincidental with rivers (Howard & Berlocher, 1998).

There are many examples of rivers acting as natural barriers to interaction between groups, and in some cases, the separation has continued over long periods, long enough to be apparent in genetic studies. An example of this is in central Africa, where the Congo River has made a clear divide between bonobos and chimpanzees. As chimpanzees are not known

to swim, the river has effectively isolated the two groups for about 1.3 million years according to genetic analysis (Middleton, 2012).

Here we argue that major rivers or smaller tributaries are likely to drive speciation in primates of the genus *Cheracebus*. Phylogeographic patterns across taxa can be used to identify the geographic features responsible for genetic divergence, but that information is not available for large areas of the Amazon basin. However, because there are concordant phylogeographic patterns for similar taxa, a fair assessment could be made using abundant species for intraspecific divergence (Simoes et al, 2014).

It has been commonly suggested that the ranges of rain forest animals are constrained by rivers, but there have been very few attempts to suggest the hypothesis or to investigate the effects of different rivers on different species. A study carried out by Capparella (1987) based around the distributions of understory birds in Amazonia, demonstrated both that species ranges regularly coincide with river boundaries, and that within species, populations on opposite banks of the river commonly show marked differences in gene frequencies.

Although phylogeographic studies of *Cheracebus* are few (Casado et al 2005, Boubli et al 2015), studies of other primate genera have shown that rivers present significant geographic barriers, as was the case of the *Saguin us niger* population across the Rio Tocantins, studied by Vallinoto et al (2006). Results showed that analysis from the D-loop region of the mitochondrial genome displayed higher divergence between populations.

1.6 Titi monkeys

The titi monkey of subfamily *Callicebinae* is a diverse group of Neotropical primates primarily found in the tropical forests of the Amazon and Orinoco basins, but also found extending into the Atlantic forest regions of Brazil dry forests of Paraguay and Bolivia (Van Roosmalen et al, 2002). They are in the Pithecidade family which, also include uacari and

saki monkeys. Primates in this family exhibit a range of variation in terms of diet, group size, use of space and social dynamics (Campbell et al, 2011).

Titi monkeys were once included in the genus *Callithrix* E. Due to the name already being in use, Thomas (1903) proposed the name *Callicebus Thomas*, which has been in use ever since. Kobayashi et al (1995) maintained that out of three groups he separated (*Callicebus torquatus, Callicebus donacophilus* and *Callicebus modestus*) *Callicebus torquatus* had a much higher degree of character differentiation, and based on dental patterns it was concluded the *torquatus* group was in fact the earliest lineage.

Titi monkeys are generally small to medium in size, weighing around 1-2 kilograms and ranging from 270-450mm in head body length (Van Roosmalen et al, 2002). They have long reddish to black fur, with a long non-prehensile tail. Diurnal and arboreal, Titi's are very agile climbers and can jump with ease from tree to tree making them efficient foragers, especially since their diet mainly consists of fruit (Vaughan, 2011). Coat colour and pattern vary substantially within this group and this has led to the description of numerous species and sub-species (Hoyos, et al, 2016).

Table 1.1. Taxonomic clades and species of titi monkey genera following Byrne et al (2016)
The three species in bold *(C. lugens, C. purinus* and *C. lucifer)* are the three species in this
study.

Plecturocebus	*Plecturocebus*	*Plecturocebus*	*Cheracebus*	*Callicebus*
P. donacophilus	P. cupreus	P. moloch	C. torquatus	C. personatus
P. modestus	P. discolour	P. vieirai	**C. lugens**	C. melanochir
P. pallescens	P. ornatus	P. cinerascens	**C. purinus**	C. Nigrifrons
P. olallae	P. caligatus	P. hoffmannsi	**C. Lucifer**	C. barbarabrownae
P. oenanthe	P. bruneus	P. baptista	C. regulus	C. coimbrai
	P. stephennashi	P. bernhardi	C. medemi	
		P. miltoni		

Titi monkeys are frugivorous, territorial and monogamous in social structure, and because of these traits it is strongly believed that these are some of the most diverse primates to be studied (Byrne et al 2016). They are organised into small family groups consisting of one adult male, one adult female and one or two young. Each group occupy a small area exclusively, frequently threatening neighbouring groups (Wilson, 2000). Titi monkeys have a gestation period of around 5 months, the reach sexual maturity at 24-36 months, and they have a life span of around 12 years (Redmond, 2010). Fathers invest a lot of time into their offspring, and monogamy is the reason behind this. The males assume primary caretaking responsibilities including: carrying, guarding, playing and food sharing. While males invest so heavily in their young, females only interact with their offspring to offer milk, and generally avoid contact at any other time (Baer et al, 2012). Bi-parental care is usually associated with monogamy, simply because if mating is monogamous males are more likely to care for the offspring, as they are more confident of their paternity (Arnhart, 1998). The genus *Cheracebus* is the most northwesterly distributed, restricted to the Orinoco and Negro basins. Currently there are 6 known species in the genus *Cheracebus* including: *C. medemi, C. lucifer, C. lugens, C. purinus, C. regulus* and *C. torquatus*. Of these, *C. lugens* has the widest distribution including the entire span of Rio Negro left bank and upper Orinoco basin. Figure 3 depicts the range maps of all currently recognised species of *Cheracebus* as presented by the IUCN (2002) and illustrated in van Roosmalen et al (2002). Figure 4 presents an schematic representation of all important Amazonian rivers that limit the distributions of all *Cheracebus* species.

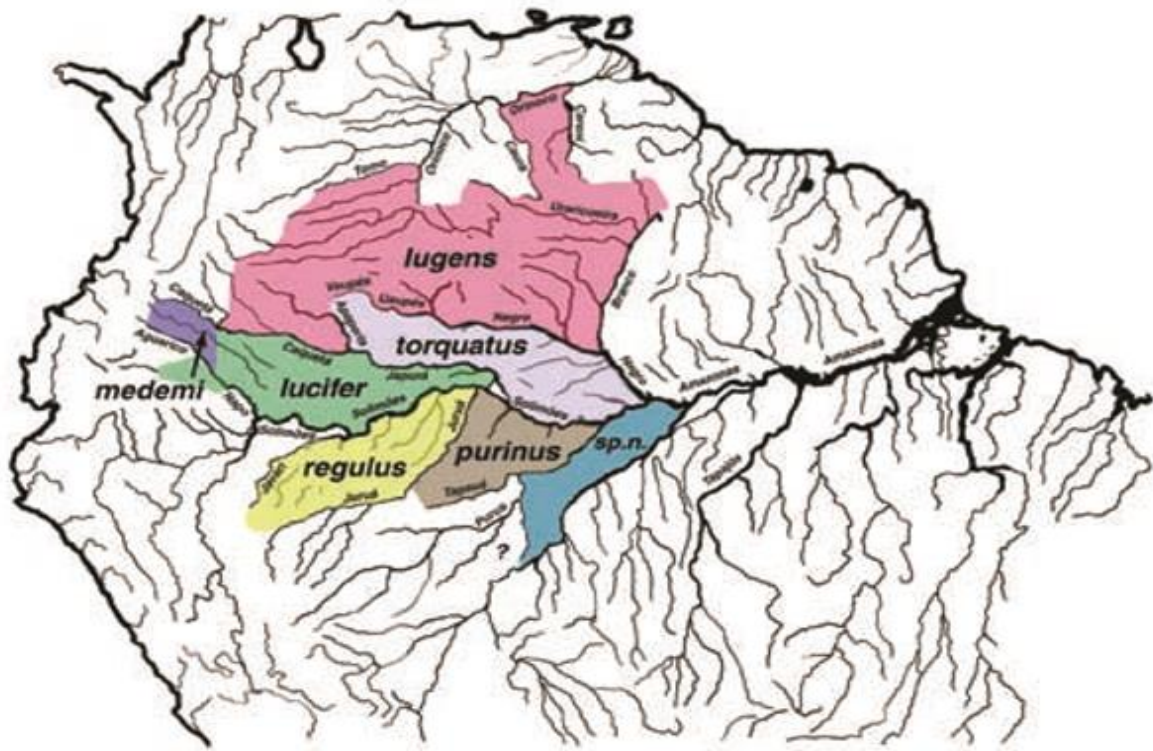**Fig 1.0.** Distributions of the known species of Amazonian titi monkeys, genus *Cheracebus,*

(taken from Van Roosmalen et al, 2002 with permission).

Fig 1.1. Schematic map of the distribution of the Amazonian titi monkeys belonging to *Cheracebus* genus depicting the important known rivers that delimit their distributions (taken from Van Roosmalen et al, 2002 with permission).

1.7 Research question and objectives

In this study, we will investigate if populations of widow titi monkeys *Cheracebus lugens* inhabiting the Rio Negro basin in Brazil present any genetic differentiation across this river's left bank tributaries, namely Cauaburi, Marauia, Padauari, Araca and Demini (Figure 5). Although Boubli et al (2015) show that the Rio Negro is responsible for the vicariant event that gave rise to two cryptic taxa of *Cheracebus lugens*, we still don't know if the Negro tributaries further split this widely-distributed species. We are also going to examine for the first time using molecular phylogenetics, the phylogenetic relationships of *C. lugens* with other species of this genera and the role of Japura and Solimoes rivers in the separation of these different species.

Figure 1.2. Map showing all the left bank tributaries of the Rio Negro that cuts across the current distribution of widow titi monkeys *Cheracebus lugens*. Red trianles represent the localites where *Cheracebus lugens* samples came from.

Methods

All samples come from a Museum collection in Brazil (INPA, Manaus) and stored in 95% ethanol before DNA extraction. All tissue samples were taken from muscle tissue of species of *Cheracebus purinus, C. lucifer* and *C. lugens*. All samples were collected from the banks of the Rio Negro, Rio Japurá and Rio Purus, all in the Amazon, Brazil, Figure 6 depicts all sample locality. Each locality might have more than one sample. Sample information including locations are listed on Table 2 below and make up a total of 20 samples.

Some of the samples were taken from used from genebank, and other samples were chosen by Hazel Byrne who also helped generate and prepare the new sequences, as well as assisting with designing the lab work.

Table 1.2: Samples chosen from genebank showing the species, sample ID and accession number.

Species	Sample ID	Primer CytB	Primer COI	Primer Rag1
C. purinus	CTGAM154	KU694254	KU694207	KU694901
C. purinus	CTGAM195	KU694255	KU694208	N/A
C. purinus	CTGAM209	KU694256	KU694209	KU694902
C. lugens	JPB81	KU694250	KU694203	KU694897
C. lugens	JPB119	KU694251	KU694204	KU694898
C. lugens	JPB124	KU694252	KU694205	KU694899
C. lugens	JPB136	KU694253	KU694206	KU694900

1.8 DNA extractions

All samples were left in a controlled area for thirty minutes for the samples to dry out from the ethanol. DNeasy blood & tissue kit was used for DNA extraction according to manufacturer protocol. Approximately 25mg of tissue was cut into small pieces and placed into a microcentrifuge tube. 20µL of proteinase was added and mixed thoroughly via vortex (5-10 seconds) then left in a water bath to incubate at 56˚C until the tissue had completely dissolved (occasionally placing the tube on the vortex to mix any undissolved tissue). 200µL

of buffer AL was added (vortexed for 5-10 seconds) as well as 200µL of 96-100% ethanol (also vortexed for 5-10 seconds).

Finally, the mixture was placed into a 2ml collection tube and placed in to the centrifuge at <6000 x 8000rpm for 1 minute. 500µL of buffer AW1 was then added to a fresh 2ml collection tube, placed into the centrifuge at <6000 x 8000rpm for a cycle of 1 minute. The last cycle consisted of 500µL buffer AW2, centrifuged for 3 minutes at 20,000 x 14,000rpm to dry ready for amplification.
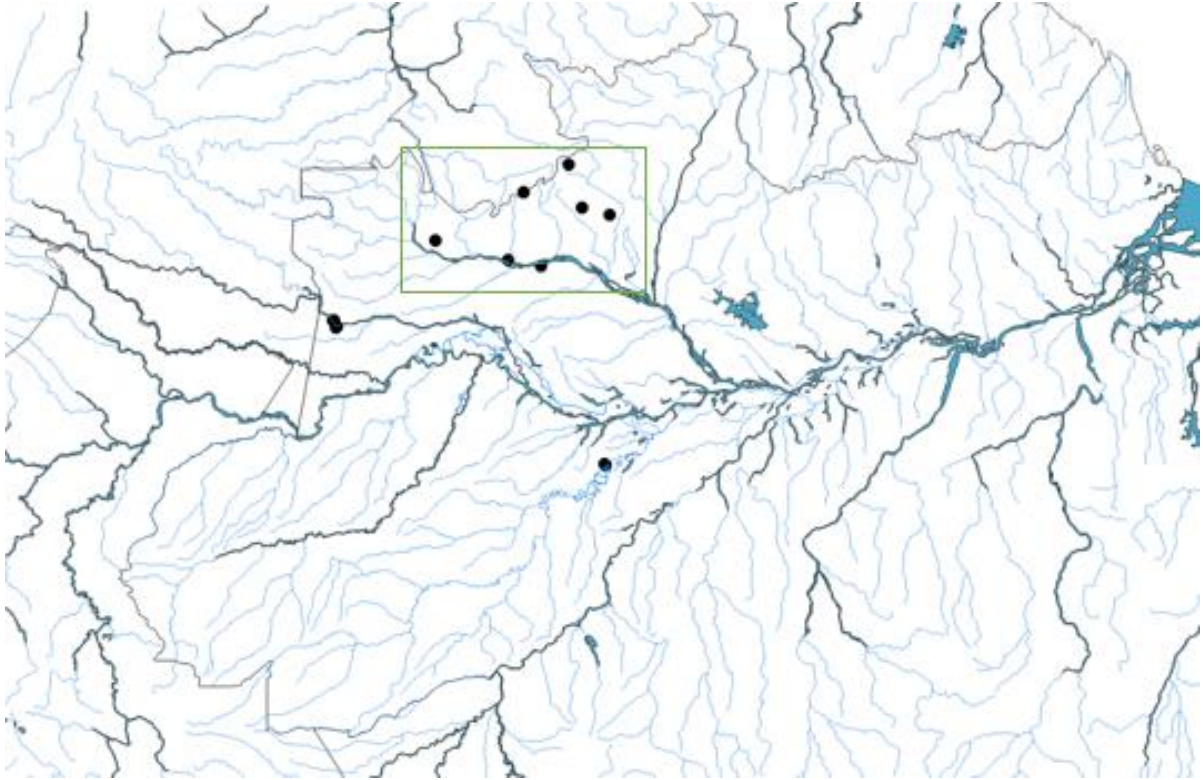
Figure 1.3. Map of Brazilian Amazon and its rivers. Black dots are localities where all 20 samples used in this study were collected. Species in the highlighted square are species of *C. lugens.*

Table 1.3. List of samples used in this study including localities marked in map of Figure 5.

| SPECIES | SAMPLE ID | SOURCE/LOCATION | COORDINATES | |
			LATITUDE	LONGITUDE
C lugens	CCM105	NA		
C lugens	JPB81	Igarapé Mandiquie, R bank of the Rio Negro, Amazonas, Brazil	-0.48	-64.41
C lugens	JPB108	Ze maria, Rio Marauia ,L bank of the Rio Negro, Amazonas, Brazil	-0.34	-65.15
C lugens	JPB111	Balawa-u, L bank of the Rio Negro, Amazonas, Brazil	1.81	-63.78
C lugens	JPB119	Marari, L bank of the Rio Negro, Amazonas, Brazil	1.19	-64.81
C lugens	JPB121	Marari, L bank of the Rio Negro, Amazonas, Brazil	1.19	-64.81
C lugens	JPB123	Igarapé Anta, Pé da Serra do Aracá, L bank of the Rio Negro, Amazonas, Brazil	0.85	-63.48
C lugens	JPB124	Igarapé Anta, Pé da Serra do Aracá, L bank of the Rio Negro, Amazonas, Brazil	0.85	-63.48
C lugens	JPB136	Igarapé Cuieiras, L bank of the Rio Negro, Amazonas, Brazil	0.69	-62.86
C lugens	JPB160	Neblina, L bank of the Rio Negro, Amazonas, Brazil	0.09	-66.82
C lugens	JPB161	Neblina, L bank of the Rio Negro, Amazonas, Brazil	0.09	-66.82
C lugens	CTGAM733	L bank of the Rio Japurá, Amazonas, Brazil	-1.71	-69.11
C lugens	CTGAM753	L bank of the Rio Japurá, Amazonas, Brazil	-1.70	-69.13
C purinus	CTGAM154	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	-4.99	-62.96
C purinus	CTGAM195	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	-4.98	-62.96
C purinus	CTGAM209	Rebio Abufari, L bank of the Rio Purus, Amazonas, Brazil	-4.97	-62.98

C lucifer	CTGAM703	R bank of the Rio Japurá, Amazonas, Brazil	-1.84	-69.03
C lucifer	CTGAM726	R bank of the Rio Japurá, Amazonas, Brazil	-1.87	-69.05
C lucifer	CTGAM727	R bank of the Rio Japurá, Amazonas, Brazil	-1.87	-69.05
C lugens	CTGAM734	L bank of the Rio Japurá, Amazonas, Brazil	-1.71	-69.11

The PCR takes advantage of an enzyme that is set to a temperature and cycle time appropriate for the gene required to isolate, in this case the cytochrome B, DLoop, CO1 and RAG1 genes. Each gene requires a different optimum temperature to run at, meaning the temperature that will give the best results. However ultimately the cycle time for the PCR remains the same for each gene. The PCR machine used was the Veriti 96 well thermal cycler. Amplicons of cytochrome oxidase 1 (CO1), cytochrome B (CytB), recombination activating gene 1 (RAG1) and D-Loop all require specific primer sequences and optimum annealing temperatures during a polymerase chain reaction. We obtained primer sequences for the chosen genes in the literature (Table 3). The reagents required for a PCR include a buffer, dNTP, primer (forward and reverse), taq and then DNA from each sample need to be added.

Every tube had 23.5 µL of master mix, meaning every component of PCR grade H2O, buffer, DNTP, Primer (forward and reverse), ExTaq Takara was mixed prior and then 1.5µL of DNA was added to the master mix. A negative control was added in to each run, consisting of 23.5µL of master mix. This is to ensure any contamination can be detected quickly and the run be re-done if necessary. The master mix comprised of 17.875µL H2O, 2.5µL buffer, 2.0µL deoxynucleoside triphosphate (DNTP), 0.5µL forward primer, 0.5µL reverse primer and 0.125µL ExTaq Takara.

Table 1.4. The gene, forward and reverse primers used in this study, along with the temperature and reference.

Gene	Forward	Reverse	Temperature	Reference
CO1	5' TCCATTACCAGGCCAGCTAG 3'	5' GAACTTGCTGGCTTTCATATC 3'	45 °C	Ward et al, 2005
D-Loop	5' CTACCATCAACACCCAAAG 3'	5' CATCCAGTGACGCGGTTAAGA '3	60°C	Oliveira et al., 2011
RAG1	5' GCTTTGATGGACATGGAAGAAGACAT 3'	5' GAGCCATCCCTCTCAATAATTTCAGG 3'	57°C	Teeling et al., 2001
CytB	GCACAACCTACAGCACCACTA	CAGCTTTGGGTGTTGAGGTRGAA	60°C	Byrne et al., 2016

Finally, once all tubes contained 25µL (23.5µL of master mix and 1.5µL of DNA) I then vortexed to properly mix the samples and placed in to the PCR machine adjusting the settings depending on the gene to.

The optimised PCR cycling profiles were set at 95˚C for 5 minutes, 94˚C for 5 minutes, then the temperature altered for a cycle of 1 minute depending on which gene was running, followed by 72˚C for 1 minute and finally another cycle of 72˚C. CO1 required a temperature of 45˚C, CytB a temperature of 60˚C, RAG1 a temperature of 57˚C and finally D-Loop ran at 60˚C.

2.0 Agarose gel electrophoresis

DNA loading buffers are dense, allowing the DNA sample to sink in to the gel well and provide a visible marker of DNA migration (Nadeau, 2016). 3µL of loading buffer was loaded on to a film surface as a small dot, followed by 3µL of DNA mixed straight into the buffer. The amount of buffer was dependent on how many samples of DNA were used. Before DNA could be loaded into the gel, then 3µL of ladder was added. The ladder simply contains a marker dye which would enable me to the results as the DNA migrates through the gel. Before the gel tray containing the applied samples was placed into the electrophoresis apparatus, approximately 180ml of buffer was added. Making sure the gel was completely submerged and in place, a voltage of 70 watts and a time of 60 minutes was set for accurate results and good migration. Once the gel had completed the cycle of 70 watts and 60 minutes GeneSnap was used, which is a piece of software designed to enhance images in gels via UV light (Figure 7).
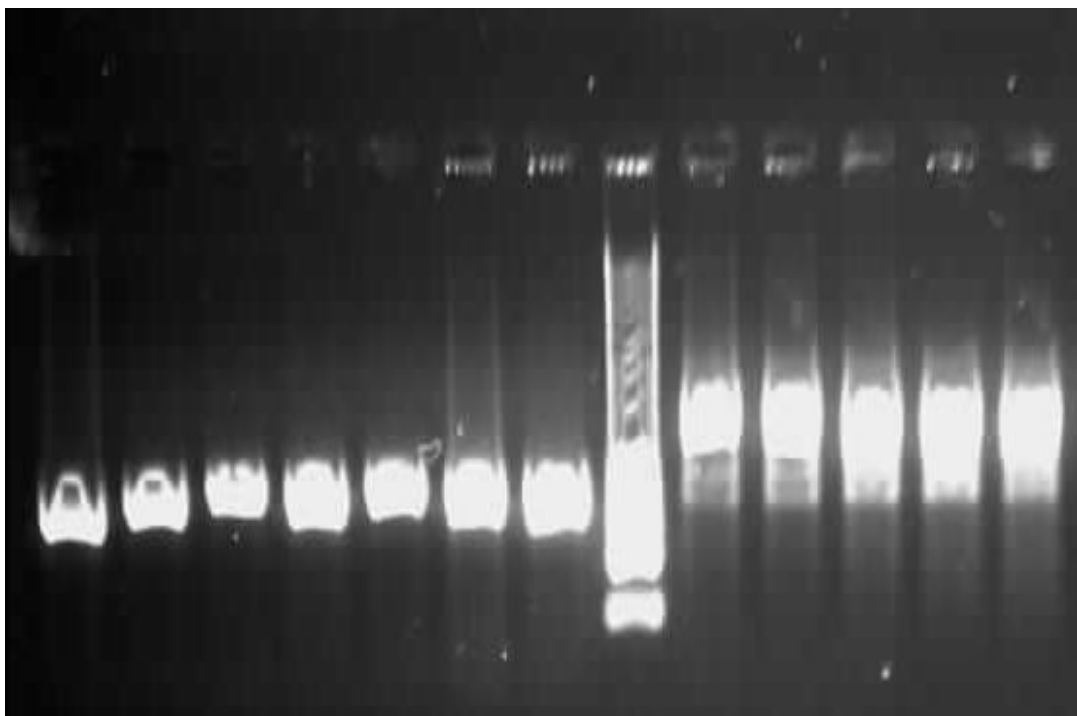
Figure 1.4. Agarose gel showing the results of a run with PCR products for 12 *Cheracebus lugens* samples

Data analysis

Neighbour joining method

Neighbour joining is a clustering algorithm that optimizes the so called "balanced minimum evolution" (Kurtzman et al, 2011) however the neighbour joining method proposed by Naruya Saitou and Masatoshi Nei does not assume all sequences have the same rate constant rate of evolution over time (Zvelebil & Baum, 2008). This method has become the distance approach of choice for many types of molecular data as it can incorporate different rates of evolution in different lineages (Hedrick, 2005).

Maximum likelihood

The statistical principle of maximum likelihood is to estimate unknown parameters of the probability of a model in such a way that it can observe data gain maximum probability. The computer basis was then laid by Felsenstein (1981) where the observed data are given by the sequence alignment and the unknown parameters of the probability model compromise the unknown true phylogeny of the given taxa. It is shown that maximum likelihood is a consistent method of phylogenetic reconstruction (Kurtzman et al, 2011).

2.1 Cutting sequences

ProSeq

ProSeq is a programme used for sequence editing and molecular analysis. It is designed to simplify preparation and analysis of DNA sequence data sets in population genetic, phylogenetic and molecular ecology studies. Among its features, it allows to calculate genetic diversity, divergence, population subdivision, gene flow and sequence alignment (Filatov, 2002). This software enabled me to align each of my sequences.

2.2 Creating haplotypes

A haplotype network is an unrooted phylogenetic network in which the nodes represent different haplotypes within a (usually closely related) group of taxa and the edges join those sequences or haplotypes that are very similar (Hudson & Scornavacca, 2010).

To create a haplotype network there are several steps, and several different types of software must be used to produce such diagrams. To create my haplotype networks, Clustal, MEGA (Molecular Evolutionary Genetics Analysis) Fluxus-engineering and DnaSP (DNA sequence polymorphism) were used.

Clustal:

Sequence alignment is an essential step in molecular analysis of sequences as it identifies homologous sites despite insertions and deletions (Paradis, 2011). Because of this it is extremely important that any software used is accurate and efficient, thus providing the most reliable results.

A sufficiently fast method, Clustal is a programme used widely in molecular biology for multiple alignment of both nucleic acid and protein sequences, as well as the ability to generate accurate phylogenetic trees (Chenna et al, 2003).

Using seeded guide trees and profile techniques to generate alignments between three or more sequences (EMBL-EBI, 2016) Clustal is a programme that can be used to align existing profiles to each other, or add new sequences to an existing alignment (Misener & Krawetz, 1999).

DnaSP

The analysis of DNA sequence polymorphisms can provide great insight into the evolutionary forces acting on populations and species. Available population genetics methods have become the framework to analyse such DNA polymorphism data (Rozas, 2009).

DnaSP, otherwise known as DNA Sequence Polymorphism, is a software package designed for the analysis of nucleotide polymorphism from aligned DNA sequence data. DnaSP can estimate several measures of DNA sequence variation within and between populations, as well as gene flow and gene conversion parameters (Rozas, 2010). If more than one nucleotide difference is present between two codons (codon meaning nucleotides that form a genetic code in a DNA molecule) DnaSP considers all possible pathways of substitution with equal probability, deleting those that lead to stop codons (Munoz-Torres, 2009).

MEGA

Molecular Evolutionary Genetics Analysis, (MEGA) is a software that projects its focus on facilitating the exploration and analysis of DNA and protein sequence variation from an evolutionary perspective. The comparative DNA and protein sequence analysis plays a central role in reconstructing the evolutionary histories of species and multigene families, estimating rates of molecular evolution and inferring the extent of selective forces shaping the evolution of genes and genomes (Kumar, 2004).

A sophisticated piece of software used for analysing DNA and protein sequence data from species and populations, with several features including multiple sequence alignment, substitution model selection, evolutionary distance estimation, phylogeny inference, substitution rate and pattern estimation, tests of natural selection and ancestral sequence inference, which is all essential for molecular analysis (Kumar, 2012).

Maximum likelihood selects the most likely phylogeny by assigning a maximum likelihood value to each character arrangement within the sequences being studied, and selects the tree topology with the highest maximum likelihood value. Specifically, phylogenetic trees generated under different models are compared. A likelihood value is generated for each phylogenetic tree, and a search using these values determines the most likely tree topology under each model (Pometto et al, 2005).

MrBayes

MrBayes is a widely-used programme for phylogenetic inference. It uses Bayes theorem to estimate the probability of a phylogenetic tree, which is called Bayesian inference of phylogeny (Yeo et al, 2010). MrBayes combines information from different data partitions or subsets evolving under different stochastic evolutionary models. This allows the user to analyse different data types such as morphological, nucleotide and protein, as well as exploring a wide variety of structured models (Ronquist & Heulsenbeck, 2003).

Fluxus engineering

Fluxus engineering allows the user to conduct DNA alignment and editing, generating evolutionary trees from genetic and linguistic data. From this fluxus can provide age estimates for ancestors in a tree (http://www.fluxus-engineering.com/sharenet.htm).

Results

A total of 20 individuals were sampled, three *Cheracebus lucifer,* three *C. purinus* and 14 *C. lugens*. haplotype networks were obtained using MEGA, clustal, DnaSP and fluxus engineering. By using each piece of software, a haplotype network was formed for each gene sequenced for this study, i.e., COI, cytochrome B, D-Loop and RAG1. Cytochrome B and D-Loop were ran as individual species and not as all three-species combined, as we found too many mutations between species obscuring the overall results for *C. lugens*, which was the focal species in this study. COI and RAG1 were ran with all three-species combined since numbers of mutations was small enough even between the three different species. Colour coding was used to identify each species and the number of mutations were placed along the lines of the haplotype networks. Figure 7 shows each haplotype network. Tables 4 to 6 present the summary results from the haplotype network analyses.

In the mitochondrial gene COI, there are five *C. lugens* haplotypes seven *C. lugens* in haplotype 4, three *C. lugens* in haplotype six and the rest of the individuals are single haplotypes). Three *C. lucifer* all individual haplotypes, and one single haplotype for *C. purinus,* with all three individuals in the same haplotype. There are 60 mutations for *C. lugens*, 228 mutations for *C. lucifer* and no mutations for *C. purinus*.

In the mitochondrial gene cytochrome B, there were eight *C. lugens* haplotypes (two individuals in haplotype 1, three in haplotype three and two in haplotype seven) three *C. lucifer* in indivudal haplotypes and one *C. purinus* haplotype containing all three individuals. There are 39 mutations in total for this gene.

In the mitochondrial gene D-Loop, there were three *C. purinus* in individual haplotypes, one *C. lucifer* haplotype containing all three individuals, and 11 *C. lugens*

haplotypes, with three individuals in one haplotype, but the rest are in single haplotypes. There are ten mutations for *C. purinus*, 76 for *C. lugens* and 0 for *C. lucifer*.

In the nuclear gene RAG1, there were four *C. lugens* haplotypes (one individual in one haplotype, five in haplotype two, two in haplotype five and three in haplotype seven. Two *C. lucifer* in haplotype seven, and two *C. purinus* in haplotype two. There are 9 mutations in total for this gene.

Summary statistics were generated in Arlequin, and each table was made per species (*C. lugens*, *C. lucifer* and *C. purinus)* calculating each value per gene (COI, cytochrome B, D-Loop, RAG1)*.* Arlequin used Tajima D's and Fu's F to generate haplotype diversity and nucleotide diversity, which determined whether populations were stable, as well as stating if species populations showed signs of expansion.

The Tajima's D value can be regarded as a control measure of severity of demographic effects. If Tajima's D is a positive number, it could imply a long-term reduction of the population size. If it is negative, it could imply a long-term expansion of population size or recent incorporation of genetically differentiated minority (Hirohisha et al, 2012).

Nucleotide diversity is the average number of nucleotide differences per nucleotide site between sequences (Hirohisha et al, 2012) and it is used used to measure polymorphisms in a population. Haplotype diversity is defined as measure of the uniqueness of a haplotype in each population (Jobling, 2016). Molecular data from the mitochondrial genome gives important insight about the genetic uniqueness and genetic diversity of population fragments. Data obtained from a mitochondrial genome differs from the nuclear genome, as mitochondria are maternally inherited (Marsh & chapman, 2013).

Table four is specifically for species *C. lugens*, Table five *C.lucifer* and finally Table six

shows *C.purinus*. The same values were used for each species, e.g. each neutrality test was

Tajima's D and Fu's F, the number of samples was set at 10000, bootstrap was set at 1000,

standard diversity, molecular diversity indices and Tamura Nei was always used.

2.4 Phylogenetic analyses

Using a likelihood function to create trees using a model of evolution, MrBayes was

used to create the most likely phylogenetic trees that show the estimation of phylogeny,

MEGA was used to create maximum likelihood to estimate parameters of probability, and

neighbour joining trees to estimate balanced minimum evolution. Figures 8, 9, 10, 11 and 12

show these phylogenetic trees with all three criteria (Bayesian inference, maximum

likelihood and Neighbour Joining). All species are colour coded with the same colours from

the haplotype networks (*C. lugens* orange, *C. Lucifer* green and *C. purinus* blue) and *Cacajao*

calvus used as the outgroup which remains black. Using all three species for each figure,

figure 8 shows the COI gene, Figure 9 shows the cytochrome B gene, figure 10 shows the D-

Loop gene, Figure 11 shows the RAG1 gene and finally figure 12 shows concatenated data.

Table 7 shows models of evolution used to generate phylogenetic trees on MrBayes.

The software J-Model test was used to carry out statistical selection of best fit models of

nucleotide substitution (Hancock & Zvelebil, 2014) resulting in a model of evolution for each
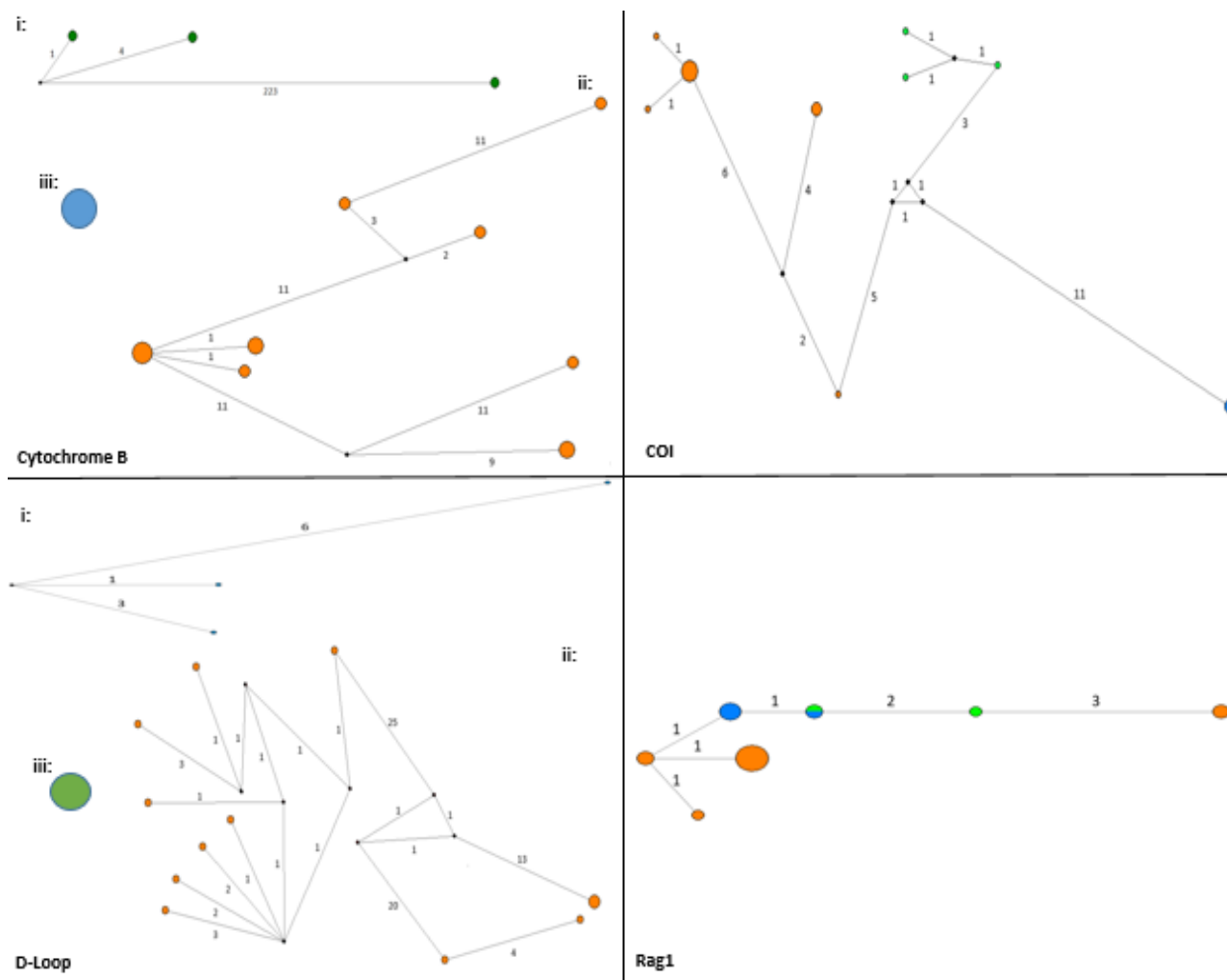
individual gene.

Figure 1.5. Haplotype networks for the four genes analysed here, COI, *Cytochrome b*, RAG1 and D-Loop for specimens of *Cheracebus lucifer* (green), *C. lugens* (orange) and *C. purinus* (blue). For cytochrome b and D-Loop, networks were produced separately for each species due to the great sequence divergences in these cases. Numbers on top of lines represent the numbers of mutations and the sizes of coloured circles reflect the number of samples sharing the same haplotype.

Table 1.5. Summary statistics for *C. lugens*.

Marker	No. of samples	Haplotypes	Haplotype diversity +/- (SD)	Nucleotide diversity +/-	Polymorphic sites	Tajima D	Fu's F	No. of loci
Rag 1	12	6	0.8182 (0.0957)	0 (0)	0	0	∞	853
Dloop	13	11	0.9615 (0.0496)	0.046130 (0.024284)	68	1.2613	0.49205	626
Cytb	12	8	0.9242 (0.0575)	0.0096 (0.0053)	58	0.77855	2.16240	1089
CO1	14	6	0.7363 (0.1092)	0.008615 (0.005050)	13	0.42132	1.22229	535

Table 1.6. Summary statistics for *C. lucifer*.

Marker	No. of samples	haplotypes	Haplotype diversity +/-	Nucleotide diversity +/-	Polymorphic sites	Tajima D	Fu's F	No. of loci
Rag 1	3	3	1.0000 (0.2722)	0	0	0	∞	853
Dloop	2	2	1.0000 (0.5000)	0.028179 (0.028973)	17	0	2.83321	626
Cytb	3	3	1.0000 (0.2722)	0.004226 (0.003537)	223	0	3.89862	1089
CO1	3	3	1.0000 (0.2722)	0.003757 (0.003539)	3	0	-0.69315	535

Table 1.7. Summary statistics for *C. purinus*.

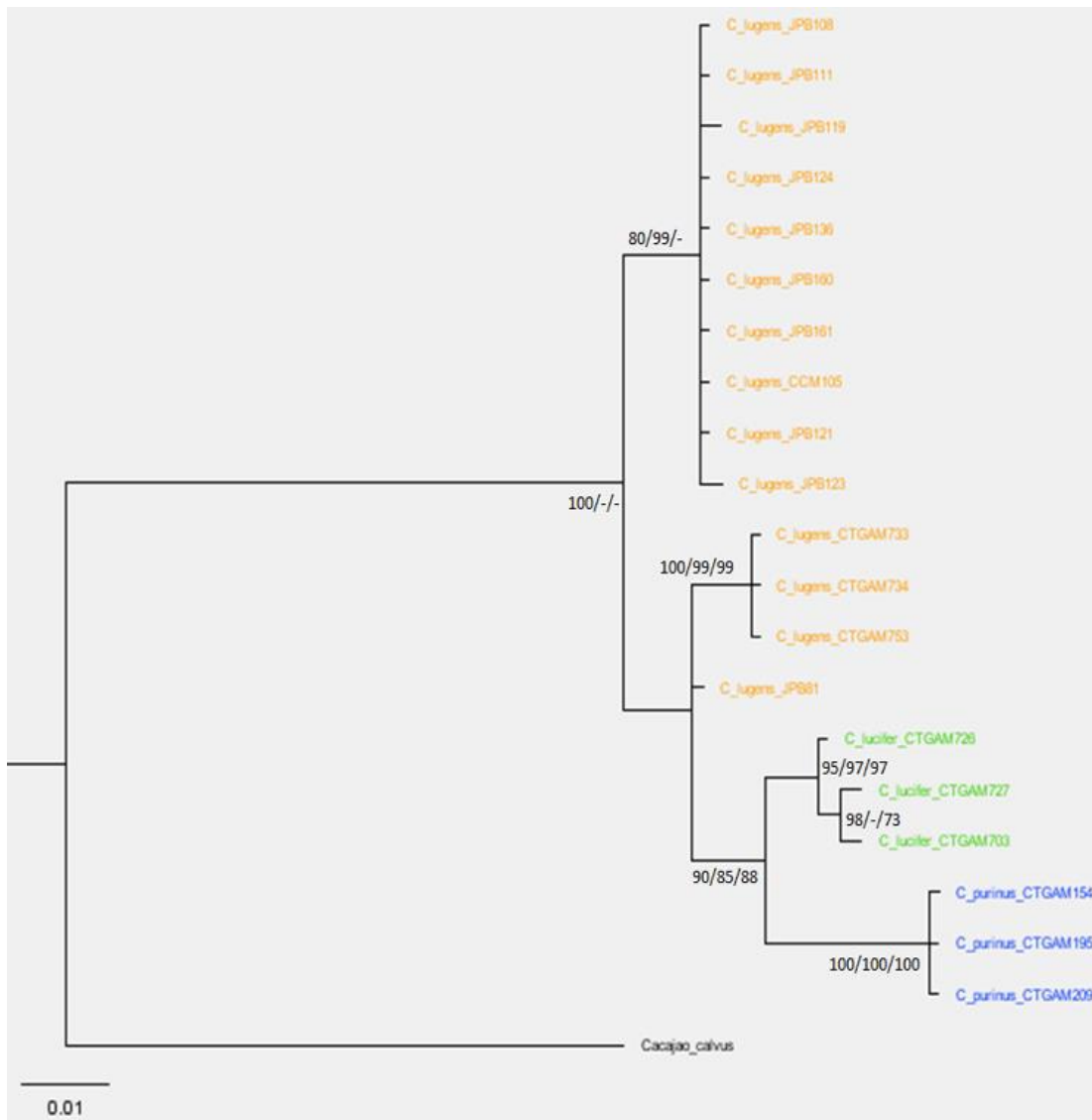| Marker	No. of samples	Haplotypes	Haplotype diversity +/-	Nucleotide diversity +/-	Polymorphic sites	Tajima D	Fu's F	No. of loci
Rag 1	3	1	0	0	0	0	0	853
Dloop	3	3	1.0000 (0.2722)	0.012968 (0.010357)	12	0	0.90079	6269
Cytb	3	2	0.6667 (0.3143)	0.001853 (0.001745)	3	0	1.60944	1089
CO1	3	1	0	0	0	0	0	535

Figure 1.6: A phylogenetic tree for samples representing three species of titi monkeys, *Cheracebus lugens, C. lucifer* and *C. purinus* using the COI gene, with values from MrBayes/Maximum likelihood/neighbourhood joining methods. The white uakari, *Cacajao calvus* is used as an outgroup.

Figure 1.7: A phylogenetic tree for samples representing three species of titi monkeys, *Cheracebus lugens, C. lucifer* and *C. purinus* using the cytochrome b gene, with values from MrBayes/Maximum likelihood/neighbourhood joining methods. The white uakari, *Cacajao calvus* is used as an outgroup.
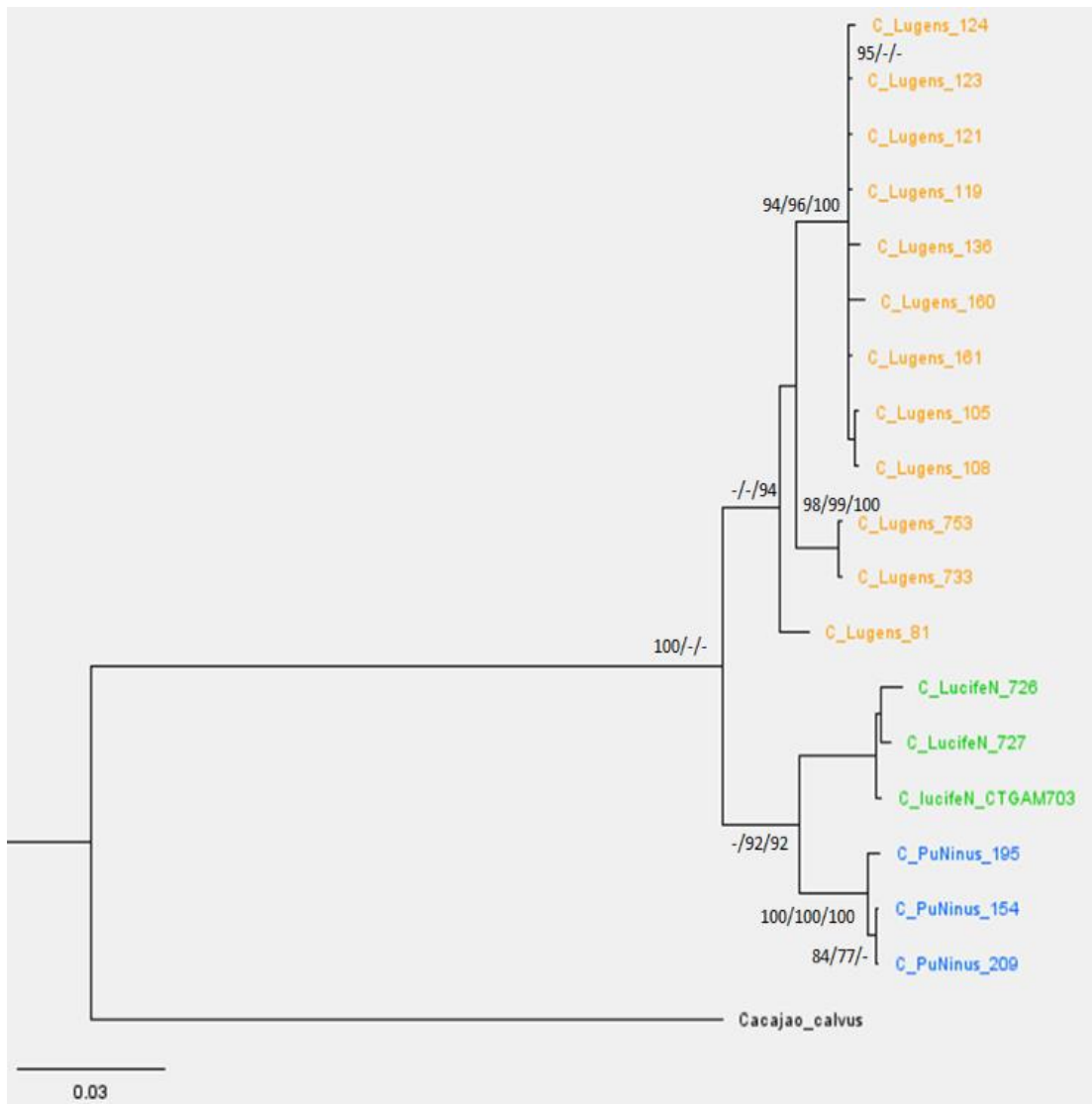
Figure 1.8: A phylogenetic tree for samples representing three species of titi monkeys, *Cheracebus lugens, C. lucifer* and *C. purinus* using the D-Loop gene, with values from MrBayes/Maximum likelihood/neighbourhood joining methods. The white uakari, *Cacajao calvus* is used as an outgroup.
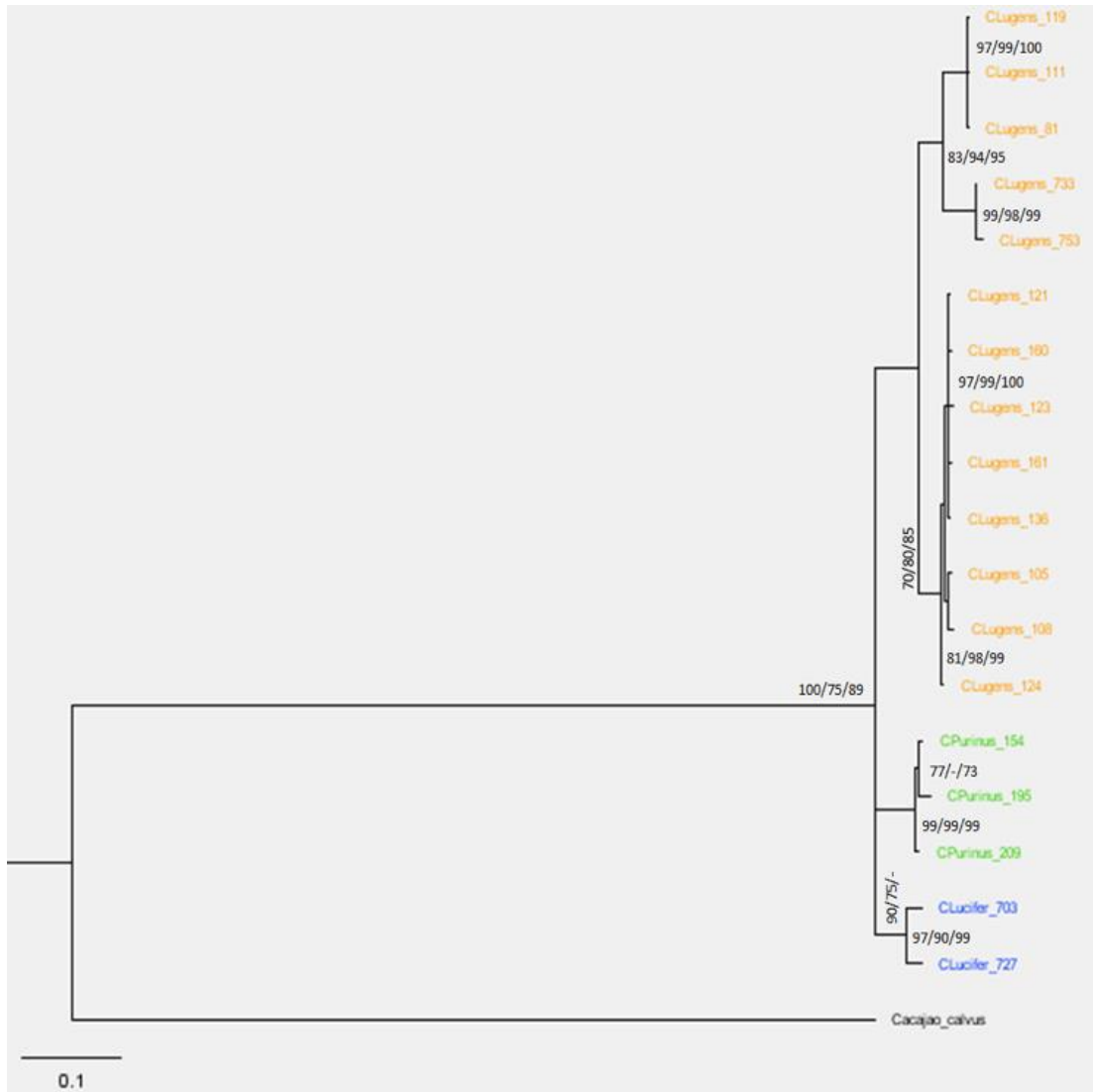
Figure 1.9: A phylogenetic tree for samples representing three species of titi monkeys, *Cheracebus lugens, C. lucifer* and *C. purinus* using the RAG1 nuclear gene, with values from MrBayes/Maximum likelihood/neighbourhood joining methods. The white uakari, *Cacajao calvus* is used as an outgroup.
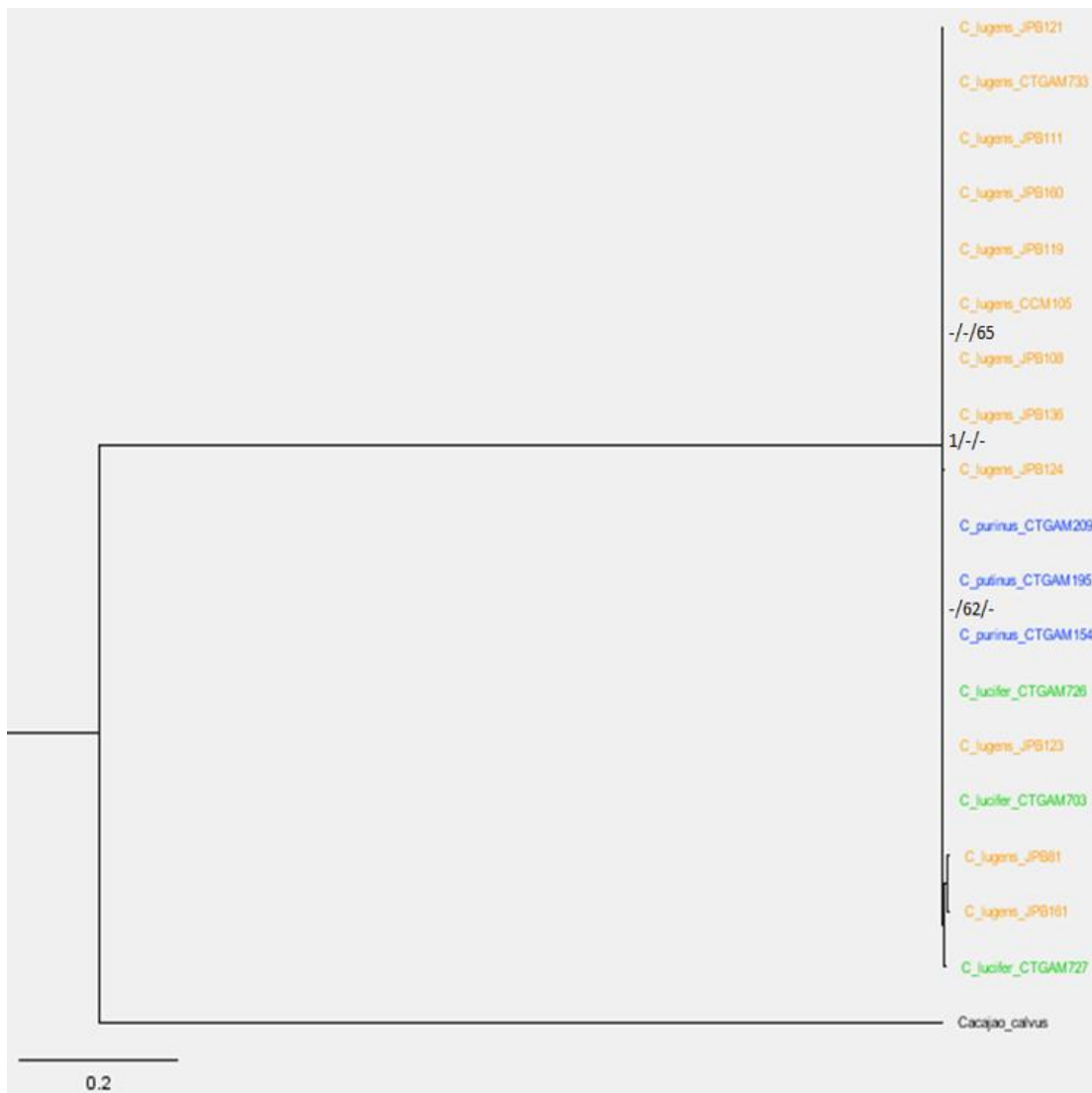
Figure 2.0: A phylogenetic tree for samples representing three species of titi monkeys, *Cheracebus lugens, C. lucifer* and *C. purinus* using the concatenated data, with values from MrBayes/Maximum likelihood/neighbourhood joining methods. The white uakari, *Cacajao calvus* is used as an outgroup.

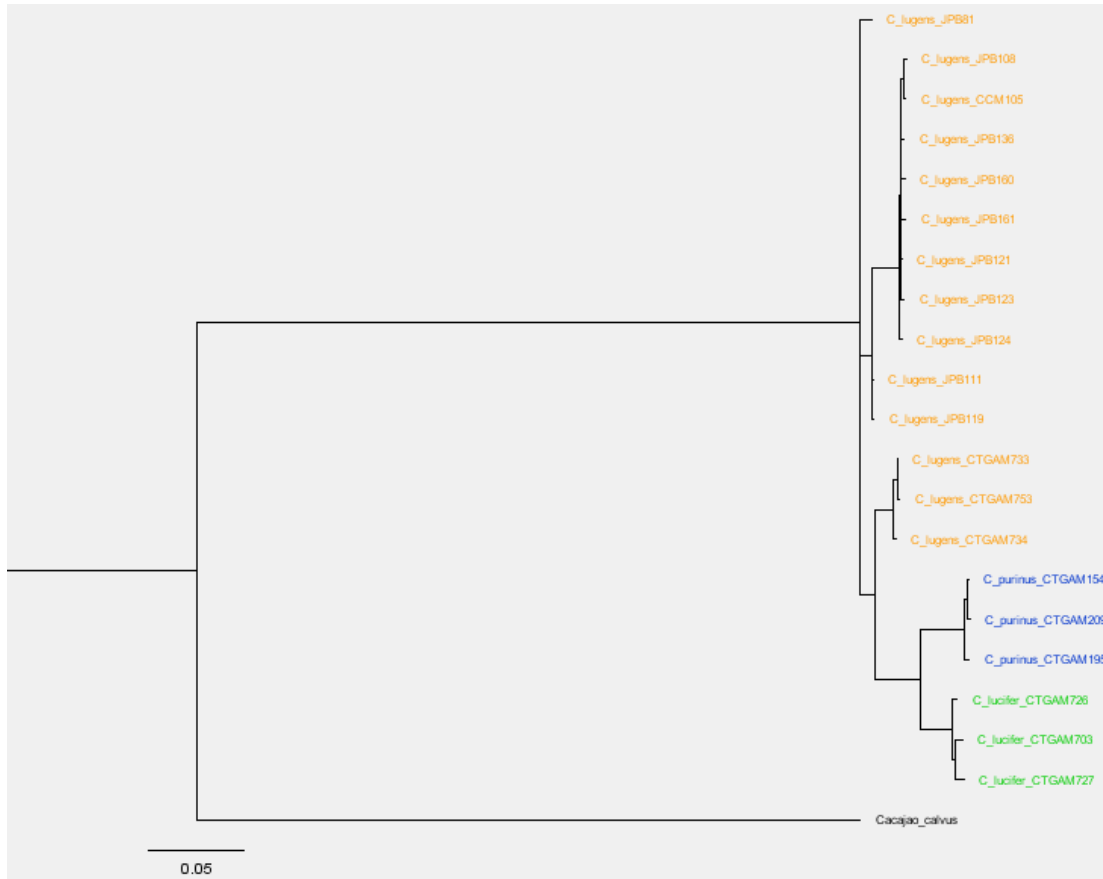**Table 1.8.** Models of evolution for each individual gene. The models of evolution were obtained from jModelTest.

Gene	Model of evolution
CO1	HKY+I
Cytochrome B	HKY+G
D-Loop	HKY+G
Rag1	JC+I

Discussion

The present study is new research that focuses on how rivers act as genetic barriers to different species of *Cheracebus,* as well as different populations of *C. lugens* in particular. Our results showed that populations of *C. lugens* inhabiting the left bank of the Rio Negro showed little genetic differentiation across this river's left bank tributaries namely Cauaburi, Marauia, Padauari, Araca and Demeni. The study showed divergence times for closely related species of *Cheracebus* coincided with the formation of the Rio Negro, offering strong support to the riverine hypothesis for large Amazonian rivers.

Data suggests that none of the rivers on the left bank of the Rio Negro appears to be a barrier to gene flow between species of *C. lugens*. As expected however, both the Japura and Solimoes showed compelling evidence that shows these rivers are important barriers separating *C. lugens*, *C. lucifer* and *C. purinus*. In support of the findings, Casado et al (2006) discovered that the *C. lugens* from different banks of the Rio Negro were genetically divergent between right bank haplotypes and left bank haplotypes although karyotypically they were identical. However Distance estimates between left and right banks showed these two groups represent different evolutionary lineages. Their findings, and our own, point to the Rio Negro as a natural geographic barrier to gene flow between *C. lugens*. Boubli et al (2015) noticed although morphologically similar, *C. lugens* samples from opposite banks were genetically distinct. Divergence times estimate that *C. lugens* of different Rio Negro margins split approximately 2 mya, a time that is compatible with the river dynamics of the Amazon region, indicating the Rio Amazonas originated approximately 2-3 mya. Divergence may be explained the riverine hypothesis (Casado et al, 2006) which is what a large proportion of our research is based upon.

As explained in the results section of this thesis the phylogenetic trees containing Mrbayes, maximum likelihood and neighbour joining are all set out per gene, each containing all 20 of the 3 species of *Cheracebus* as well as the outgroup species, *Cacajao calvus*.

When conducting the phylogenetic trees, it was noted that *C. purinus* and *C. lucifer* were more closely related to each other than they were to *C. lugens*. A divergence time of 1.63 mya between *C. lugens* and *C. purinus* was discovered, and it is thought the divergence between the these two titi monkey species resulted after the formation of the river Amazonas (Boubli et al, 2015). The same goes for *C. Lucifer*, *C. lugens* is by far the biggest sample size therefore it is possible to learn the most information from this group based on this.

The mitochondrial gene COI, cytochrome B and D-Loop shows very steady rates of divergence, and correlates well with the positioning of the titi monkey locations (see figure 6). We can see as expected all three species have remained together in their own groups and no species are mixed. Uniquely however we can see sample *C. lugens* JPB81 has branched off slightly by itself, away from other species of *C. lugens*. Potential reasons for this are formation of the river Japura which has separated groups of *C. lugens*. *C. lugens* and *C.lucifer* are more closely related than *C. lugens* and C. *purinus* which is unexpected, this means *C. lugens* is paraphyletic. Each branch has very strong values of MrBayes, Maximum likelihood and neighbourhood joining, none of which fall below 70%. If any were to fall below that the values were left out. This indicates the predicted phylogenetic relationships and divergence accuracy are high.

The nuclear gene RAG1 is not clear in terms of divergence, and there only appears to be three divergences (*C. lugens* JPB161, *C. lugens* JPB81 and *C. lucifer* CTGAM727). Unlike the mitochondrial genes, individuals of the same species are no longer grouped together, meaning more genetic differences and closer relationships between unexpected individuals in this particular gene. The values are all relatively low, and there are several values missing all together as they were too low to include.

For the mitochondrial gene cytochrome B, I notice it is quite similar to the COI gene in the fact that groups of the same species are all grouped together, with *C. lugens* JPB81 again branching off on its own (meaning it is paraphyletic). *C. lucifer* is closest related to *C. purinus*, followed by *C. lugens* nearest to *C. lucifer*, and finally *C. lugens* is the furthest related genetically to *C. purinus.* All values are strong, showing 77% plus probability (values below 70 were not included).

For the mitochondrial gene D-Loop, all groups of species are together, however it is noted *C. lugens* has branched leaving what appears to be three different groups. For every other gene, *C. lugens* is by itself, however in this gene this individual is grouped with two other individuals. Unlike the gene cytochrome B, *C. lugens* is more closely related to *C. purinus* and further related to *C. lucifer*. Genetic divergence remains the same between *C. lucifer* and *C. purinus*. Again, values are strong ranging from 70% plus, and values below this were left out.

Many studies have been conducted over time based on finding out relationships between different species and their divergence. Opazo (2006) conducted a study on different platyrrhine families (Atelidae, Cebidae and Pitheciidae) and found in his results the two most closely related families were Atelidae and Cebidae. Another study conducted by Schrago (2006) based on the divergence of new world primates from old world primates, discovered platyrrhines diverged from old world arthropods approximately 35mya. He also discovered living new world primates are descendants of ancestors that lived in the early Miocene period around 20 mya. He concludes the divergence times for primate families (Cebidae 16.9mya, Pitheciidae 15.6 mya and Atelidae 12.4 mya) may be a consequence of environmental changes.

Through evolutionary history of a species, individuals diverge by accumulating mutations in their DNA called polymorphisms, and the sequence containing all polymorphisms in an individual is called a haplotype. These mutations can be homoplasic, meaning that they occurred more than once during evolution (Rocha et al, 2012).

The average nucleotide per species for all genes were similar and very low. *Cheracebus lugens* had an average of 0.05 +/- 0.03, *C. lucifer* had an average of 0.03 +/- 0.03 and *Callicebus purinus* had an average of 0.01 +/- 0.01. This means that this measurement, which is the degree of polymorphisms within these species is low.

Tajima's D averages are as follows; *C. lugens* had an average of 2.14, *C. lucifer* had an average of 0, and finally *C. purinus* had an average of 0. As all species have a very low average (especially with *C. lucifer* and *C. purinus* having a value of 0) but all remain positive, it suggests a long-term reduction of the population size. This leads me to believe the reason for this is a conservation issue, possibly fragmentation causing separation of individuals of the same species.

Literature based around south American new world primates showed Garber et al (2008) focused on male and female genetic diversity of woolly monkeys. They discovered that haplotype divergence overall for both male and female was 3.4 +/- 1.5%, a high diversity meaning the population size of this troop of woolly monkey are at a healthy level. In comparison, the 3 species of *Cheracebus* used for my study are at a slightly lower level. Average Diversity for *C. lugens* is 2.88 +/- 0.23, for *C. lucifer* average diversity is 3.25 +/- 1.11 and finally the average diversity for *C purinus* is 1.6 +/- 0.58 across all genes. *C. purinus* was a relatively small sample size, while *C. lugens* was a much larger sample size which could be an explanation for why the diversity is slightly more for *C. lugens*. Another explanation could be that the *C. lugens* have a much more stable population showing very

slight expansion, and *C. purinus* is showing little signs of expansion, show a less stable population.

Collins & Dubach (2000) conducted a study on three genetic regions in an analysis of spider monkeys. Sequencing DNA in the cytochrome c gene and COII gene they could determine relationships. They recognised *Ateles belzebuth* as having high genetic diversity amongst geographically widespread haplotypes including former species. They concluded however although there was high diversity, they noticed no significant gene flow barriers between these populations, and morphological analysis supported their findings. They also stated it was not possible to define a geographic boundary between variable haplotypes within the group even though genetic variation was high (Campbell, 2008).

Four genes, Cytochrome B, D-Loop, RAG1 and CO1. RAG1 and CO1 have each species combined into one figure, however Cytochrome B and D-Loop show each individual species. *C. lucifer* in D-Loop and *C. purinus* in Cytochrome B both show one individual haplotype. This is because we were restricted by the number of samples we could use. As we had restricted DNA samples from both *C. purinus* and *C. lucifer*, it only showed one haplotype. The sample size can make a difference in population and if our sample size was bigger, we would have shown more results. As sample size was so small in this case for *C. purinus* and *C. lucifer* it was difficult to show population expansion with the results only showing one haplotype per species.

C. purinus has by far the most mutations with 223. Explanations for this could be possible contamination, however this is the only gene and species that appears to be affected, all other genes are relatively close together and have a lot less mutations. Another reason may be *C. purinus* being highly variable, or perhaps because we only have 3 individual species.

2.5 Conservation status

Titi monkeys have a very good predator avoidance strategy, with potential threats being diurnal birds of prey, felids, snakes and capuchin monkeys. However, predators are not the only threat to titi monkeys, with habitat destruction hunting and the pet trade being a high contributing factor to a larger number of titi monkey classified as critically endangered and endangered (Veiga et al, 2013). Amazonian primates including all species of *Callicebus* are hunted for a variety of reasons such as for food, traditional medicines, trophies and ornamentation. Although such hunting is prohibited, enforcement of any protection is rare and sometimes non-existent (Nowak, 1999).

As well as the threat of hunting, genetic consequences of habitat fragmentation depend on the level of migration or gene flow occurring between the population fragments. Not surprisingly, gene flow is affected by the number, size and patterns of fragments, as well as the distance between fragments and the time since fragmentation. If there is restricted gene flow, fragmentation usually leads to loss of genetic diversity within fragments and greater inbreeding. In these fragments, greater genetic differentiation can lead to risk of extinction. Loss of genetic diversity due to genetic drift or gene flow between fragments can lead to negative consequences for disease resistance and ultimately fitness. Studies of the genes in light of fragmentation is relatively new, but may become increasingly important in the future of primate conservation research (Marsh & Chapman, 2013).

2.6 Conservation

Conservation of biodiversity in the Amazon faces many challenges, including the large size of the region, poor taxonomic knowledge and the rapid pace of regional development and environmental change. The basin covers approximately six million km^2, with predictions that 40-55% of its rainforest will be lost within the next few decades (Simoes et al, 2014). The

conservation of river systems is considered high priority, with the region containing 13 major river sytems: Amazonas, Madeira, Tocantins, Negro, Xingu, Tapajós, Purus, Marañon, Ucayali, Japurá, Juruá, Putumayo, Trombetas and Napo rivers (Simoes et al, 2014).

The utility of information on genetic diversity has led to a growing interest within the field of conservation biology in understanding the factors driving variation in intraspecific genetic diversity across space and time. Spatial variation in genetic diversity is often driven by the presence of barriers that restrict gene flow between groups, therefore leading to the accumulation of genetic differences between groups. A variety of factors can act as barriers to gene flow, including geographic influences such as mountain ranges or oceans, or ecological influences such as habitat differences or behavioural factors (Yamagiwa & Karczmarski, 2013).

2.7 Conclusion

In conclusion, we conducted our research over a period of a year in which we discovered results that supported our hypotheses, and discovered results we were not expecting. The aim was to identify whether or not a variety of Amazonian rivers acted as genetic barriers between individuals of the same species and individuals of different species, and I believe that the hypotheses has been supported.

We have shown that large Amazonian rivers do in fact act as genetic barriers across species of *Cheracebus,* and we even found slight genetic differences in species assumed to be the same due to morphology. I think it is important to remember just how vast the Amazon is and keep in mind how species rich and diverse it compares to anywhere else in the world. We are still to discover much, much more of the Amazon basin, and because of this if the theory of my research is applied to even a fraction of what wildlife is based in the Amazon, I believe a lot of taxonomic revisits would be in order.

There is a significant amount of research still to be done in this area, and I feel it would be beneficial to significantly expand the sample size. I think if we had of had more species to work with we would have seen a much bigger genetic difference and show that genetic differences occur over a much bigger area.

References

Arnhart. A (1998) *Darwinian natural right. The biological ethics of human nature*, Pg 106. State university of New York press, United Sates of America.

Ayres, J. M, Clutton-Brock, T. H (1992) River boundaries and species range size in Amazonian primates. *The American naturalist*. 140(3) pp. 531-537.

Avise, J. C (2000) *Phylogeography. The history and formation of species*, Pg 3. Library of Congress, United States of America.

Baer, J. F., Weller, R. E., K. Ibulaimu (2012) *Aotus: The owl monkey*, Pg 103. Academic Press, United Kingdom

Baum. D (2015) *Tree thinking: An introduction to phylogenetic biology*, 1st edition. Content technologies publication, United Kingdom.

Beebee. T, Rowe. G (2005) *An introduction to molecular biology*, Pg 165, 170. Oxford University Press, Oxford.

Borodovsky. M, Gogarten, J. P, Przytycka, T. M, Rajasekaran. S (2010) *Bioinformatic research and applications*, Pg 113. Springer, United States of America.

Boubli, J. P., Ribas. C, Lynch Alfaro., J. W, Alfaro., M. E, Da Silva., M. N. F, Pinho., G. M, Farias, I. P. (2015). Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular phylogenetics and evolution*. 1 (1), 4-13

Byrne. H, Rylands, A. B, Carneiro, C. J, Alfaro, J. W. L, Bertuol. F, Da Silva, M. N. F, Messias. M, Groves, C. P, Mittermeier, R. A, Farias. I, Hrbeck. T, Schneider. H, Sampaio. I, Boubli, J. P (2016) Phylogenetic relationships of the new world titi monkey (*Callicebus*): First appraisal of taxonomy based on molecular evidence. *Frontiers in zoology.* Pg. 1-5.

Campbell, C. J., Fuentes, A., Mackinnon, K. C., Bearder, S. K., Stumpf, R. M (2011) *Primates in perspective*, Pg 122. Oxford University Press, New York.

Campbell, C. J (2008) *Spider monkeys: Behaviour, ecology and evolution of the genus Ateles.* Pg. 67. Cambridge University Press. United States of America.

Casada. F, Bonvicino, C. R, Seuanez, H. N (2006) Phylogeographic analysis of *Callicebus Lugens* (Platyrrhini, Primates) *Journal of heredity*. 98(1) Pg. 88-92

Cheng. L, Zhang, D. Y (2010) Molecular genetic pathology, Pg 98. Springer. United States.

Rapley. R (1994) Polymerase chain reaction. *Molecular biomethods handbook*. Pg. 325

Chenna. R, Sugawara. H, Koike. T, Lopez. R, Gibson, T. J, Higgins, D. G, Thompson, J. D (2003) Multiple sequence alignment with the Clustal series of programmes. *Nucleic acids research*. 31(13) Pg. 3497-3500.

Cockell, C. S (2015) *Astrobiology: Understanding life in the universe*, Pg 113. John Wiley & Sons, United Kingdom.

Cooper, N. C (1994) *The human genome project: Deciphering the blue print of hereditary*, Pg 128. University science books. United States of America.

DeLuycker, A. M (2007) *The ecology and behaviour of the Rio Mayo titi monkey (Callicebus oenanthe) in the Alto Mayo, northern Peru*, Pg 22, 9. ProQuest LLC, United states of America.

EMBL-EBI (2016). Clustal omega, multiple sequence alignment. European Molecular Biology Laboratory. Available at: http://www.ebi.ac.uk/Tools/msa/clustalo/

Filatov, D. A (2002) PROSEQ: A software for preparation and evolutionary analysis of DNA sequence data sets. *Molecular ecology resources*. 2(4) Pg 621-624.

Forster. M, Forster. P (2017) http://www.fluxus-engineering.com/sharenet.htm

Freeland, J. R, Peterson, S. T, Kirk. S (2011) *Molecular ecology*, Pg 179. John Wiley & Sons LTD, United Kingdom

Garber, P. A, Estrada. A, Bicca-Marques, J. C (2008) *South American primates: Comparative perspectives in the study of behaviour, ecology, and conservation*. Pg. 234, 5. Springer. New York.

Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, et al (1998) Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Molecular Phylogenetic Evolution*. 1998;9(3):585–98.

Gupta. S, Nadarajan. N, Gupta, D. S (2013) *Legumes in the omic era*, Pg. 258. Springer. United States of America.

Habel, J. C, Assmann. T (2009) Relict species. Phylogeography and conservation biology, Pg 179. Springer, Germany.

Hancock, J. M, Zvelebil, M. J (2014) *Concise encyclopaedia of bioinformatics and computational biology*. Wiley Blackwell. United Kingdom.

Hedrick, P. W (2005) *Genetics of populations*, 3rd edition. Pg. 605. Jones and Bartlett Publishers.

Heulsenbeck, J. P, Ronquist, F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics application note. 17(8) Pp. 754-755.

Huson, D. H, Scornavacca. C (2010) A survey of combinatorial methods for phylogenetic networks. *Genome biology and evolution*. 3 Pp. 23-35.

Howard, D. J, Berlocher, S. H (1998) *Endless forms: Species and speciation*, Pg 205. Oxford University Press, United Kingdom.

Hoyos. M, Bloor. P, Defler. T, Vermeer. J, Rohe. F, Farias. I (2016) Phylogenetic relationships within the *callicebus cupreus* species group (pitheciidae: primates) Biogeographic and taxonomic implications. *Molecular phylogenetics and evolution*. 102. Pg. 208-219.

Huang. Q, Baum. L, Fu, W. L (2010) Simple and practical staining of DNA with GelRed in agarose gel electrophoresis. *Clinical laboratory* 56(3-4) Pg. 149-152.

Hummel. S (2003) *Ancient DNA typing: Methods, strategies and applications*, Pg 113. Springer. Germany.

Innis, M. A, Gelfand, D. H, Sninsky, J. J (1995) *PCR strategies*, Pg, 3-4. Academic press. United Kingdom

Jobling. M (2016) *Facts 101: Human evolutionary genetics: Origins, people & disease.* Content technologies Inc. United Kingdom.

Kinzey, W. K (1997) *New world primates. Ecology, evolution and behaviour*, Pg 213. Library of Congress, United States of America.

Kobayashi S. (1995) A phylogenetic study of titi monkeys, genus *Callicebus*, based on

cranial measurements: I. Phyletic groups of *Callicebus*. Primates. 36(1) Pg. 101–20.

Kumar. S, Tamura. K, Nei. M (2004) MEGA3: Integrated software for molecular

evolutionary genetics analysis and sequence alignment. *Bioinformatics*. 5(2) Pg. 150-163.

Kumar. S, Stecher. G, Peterson. D, Tamura. K (2012) MEGA-CC: Computing core of

molecular evolutionary genetics analysis program for automated and iterative data

analysis. *Bioinformatics*. 28(20) Pg. 2685-2686.

Kurtzman. C, Fell, J. W, Boekhout. T (2011) *The Yeasts: A Taxonomic study*. Pg. 164.

MacPhee, R. D. E (2013) *Primates and their relatives in phylogenetic perspective*, Pg 296.

Plenum Press, New York.

Marsh, L. K, Chapman. C (2013) *Primates in fragments: Complexity and resilience*. Pg. 363.

Springer. New York.

Middleton. N (2012) *Rivers: A very short introduction*, Pg 53. Oxford University Press,

United Kingdom.

Misener. S, Krawetz, S. A (1999) *Bioinformatics Methods and protocols*, Pg 236, volume

132. Humana Press, New Jersey.

Murphey, J. M, Eizirik. E, Johnson, W. E, Zhang, P. Y, Ryder, O. A, O'Brien, S. J (2000)
Molecular phylogenetics and the origins of placental mammals. *Nature*. 409, pg. 614-
618.

Munoz-Torres, M. C (2009*) Comparative genomics and molecular evolution: New genomic
resources for the hymenoptera and evolutionary studies on the genes of Nasonia
vitripennis hox complex*, Pg 121. ProQuest, United States of America.

Nadeau. J (2016) *Introduction to experimental biophysics: Biological methods for physical
scientists*, Pg 574. CRC Press. United States.

Nei. M, S. Kumar (2000) *Molecular evolution and phylogenetics*, Pg 3, 73. Oxford
University Press, New York

Nowak, R. M (1999) *Walker's primates of the world*. Pg. 33. John Hopkins University pess.
United States of America.

Opazo, J. C, Wildman, D. E, Prychitko. T, Johnson. R. M, Goodman. M (2006) Phylogenetic
relationships and divergence times among new world monkeys (platyrrhini, primates).
Molecular phylogenetics and evolution. 40(1) Pg. 274-280.

Paradis. E (2011) *Analysis of Phylogenetics and Evoluton*, Pg 65, 2nd edition. Springer, New
York.

Patton, J. L, Da Silva, M, N. F, Malcolm, J. R (1994) Gene Genealogy and differentiation among arboreal spiny rats (Rodentia: Echimyidae) of the Amazon basin: A test of the riverine barrier hypothesis. *Evolution*. 48(4) pg. 1314-1323.

Pevsner. J (2008) *Bioinformatics and functional genomics.* Pg. 368. John Wiley and Sons. New Jersey.

Pometto. A, Shetty. K, Paliyath. G (2005) *Food biotechnology,* 2nd edition. Pg. 1272. Taylor and Francis. United States of America.

Redmond, I (2010) *Primates of the world*, Pg 86. New Holland Publishers, UK.

Rozas. J, Librado. P, Sanchez-DelBarrio, J. C, Messeguer. X, Rozas. R (2010) DnaSP. DNA Sequence Polymorphism. Available at: http://www.ub.edu/dnasp/

Rozas. J (2009) *DNA sequence polymorphism analysis using DnaSP. Bioinformatics for DNA sequence analysis.* 537. Pg. 337-350.

Ronquist. F, Heulsenbeck, J. P (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 19(12) Pg. 1572-1574.

Rylands, A. B, Mittermeier, R. A (2009) The diversity of the new world primates (platyrrhini): An annotated taxonomy. *South American primates.* Pg. 23-24.

Schrago, C. G (2006) On the time scale of new world primate diversification. *American journal of physical anthropology*. 132(3) Pg. 344-354

Schochetman. G, Ou. Chin-Yih, Jones, W. K (1988) Polymerase chain reaction. *The journal of infectious disease*. 158(6). Pg. 1154-1157.

Simoes, P. I, Stow. A, Hodl. W, Amezquita. A, Farias, I. P, Lima, A. P (2014) The value of including intraspecific measures of biodiversity in environmental impact surveys is highlighted by the Amazonian brilliant-thighed frog (*Allobates femoralis*). *Tropical conservation science*. 7(4) Pg. 811-828.

Stewart. M (2007) *New world monkeys*, Pg 6. Lerner Publications Company, United States of America

Thomas O (1903) *Notes on South-American monkeys, bats, carnivores, and rodents, with descriptions of new species*. Ann Mag Nat Hist, 7th series. 1903;(12):455−464.

Van Roosmalen, Marc. G. M., Van Roosmalen, T., Mittermeier, R. A. (2002). A taxonomic review of the titi monkeys, genus *Callicebus thomas*, 1903, with the description of two new species, *Callicebus bernhardi* and *Callicebus stephennashi*, from Brazilian Amazonia. *Neotropical primates*. 10 (1), 1-2, 32.

Vaughan, T. A., Ryan, J. M., Czaplewski, N. J (2011) *Mammology*, Pg 185, 5th edition. Jones and Bartlett publishers, United States of America.

Veiga, L. M, Barnett, A. A, Ferrari, S. F, Norconk, M. A (2013) *Evolutionary biology and conservation of Titis, Sakis and Uacaris*, Pg 203, 34. Cambridge University Press, United Kingdom.

Weiss. S, Ferrand. N (2007) *Phylogeography of southern European refugia*, Pg 7. Springer, the Netherlands.

Wink. M (2006) *A introduction to molecular biotechnology: Molecular fundamentals, methods and applications in modern biotechnology*, Pg 173. WILEY-VCH. Germany.

Wilson, E. O (2000) Sociobiology. *The new synthesis*, Pg 26. Library of Congress, United States of America.

Xiong. J (2006) *Essential bioinformatics*, Pg. 129. Cambridge University Press. United States of America.

Yamagiwa. J, Karczmarski. L (2014) *Primates and cetaceans: Field research and conservation of complex mammalian societies*, Pg 290. Springer, Japan.

Yeo, S. S, Park, J. J, Park, J. H, Hsu, C. H, Yang, L. T (2010) *Algorithms and architectures for a parallel processing, 10th international conference*. Part 1, Pg. 102. Springer. Germany.

Zvelebil, M. J, Baum, J. O (2008) *Understanding bioinformatics*, Pg 227, 282. Garland Science, Taylor and Francis Group, United States of America.

.