Primate viability in a fragmented landscape: genetic diversity and parasite burden of long-tailed macaques and proboscis monkeys in the Lower Kinabatangan Floodplain, Sabah, Malaysia

MILENA SALGADO LYNN

Thesis submitted to Cardiff University in candidature for the degree of Doctor of Philosophy

September 2010

Cardiff School of Biosciences

Cardiff University

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

"As for me, all I know is that I know nothing"

Socrates

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date 26/4/11

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed ..... (candidate) Date 26/4/11

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by giving explicit references.

Signed ..... (candidate) Date 26/4/11

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date 26/4/11

STATEMENT 4 - BAR ON ACCESS APPROVED

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans *after expiry of a bar on access* approved by the Graduate Development Committee.

Signed ..... (candidate) Date 26/4/11

ABSTRACT

This study investigates the genetic and parasite diversity of two primate species living in the Lower Kinabatangan Wildlife Sanctuary (LKWS), Sabah, Malaysia. Based on non-invasive samples (faeces), the effects of forest fragmentation and geographical barriers, especially the Kinabatangan River, on these two species of primates with different social systems and dispersal abilities were examined. While the proboscis monkey is an endangered primate, the long-tailed macaque is considered one of the most successful invasive alien species. The genetic diversity and the potential effect of the Kinabatangan River on the population structure were examined using microsatellites and a microsatellite library specific for the proboscis monkey was developed during this study. High and moderate levels of genetic diversity were found for the long-tailed macaque and the proboscis monkey respectively. As predicted from the dispersal pattern of these primates, microsatellite analysis revealed low genetic differentiation among sites, suggesting high levels of gene flow as well as regional admixture with one genetically-based cluster inferred from Bayesian analyses. In addition to the neutral genetic marker, as a preliminary approach to study adaptive genetic variation in these populations, *Mhc-DRB* loci were identified in both species using generalist *-DRB* primers. High levels of diversity and evidence of positive selection were found in the long-tailed macaque sequences, which included representatives of several *-DRB* loci/lineages according to phylogenetic analyses. In contrast, only five *-DRB* sequences were detected in the proboscis monkey, all belonging to a single *-DRB* locus; although few, these are the first MHC reported sequences for this species. MHC variability is believed to be maintained by pathogen-driven selection, mediated either through heterozygote advantage or frequency-dependent selection. Using the same samples as for the genetic analysis, a survey of the gastrointestinal parasite fauna of both primates revealed 14 taxa. Parasite richness was higher in proboscis monkeys, and prevalence of particular parasites differed between the primates. Potential effects of natural and anthropomorphic mediated habitat fragmentation on parasite species richness, proportion of individuals with mixed infections and the prevalence of particular parasites between the two primate species were explored. Natural fragmentation (the Kinabatangan River) did not affect parasite distribution. Although it was expected that areas with high rates of human – non-human primate contact would have a positive correlation with the assessed parameters this was not always the case as host-parasite dynamics are likely to be affected by complex interactions between environmental, and host demographic, behavioural and genetic factors.

The results of this study can be used as a baseline for conservation and management measures for the proboscis monkeys and long-tailed macaques of the LKWS.

ACKNOWLEDGEMENTS

Tu Papá, Tu Mamá y Lucalú: esto, como todo, es por y para ustedes.

I am very grateful to all the people who have contributed to this thesis in one way or another. Mike, Jo and Ben, my supervisors, kindly accepted me as a student and guided, supported and encouraged me throughout my studies. I have learnt much from you and can't thank you enough. A standing ovation goes to "The G10/C5.18 Family"; particular thanks go to XiangJiang, Andy, Dave B, Jo B, Rhys, Geoff, Tim, Loys, Jeff, Viola, Alex, Sigrid, Bettina, Dan, Katherine, Wendy, Pati, Eddie and "Little Portugal": Rui, Tania, Joana, Mafalda B and Mafalda C, willing to help and make suggestions all the time, above all, willing to be friends. My deepest gratitude goes to the "Cardiff Ladies", Mireille, Gabi, Renata and Coralie, who provided knowledge, support, laugh, comfort (even housing), friendship, and helped me keep sane throughout the years in Cardiff. I also want to acknowledge Yvonne Linton and John Deeming for their enthusiasm, guidance, help and support for "the mosquito project", pity we were stopped! Additional thanks go to Xiudeng Zheng, Steve Hope & Steve Turner from the sequencing core, Azucena Chaidez, Jannine Manjarrez, and Ian Vaughan, who helped with technical and statistical problems. To Jo Lello and Cock van Oosterhout, who gave very useful comments and advice to make this manuscript a decent looking one. Jo actually ran a few statistical models for me and I'm infinitely grateful.

In Malaysia, I would like to acknowledge the support of the Sabah Wildlife Department and Danau Girang Field Centre, the members of the Kinabatangan Orang-utan Conservation Project and of the Institute for Tropical Biology and Conservation from University Malaysia Sabah for their kind assistance during my sampling seasons. In particular I would like to thank Rosman Sakong and Dave Stanton who worked weird and tough schedules looking for "najis" and parasites.

This project would not have been possible without the financial support provided from the Mexican Government, through CONACYT, and the Rufford Small Grants for Nature Conservation, North of England Zoological Society (Chester Zoo) and the British Society for Parasitology.

A special mention goes to Fairus... you started this, I wish you could have seen it through and continue it with me...

And last but not least: to Ben, who witnessed all and gave me his wholehearted support through the best and the worst of this madness. Thank you for everything, let's keep filling the box!

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

LIST OF FIGURES

CHAPTER ONE

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER ONE

Introduction

This study aims to assess the interaction between habitat loss and fragmentation on the population genetic structure (based on microsatellite markers and MHC genes) and endoparasite diversity of long-tailed macaques (*Macaca fascicularis*) and proboscis monkeys (*Nasalis larvatus*) in the Lower Kinabatangan Wildlife Sanctuary (LKWS). The Kinabatangan region in Sabah, Malaysia, is a major wildlife conservation area. Its broad habitat diversity harbours rare and endangered animals, including 10 primate species. Logging and changes in land use in the region led to the creation of the LKWS in an effort to preserve the natural resources and wildlife from further over-exploitation and extinction. As part of ongoing conservation efforts in the sanctuary, orang-utans, proboscis monkeys and long-tailed macaques have been studied at the population genetic level with the intention of incorporating the resulting data into long-term conservation management programs (Goossens *et al.*, 2005, 2006a; Jalil, 2007), already with success in the case of orang-utans (Orangutan Action Plan, Sabah Wildlife Department, 2010). However, these studies focused only on neutral genetic variation and, even in that context, are still incomplete for proboscis monkeys and long-tailed macaques. Given the need for an assessment of the viability of the remaining populations and their responsiveness to environmental pressures, studies regarding adaptive genetic variation and about specific selection pressures upon them, such as parasites, are more than pertinent. The geography, biodiversity of the LKWS, as well as the impact of habitat degradation on the study site are described in this chapter along with conservation measures taken. The ecology and previous genetic studies of the two primates are also reviewed. Additionally, the utility of different molecular markers and the suitability of non-invasive sources of DNA for population genetic studies are considered. In the context of conservation, the role of parasites as a cause of wildlife declines, with special emphasis on primates is also addressed in this section. Finally, the hypotheses I aimed to test and a description of the layout of the thesis are presented at the end of the chapter.

1.1 The Lower Kinabatangan Wildlife Sanctuary

The Kinabatangan River is the longest river in Sabah and its floodplain is of major importance as a wildlife conservation site. The region consists of a variety of habitats, such as riverine forest, seasonally flooded forest, swamp forest, dry dipterocarp forest, nipah palm and mangrove. Being so diverse, it is a key region for some of Southeast Asia's rarest animals, such as the Bornean elephant and the estuarine crocodile. It is also one of only two places on earth where 10 primate

species are known to live in sympatry (WWF – Asia/Pacific, 2005). Over the past 40 years, the forest surrounding the Kinabatangan River has been extensively logged and replaced by oil palm plantations with notable effects on wildlife populations (e.g. Goossens *et al.*, 2005, 2006a). These activities could also pose new threats to these species in terms of the emergence of infectious diseases.

1.1.1 Geography

The Kinabatangan River is located in the state of Sabah, Malaysia, on the island of Borneo (Fig. 1.1). Its catchment area covers almost 23% of the total area of Sabah. The main channel length is approximately 560 km and flows eastwards towards the Sulu Sea, draining an area of about 16,800 km^2, of which about 3,000 km^2 forms the coastal plain and 13,800 km^2 forms the upper catchment (Regip *et al.*, 2004). The Millian River and the Kuamut River are the main tributaries of the Kinabatangan. Its floodplain is the largest still forested in Sabah, and contains some of the most important remaining freshwater swamp rainforests and oxbow lakes in South East Asia, meaning that the river is geomorphologically active. The mean annual rainfall in the catchment is about 2500 mm, although some zones receive up to 3000 mm. Rainfall is heaviest during the north-east monsoon (October – March) which results in substantial floodings. Daily temperatures fluctuate between 23°C to 32°C, with almost no monthly variation (Vaz, 1997). Both, climate and rainfall directly influence the diverse habitats and wildlife along the river.

Figure 1.1. Map of the Kinabatangan River Basin (http://assets.panda.org/img/original/kinabatanganmap.gif)

1.1.2 Wildlife

The vegetation in the area is distributed according to its resistance to the frequency of flooding, salinity and waterlogging. Therefore, a variety of habitats can be found along the river and catchment area. The mouth of the river is predominantly covered by mangrove forests. Above the tidal and brackish waters, riverine forest develops along the main channel. Scattered across the floodplain, freshwater swamp forests develop on soils that are permanently or occasionally flooded. Additionally, about 30 oxbow lakes are distributed along the region (Davison, 2002). Limestone outcrops (karsts) from sparse higher elevation features throughout these lowlands. Away from the river, on drier land, the main primary vegetation is dryland forests comprising dipterocarp trees, however, these forests have been replaced by oil palm plantations almost entirely (Vaz, 1997) and what remains is heavily degraded and/or isolated as extensive logging and fires have spread in the region (Majail & Webber, 2006). Nevertheless, the broad habitat spectrum in the Lower Kinabatangan supports a high diversity of fish, birds, reptiles and mammals, some of which can only be found in Borneo. However, perhaps the most notable group of mammals in the region is the primates (Rautner *et al.*, 2005).

The primate fauna of the Kinabatangan includes a strepsirrhine (slow loris), seven catarrhines (proboscis monkey, two species of macaques, three species of leaf monkeys (langurs), and a tarsier), one lesser ape (gibbon) and one great ape (orang-utan) (Davison, 2002). The tarsier (*Tarsius bancanus*) and slow loris (*Nycticebus coucang*) can be found in tall and secondary forest, and sometimes in gardens and cocoa plantations. Langurs have settled in different niches to equitably share resources. Red (*Presbytis rubicunda*) and Hose's (*P. hosei*) langurs tend to occupy tall lowland dipterocarp forests (and are endemic to Borneo), while silvered langurs (*P. cristata*) tend to occupy secondary forests and swampy areas. Long-tailed and pig-tailed macaques (*Macaca fascicularis* and *M. nemestrina*) have adapted to surviving in secondary and logged forests, thus they are the most abundant primates in the region. In contrast to the macaques, the one colobine of the region, the proboscis monkey (*Nasalis larvatus*), is more affected by logging as its dietary requirements restricts it to riverine, peat swamp and mangrove forests. The apes of the region, the Bornean gibbon (*Hylobates muelleri*) and the orang-utan (*Pongo pygmaeus*), can be found in hill and lowland dipterocarp forests (Vaz, 1997), and seem to adapt well to logged and degraded forests (Ancrenaz *et al.*, 2004; Ancrenaz *et al.*, 2010, Rautner *et al.*, 2005). However, all of these primates are threatened to some extent by habitat loss and other human activities.

1.1.3 Human activity and its consequences in the region

Human settlements and activities have been recorded for centuries in the Kinabatangan (Payne, 1989). The river has traditionally been used for transport, trade and communication by the Orang Sungai (people from the river). Nowadays, Malaysia is the world's leader in palm oil production and, consequently, the ecosystem has been drastically modified (Rautner *et al.*, 2005), especially in the Kinabatangan. Pollution and threat to wildlife due to habitat loss are some of the negative effects of such changes (Regip *et al.*, 2004; WWF Malaysia, 2003). As a consequence, management plans for conservation and sustainable use of resources are now being developed and implemented (Payne, 1989; Davison, 2002; Regip *et al.*, 2004).

Since the 1950s, the Kinabatangan forests have become degraded and have been largely replaced with oil palm plantations. Some forest reserves were established in the 1960s for selective logging and natural regeneration, but a decade later they were redesignated for conversion to permanent agriculture, except for a few small forest reserves (Vaz, 1997). From then on, the land has been used to plant oil palms, which makes the oil palm industry the second major stakeholder in the region along with the Orang Sungai (WWF Malaysia, 2003). Today, about 20,000 of the 65,000 ha susceptible to flooding have been cleared for agriculture and the Kinabatangan landscape now consists of patches of protected forest surrounded by oil palm plantations and secondary forest (Davison, 2002; Rautner *et al.*, 2005). Habitat loss and fragmentation of forest habitat has resulted in many wildlife species being restricted to narrow riparian corridors. Consequently, conflict between wildlife and local people has increased, especially during flooding episodes when animals move to higher grounds (WWF Malaysia, 2003). Conflicts include the destruction of property and plantation crops by elephants and the recurrent negative interactions (i.e. crop raiding) between people and primate fauna.

1.1.4 Conservation Programs

The Lower Kinabatangan Wildlife Sanctuary (LKWS) was fully gazetted in 2005 thus reflecting the conservation efforts made since the middle of the last century. From the 1930s to the 1970s some areas in the region were designated Virgin Jungle Forest Reserves (CAIMS, 2005), however, they were not connected. Hence, a proposal to establish a Kinabatangan Wildlife Sanctuary was made in the late 1980s (Payne, 1989). The aim was to create a large (44,300 ha) sanctuary including the virgin forest patches and other sites of special interest, and enhance the value of the area for conservation of (especially) large, wide-ranging animals. However, at that time the proposal did not succeed. Over the next decade efforts in creating a sanctuary continued and in 1999, 26,000 ha of the floodplain were designated as a "Gift to the Earth" by the State Government (WWF Malaysia, 2003). Efforts continued to create a "Corridor of Life" connecting

10 disconnected forest patches (Lots), where "people, wildlife, nature based tourism, and forest industries thrive and support each other" (Pang & Prudente, 2003), and six of 10 proposed lots were gazetted as a Bird Sanctuary under the Land Ordinance (Environmental and Conservation Workgroup, 2001). In 2002, the proposed LKWS was gazetted under the Land Ordinance, comprising 10 lots (with lot 10 divided in A-C) (Pang *et al.*, 2002), however the aim was to fully gazette the sanctuary under the Wildlife Conservation Enactment 1997. Work continued and habitat restoration began to re-join the forested areas by planting native trees to serve as corridors for the movement of some species of wildlife (Davison, 2002). In parallel, an Integrated River Basin Management Plan was proposed for the sustainable management of land and water resources to reduce the impact of erosion and flooding (Mathew, 2004). Finally, in 2005, the 10 lots of the LKWS (26,103 ha) were completely re-gazetted under the Wildlife Conservation Enactment 1997 (Pang *et al.*, 2005) (Figure 1.2). The vision for this sanctuary in terms of conservation is "the stabilization of wildlife populations" by 2020 (Majail & Webber, 2006).

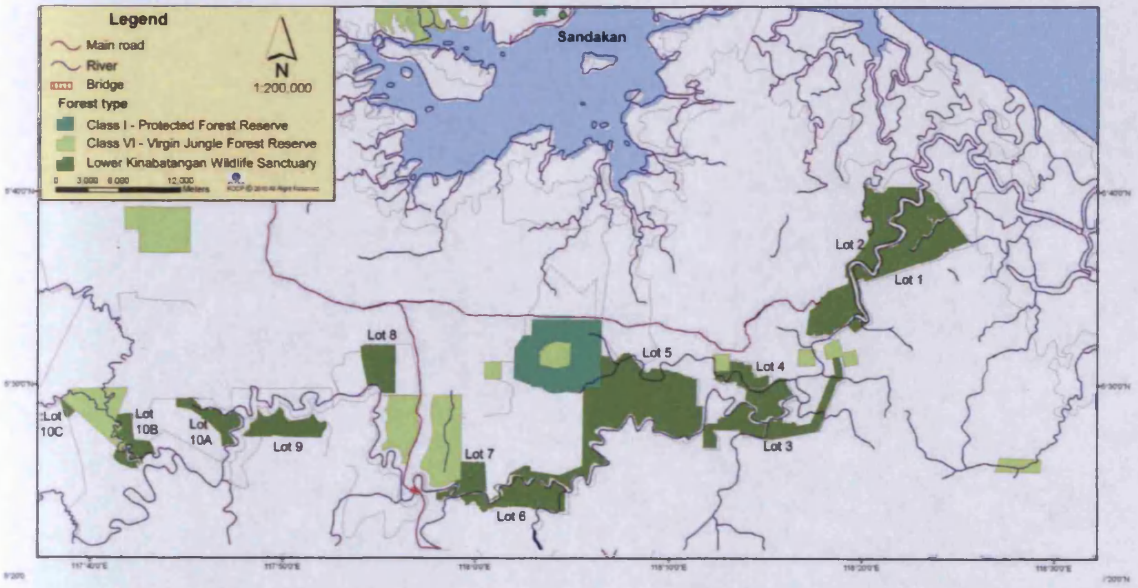


Figure 1.2. Map of the Lower Kinabatangan Wildlife Sanctuary (Courtesy of HUTAN). The ten Lots of the sanctuary (depitcted in darker green) are distributed along the North and South side of the Kinabatangan River (sinuous blue line).

In the case of primates, the necessity of establishing a monitoring system for detection of population trends over time has been already stated (Goossens *et al.*, 2003a). Recent studies of some primate populations have led to the design and implementation of management measures necessary to ensure their long-term survival in the sanctuary (Ancrenaz *et al.*, 2003; Goossens *et al.*, 2005). In the case of orang-utans, molecular analysis of mitochondrial and nuclear DNA variation have clearly shown substantial genetic differentiation between populations on either

side of the river (Goossens *et al.*, 2005; Jalil, 2007; Jalil *et al.*, 2008). Therefore to augment genetic diversity, there is a need to facilitate migration of orang-utan populations between lots in the sanctuary. To accomplish this, it has been proposed to restore forest corridors alongside the river banks and between lots, and also to translocate individuals from opposite sides of the river (Goossens, *et al.*, 2005), thus reducing the chance of extinctions providing that the habitat is also maintained.

Although many metapopulation models show that increasing movement among populations reduces the chances of metapopulation extinction, epidemiological models indicate that increased contact among populations enhances the spread of disease and can trigger epidemics (Hess, 1996). This could be done by introducing diseases or parasites to previously naïve subpopulations, or by lowering the fitness of recipient populations through the introduction of alleles that are maladapted to cope with such challenges, or by the breaking of coadapted gene complexes (Reed, 2004). With that in mind, it is important to also monitor the status of wildlife diseases in the LKWS.

1.2 Study subjects

1.2.1 The long-tailed macaque (Macaca fascicularis Raffles, 1821)

The genus *Macaca* is the most widespread from the Cercopithecidae family. It is distributed from North Africa to Japan and it inhabits a broad range of habitats. Ten macaque species inhabit the Indonesian archipelago and according to geographical distribution are grouped in three distinct phylogenetic lineages: *silenus-sylvanus*, *sinica-arctoides* and *fascicularis* (see Abegg & Thierry, 2002). The long-tailed macaque (*M. fascicularis*) is grouped in the fascicularis lineage along with the rhesus macaque (*M. mulatta*), the Formosan macaque (*M. cyclopis*) and the Japanese macaque (*M. fuscata*) (Abegg & Thierry, 2002). Adults are sexually dimorphic with males weighing 5 – 7 kg and females 3 – 4 kg (Aldrich-Blake, 1980; Harcourt & Schwartz, 2001).

Long-tailed macaques inhabit a variety of forest habitats throughout their native range, preferring edge habitats and riverine areas, but can also be found in village areas (i.e. disturbed habitat). They are considered a pest in some areas of the region because of their raids on fields and gardens and depredations on crops, and for that reason they are often killed (Long, 2003; van Schaik *et al.*, 1996). They have been listed as one of the 100 most invasive alien species with successful invasions in Sulawesi, Lesser Sunda, Palau, Mauritius, Papua New Guinea and Hong Kong (Long, 2003; Lowe *et al.*, 2000) as they are omnivorous and opportunistic feeders. These monkeys are well equipped to successfully disperse over water barriers, and rely mainly on

riverine and coastal forests including mangroves. They often feed and sleep in preferred trees above rivers, high above ground as a protection against predators (Abegg and Thierry, 2002). They are diurnal and can be totally or semi-arboreal, moving through the forest canopy and on land quadrupedally, but are also considered good swimmers (Richard, 1985). In Sumatra, long-tailed macaques occur at high densities in selectively logged forest, secondary forest and cultivated land (Supriatna *et al.*, 1996).

The social structure of long-tailed macaques is usually a multimale-multifemale group, where females and juveniles tend to be related and organized in a hierarchical matriline (de Ruiter & Geffen, 1998) where dominant males appear to father the majority of offspring (de Ruiter *et al.* 1994; Engelhardt *et al.*, 2006). However, dominant males also seem to share mate guarding and mating with the dominant male of adjacent groups (Engelhardt *et al.*, 2006). Females are philopatric, while males migrate at maturity and join other groups, and dispersing males have been observed to swim across rivers in Sumatra (de Ruiter & Geffen, 1998). New groups may be formed by fission of a large group (implying female dispersal), which will be dominated by an alpha male (ibid).

In the Kinabatangan, Goossens *et al.* (2003a) observed a social structure of multimale-multifemale with group sizes of 10-48 and occasionally up to 100 individuals; solitary males were also observed. They utilised a home range of 25-200 ha, with a population density of 6.34 individuals/km^2. Long-tailed macaques were seen ranging along riverbanks, particularly when water levels were low during the dry season where they can be seen fishing for crabs (hence their other common name, the crab-eating macaque). According to the same study the population size in the LKWS is approximately 3170 individuals, and this more likely be an underestimate than an overestimate (Goossens *et al.*, 2003a).

1.2.2 The proboscis monkey (Nasalis larvatus van Wurmb, 1787)

Nasalis larvatus is a large, sexually dimorphic, arboreal monkey, endemic to the island of Borneo with no evidence that it has ever occurred elsewhere (Harcourt & Schwartz, 2001; Meijaard & Nijman , 2000; Payne & Francis, 1998). This species rarely ranges far from rivers, is restricted to lowlands and is typically associated with coastal forest, including mangroves, and riverine peat swamp and fresh water swamp forests (Kawabe & Mano, 1972; Meijaard & Nijman, 2000). It has been reported that the only large populations in Sabah occur in the fresh water wetlands of the Kinabatangan flood plain, and around Dewurst Bay in the Eastern Deltas (Meijaard & Nijman, 2000), however, they also occur in other parts of the State as seen in Figure 1.3 (Sha *et al*, 2008). Currently, the proboscis monkey is threatened by habitat destruction and hunting, and

much of its former range has been reduced due to logging (e.g. in Kinabatangan), swamp reclamation, gold mining, shrimp farming and forest fires (Meijaard & Nijman, 2000; Sha *et al.*, 2008). Hunting is much in evidence in Sarawak and Kalimantan (Meijaard & Nijman, 2000). Its conservation status has changed over the past ten years from vulnerable to endangered according to IUCN (2010). It is currently listed in Appendix I of CITES (UNEP-WCMC, 2010) and is protected by law throughout its range (Meijaard & Nijman, 2000; Sha *et al.*, 2008).

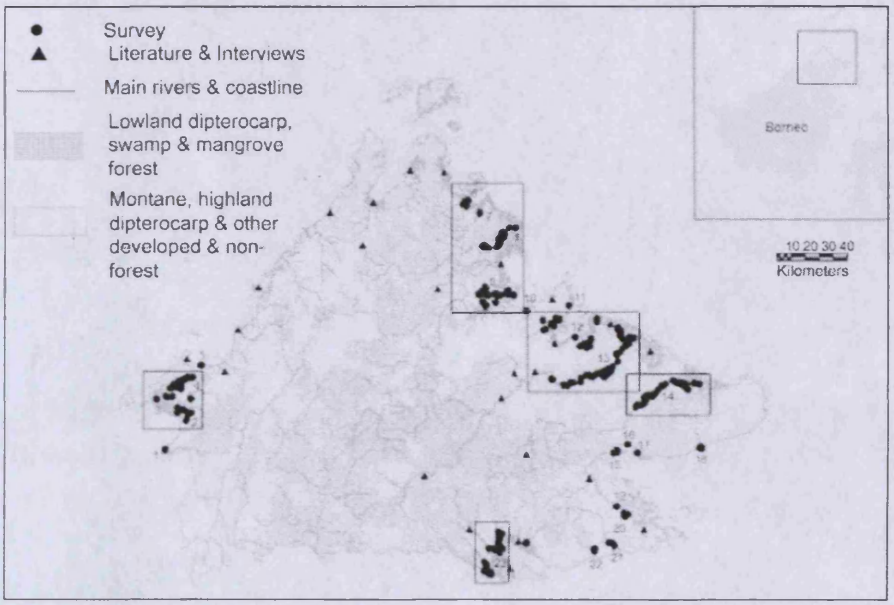


Figure 1.3. Distribution of proboscis monkey in Sabah (taken from Sha *et al.* 2008, with permission from the authors). "Survey" indicates locations of sightings from that study, "Literature and interviews" indicate sightings from literature, interviews and other sources that were not verified in that study. "Boxed" areas are identified major centres of continuous distribution. Numbers correlate to those in Table 1 of the same study.

This species belongs to the sub-family Colobinae which includes at least 30 species classified into 9 genera (Murai *et al.*, 2007). These species are characterized by a multi-chambered stomach that is an adaptation to the digestion of foliage (ibid). Of all colobine monkeys, the proboscis monkey presents the most distinct sexual dimorphism. Adult males have the largest body size of all the colobine species, and their noses are long and overhang their mouth, while adult females are about half the weight of males and have much shorter noses (Murai, 2006).

Proboscis monkeys have a flexible social structure. Harems (consisting of a single male and several females with their offspring) are predominant, while young males usually emigrate from their natal groups, subsequently forming all-male groups (consisting of juveniles, adolescents and adult males); solitary males are rare (Bennett & Sebastian, 1988; Boonratana, 2000, 2002; Murai

et al., 2007; Yeager, 1995). Males as young as 18 months will leave their natal group and join an all male group. In general, groups consist of three to 30 individuals and both types of groups (one male and all-male) usually come into close proximity during the evening as they migrate close to rivers to sleep in trees (Bennett & Sebastian, 1988). Non-breeding groups are loosely bonded, predominantly male groups with up to one female member (Boonratana, 2000). There is a secondary level of organization, the band, with fission-fusion of stable one-male groups within bands. Several one-male groups will often form a multilevel society, and male replacement seems relatively peaceful compared to other colobines, without any evidence of fighting (Murai, 2006).

In addition to male emigration, female transfer has also been confirmed by observed changes in group composition (Murai *et al.*, 2007). Female transfer has been observed in several groups of folivorous primates (e.g. Alouatta sp. and the Colobinae). Proboscis monkeys' adult females will sometimes leave their natal single male group to join another group, but may later re-join their original natal group (Bennett & Sebastian, 1988; Murai, 2006). Subadult and juvenile females are often found temporarily in otherwise all-male groups and will copulate with them (Murai *et al.*, 2007). Resident adult males are indifferent to transfers of the subadult females; whereas they call out to adult females who move from their group to another, and recouping is not aggressive. This transfer behaviour suggests that females need to disperse from their natal groups before reaching full maturation in order to avoid inbreeding. Females have a tendency to transfer to a larger group, usually in close proximity to the original one, indicating that intra-group feeding competition is weak and has no bearing on female transfer in proboscis monkeys. Consequently, females of this species are not constrained in their transfer behaviour by food resources. Infanticide has been reported in Labuk Bay sanctuary, but avoidance of infanticide may not be an important factor in female transfer (Murai *et al.*, 2007).

Groups generally move inland in the early morning, returning to trees next to the rivers before dusk. They avoid large areas lacking natural riverine vegetation or areas dominated by tall grass and scrub (Boonratana, 2000). The home ranges of various groups overlap extensively, sharing the same night trees, food sources and migration routes (Boonratana, 2000; Matsuda, 2008; Murai *et al.*, 2007). This implies a high degree of intergroup tolerance and the absence of territorial behaviour. Ranging behaviour appears to be influenced by the location of river crossing points and the availability of foods, in particular fruits (Matsuda *et al.*, 2009); scarce and unpredictable food supplies require individuals to travel over considerable distances (up to 2 km), which results in large range size (Boonratana, 2000; Matsuda *et al.*, 2009). In addition to food availability, ranging behaviours, including the selection of sleeping sites, may be affected by predation pressure (Matsuda *et al.*, 2009). According to Matsuda (2008), at least 800 m from both river banks is needed for the survival of the species.

In the LKWS, in 2003(a) Goossens *et al.* reported about 3500 proboscis monkey individuals, with a population density of 6.86 individuals/km^2, by 2008 the numbers had reduced by more than half (1450 individuals) (Sha *et al*;, 2008). Group sizes are about 4-20 individuals, with a home range of 900 ha (ibid). The social structure consists of one male-multifemale group, aggregating into multimale-multifemale, or all-male bachelor groups, which have been observed to associate; solitary males have not been observed. The proboscis monkey is one of the major tourist assets of the LKWS. However, their long-term survival is jeopardized by habitat fragmentation and land clearance, hunting and tourism activities. Despite considerable knowledge of this species, there is only one population genetic study of this endangered colobine (Jalil, 2007). It is therefore important to monitor the proboscis population along the Kinabatangan, as well as that of other primate species.

1.3 Primate population genetic studies in the LKWS

Previous population genetic analyses in primates of the LKWS have focused on neutral (demographically mediated) genetic variation. Both the long-tailed macaque and proboscis monkey showed high levels of genetic diversity based on mitochondrial DNA data (Jalil, 2007). However, further studies using more molecular markers are needed to complement the existing information and thus, sustain a more robust conservation management plan. Genetic studies were also performed in orang-utans and along with the proboscis monkeys populations they seem to have undergone historic demographic expansions (Jalil, 2007), which in the case of orang-utans correlates with the expansion of the lowland forest after the last glacial maximum (Jalil *et al.*, 2008).

Using 14 microsatellite loci to characterize patterns of genetic diversity in orang-utans, Goossens *et al.* (2005) found evidence for recent/ongoing migration along the same side of the river, but limited /no gene flow across the river. This genetic differentiation was supported by Jalil *et al.* (2008), analysing the mtDNA control region, indicating that the river is a major barrier for dispersal of orang-utans in the LKWS. The remaining Kinabatangan populations show a reasonably high level of genetic diversity (average heterozygosity, He = 0.74) (Goossens *et al*, 2005) which is the remnant of an ancient large population and the result of recent population expansions by a few founder lineages (Jalil *et al.*, 2008). Despite such levels of genetic diversity, Goossens *et al.* (2006a) detected evidence for a recent major population decline with no support for growing or even stable populations; this decline coincides with the first (colonial) forest exploitation and habitat degradation in the country. The same study estimates an extremely low population size, prone to be eliminated by genetic drift. However, these studies also showed that

populations still exhibit high levels of genetic diversity and suggested plans to encourage migration of orang-utan populations between lots in the sanctuary to augment their genetic diversity and avoid the depletion of rare alelles. The genetic data of the studies (Goossens *et al.*, 2005, 2006a) are included in the Sabah Orang-utan Management Plan.

Previous genetic studies performed on long-tailed macaques and proboscis monkeys were solely based on mtDNA (Jalil, 2007). Long-tailed macaques showed high levels of genetic diversity and the inference was that the population has remained stable for a long period of time with no evidence of a major demographic expansion in the recent evolutionary past. This long term stability could be due to adaptability to different habitats and the opportunistic nature of these monkeys. However, there was some evidence of genetic differentiation on each side of the Kinabatangan River. This could be a consequence of the females being philopatric and, perhaps, by biases in demographic estimates reached using a maternally inherited molecular marker. It was therefore important to know if this inferred structure was confirmed or whether it needed to be modified when analysed with different molecular markers, such as microsatellites.

Proboscis monkey populations also showed high levels of genetic diversity possibly as a consequence of female biased dispersal and/or a large ancestral population that inhabited Borneo (Jalil, 2007). Three mitochondrial haplogroups were found to be assorting within the population and a lack of geographical partitioning might have been the product of secondary contacts from three separate refugia (Jalil, 2007) or simply ancestral lineage retention. Although this study implies that the proboscis monkey population in LKWS is not facing immediate danger, the present levels of genetic diversity might actually reflect past demography as deforestation might have forced widely dispersed populations to disperse into the Kinabatangan, as is the case for the orang-utan in this area (Goossens *et al.*, 2006a). Furthermore, Jalil (2007) notes that a single molecular marker is insufficient to provide information for conservation purposes, and for that reason he suggested using a much faster evolving molecular marker, such as microsatellites, to incorporate male dispersal and accurately depict the genetic structure of the populations.

1.4 Molecular markers suitable for population genetic studies

In conservation biology, knowledge of population structuring potentially provides valuable information for conservation strategies and management. Molecular tools are commonly used to investigate the genetic structuring of populations, addressing specific conservation questions regarding, and helping to guide, the establishment of management programs (Balloux & Lugon-Moulin, 2002). Neutral genetic markers have great potential for investigating processes such as disperal, migration or gene flow. Microsatellites have become the preferred neutral nuclear

marker in many studies because of their high levels of variability, ease and reliability of scoring, codominant inheritance and short lengths, making them useful for studies of DNA from hair or faeces collected in the field (Luikart & England, 1999; Broquet *et al.* 2007). However, although being by far the most common class of genetic markers used, neutral genetic markers give limited insight into the specific adaptive variations and the population viability of a species, including the adaptive consequences of management of gene flow (Goossens *et al.*, 2002; Holderegger *et al.*, 2006). Non-neutral genetic markers instead focus on genes and regulatory elements that have adaptive or selective significance. In contrast, natural selection will directly act on genotypes favouring the ones that might have a higher overall fitness, genome-wide (Holderegger *et al.*, 2006). The extant of individual variation at functionally important genes such as some loci in the Major Histocompatibility Complex (MHC) is thought to be a good candidate for adaptive significance. By investigating the extent of MHC variation for species which have undergone population bottlenecks, such as is likely to have happened in proboscis monkeys and seems certain to have happened in orang-utans, we can gain an indirect insight into their potential resistance against various diseases, which could be of great importance for the re-establishment of self-sustained populations.

1.4.1 Microsatellites

Microsatellites are also known as simple sequence length polymorphisms (SSLPs), short tandem repeats (STRs) or simple sequence repeats (SSRs). They are resolved as single copy nuclear DNA consisting of tandemly repeated short sequence motifs, each between one and 10 bp in length, such as (GACA)n, (TAT)n, or (CA)n (Beebe& Rowe, 2004; Bruford *et al.*, 1998; Hancock, 1999). They are widely scattered throughout the genome and are highly polymorphic due to variation in the number of repeat units at many of the loci so far studied (Bhargava & Fuentes, 2010; Ellegren, 2004; Li *et al.*, 2002;). Therefore, microsatellites have found their application in areas such as linkage mapping, individual identification, forensics and in Hitchhiking mapping (Ellegren, 2004). These markers are inherited in a Mendelian fashion (as codominant, where alleles from both parents are traceable in the offspring) (Beebe & Rowe, 2004; Bhargava & Fuentes, 2010; Zhang & Hewitt, 2003) and the data generated are similar to that of the previously used allozyme markers, except that the number of alleles and heterozygosity revealed is almost always much higher (Chambers & MacAvoy, 2000; Ciofi *et al.*, 1998).

Microsatellites are appropriate markers for many population and biodiversity studies for several reasons. They are believed to be largely selectively neutral with largely independent evolution and high mutation rates, ranging from 10^{7} to 10^{2} mutations per locus per generation in eukaryotes (Bhargava & Fuentes, 2010; Buschiazzo & Gemmel, 2006). Additionally, by

combining many, highly variable, independent microsatellite loci a unique multilocus genotype can be obtained: a "molecular tag" or "fingerprint" for every individual in a population can be provided (Beebe & Rowe, 2004; McKelvey & Schwartz, 2004). Such characteristics make them useful for studying paternity and kinship, genetic variation, population genetic structure and gene flow, and to study recent population history (Bruford & Wayne, 1993; Jarne & Lagoda, 1996; Kohn & Wayne, 1997; Zhang & Hewitt, 2003). Many markers now exist for many species and cross-species microsatellite amplification allows the use of many microsatellite markers for unstudied related species (Allendorf & Luikart, 2007; Beebe & Rowe, 2004; MacKelvey & Schwartz, 2004), making the protocols to develop new or additional microsatellites simpler and less time consuming. Moreover, material for microsatellite analysis can potentially be sampled non-invasively from free-living populations, which in the case of endangered species is essential (Broquet *et al.*, 2007; Goossens & Bruford, 2009; Mills, 2007). Besides, the analytical methods (maximum likelihood, coalescent and Bayesian statistical approaches) for retrieving information from microsatellites allow more detailed inference about both evolutionary parameters and historical events (Allendorf & Luikart, 2007; Hedrick, 2001; Luikart & England, 1999). Such methodologies also generally provide more precise and accurate estimates of population parameters such as migration rates, effective population size, and intra- and interlocus disequilibrium. Moreover, it is also possible to identify a recently "bottlenecked" (or declining) population when no information exists on the current or historical population size and combined with other methods it is possible to simultaneously estimate the approximate date and rate of a recent reduction (or increase) in the effective population size (Allendrof & Luikart, 2007; Hedrick, 2001; Luikart & England, 1999; Zhang & Hewitt, 2003).

Methods to identify and characterize highly polymorphic loci for primate population genetic studies have included isolation of novel microsatellites directly, and more recently linkage analysis from whole genome sequences (Raveendran *et al.*, 2006). However, the screening of human primers with particular amplification conditions has proved that many loci are also informative in several non-human primates, including gorilla, chimpanzee, orang-utan, macaque and langur (Arandjelovic *et al.*, 2009; Bayes *et al.*, 2000; Bonhomme *et al.*, 2005; Bradley *et al.*, 2000; Clifford *et al.*, 1999; Coote & Bruford, 1996; Goossens *et al.*, 2005; Hadfield *et al.*, 2001; Launhardt *et al.*, 1998; Little, 2003). Microsatellites have already been used to study the genetic structure of the orang-utan population in the LKWS, but not in long-tailed macaques and proboscis monkeys. More than 150 human primers have been already tested in *Macaca fascicularis*, and at least 70 markers have successfully cross-amplified (Bonhomme *et al.*, 2005; Kikuchi *et al.*, 2007). No microsatellite markers have been reported for *Nasalis larvatus*, but human microsatellites have been tested in other colobines (Coote & Bruford, 1996; Launhard *et*

al., 1998; Little, 2003). For example, at least 20 markers have proved useful in paternity testing and relatedness analysis in Hanuman langurs (*Presbytis entellus*) (Launhardt *et al.*, 1998; Little, 2003).

Microsatellites therefore provide one of the most powerful and practical means currently available for analysing genetic diversity in threatened species. However it should be borne in mind that primers developed in some species may detect lower levels of variation among species that are not closely related due to ascertainment bias (Dakin & Avise, 2004; Primmer *et al.*, 1996).

1.4.2 MHC

Genes of the Major Histocompatibility Complex are thought to be of important adaptive significance and have thus been used many times as indicators of adaptive genetic variation (Robinson *et al.*, 2003). They have the highest known variation in any vertebrate genes, and are thought to influence many important biological traits including immune recognition, susceptibility to infectious and autoimmune diseases, individual odours, mating preferences, kin recognition, cooperation and pregnancy outcome (Penn & Potts 1999; Penn, 2002; Robinson *et al.*, 2003; Schwensow *et al.*, 2007; Sommer, 2005).

MHC-encoded molecules are transmembrane glycoproteins that bind antigens derived from non-self molecules and present them to T lymphocytes, which in turn initiate the immune cascade. There are two major groups of MHC genes: class I and class II (Figure 1.4). MHC class I genes are expressed on virtually all nucleated somatic cells and their products are essential for immune protection from intracellular pathogens. MHC class II genes are only expressed on specific antigen-presenting cells such as B cells and macrophages, and their coding proteins bind and present peptides mainly stemming from extracellular parasites (Bernatchez & Landry, 2003; Hughes *et al.*, 1997; Hughes & Yeager, 1998; Schwensow *et al.*, 2007; Sommer, 2005). Given the high number of alleles in both groups of MHC, they are sometimes functionally classified as supertypes depending on the overlapping of their peptide-binding specificities (Trachtenberg *et al.*, 2003).

MHC Class I genes are divided into classical and non-classical. MHC-A, -B and -C are the classical genes and are highly polymorphic, while the non-classical genes MHC-E, -F and -G are more conserved. Exons 2 and 3 of the classical genes encode the peptide-binding region (PBR) and therefore contain the most variable site in the MHC I genes. It is the variability of the PBR which allows the presentation of a diverse array of peptides. All classical MHC I and the MHC-E molecules are expressed on the surface of nearly all cells. In contrast, MHC-G and -F expression is restricted to specific tissues, with MHC-G molecules limited mainly to trophoblasts, whereas

MHC-F is preferentially expressed in lymphoid tissues (Hughes & Yeager, 1998; Lafont *et al.*, 2003).

The class II MHC genes of eutherian mammals are arranged in a number of separate regions within chromosome six. In humans (where MHC is commonly named HLA for human leucocyte antigen), there are three classical regions including the polymorphic class II loci, designated DR, DQ and DP. Each region contains multiple A or B loci. In addition to these regions, the *-DMA*, *-DMB*, *-DNA* and *-DOB* non-classical genes also map to the MHC class II region (Bontrop *et al.*, 1999; Hughes *et al.*, 1997, Hughes & Yeager, 1998). In a number of other placental mammals, homologues of the human class II loci have been discovered, and these are usually named following their human homologues. The genes *HLA-DPA1*, *-DPB1*, *-DQA1*, *-DQB1* and *-DRB1* exhibit a high degree of allelic variation (Penedo *et al.*, 2005). The second exon of the DRB locus codes for parts of the functionally important antigen-binding sites (ABS) and is an analogue of the PBR of class I molecules. In a variety of species, the ABS is highly variable in both the overall number of alleles and the extent of sequence variation between alleles (Hughes & Yeager, 1998; Schwensow *et al.*, 2007; Sommer, 2005).

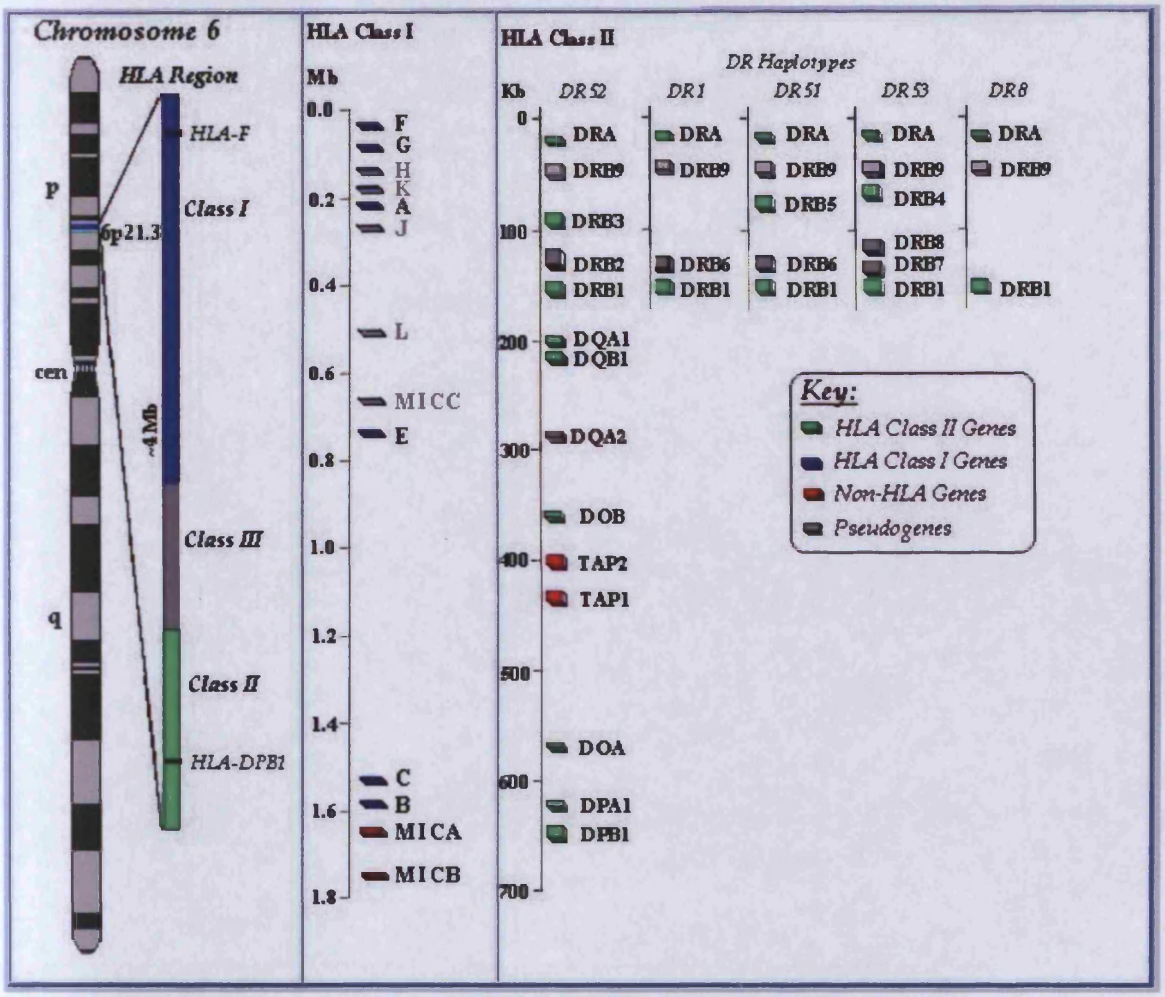


Figure 1.4. Vertebrate MHC map. Source: NCBI dbMHC Home. (http://www.ncbi.nlm.nih.gov/gv/mhc/main.cgi?cmd=init&user_id=0&probe_id=0&source_id=0&locus_id=0&locus_group=1&proto_id=0&kit_id=0&graphview=0)

MHC variability is believed to be maintained by pathogen driven selection, mediated either through heterozygote advantage (overdominance hypothesis) or frequency dependent selection (Red Queen hypothesis) (Doherty & Zinkernagel, 1975; Hughes *et al.*, 1997; Piertney & Oliver, 2006; Sommer, 2005; Schmid-Hempel & Koella, 1994). It has been suggested that populations with reduced MHC variability would be particularly vulnerable to disease, an issue of particular concern for endangered species living in small, isolated populations facing already a significant threat of extinction from exposure to pathogens and parasites (Schwensow *et al.*, 2007; Sommer, 2005; Bernatchez & Landry, 2003; Lukas *et al.*, 2004). However, relatively recent studies have addressed MHC polymorphisms in wild populations and/or attempted to test for an association between such polymorphisms and parasite resistance under natural conditions (Bernatchez & Landry, 2003; Lenz *et al.*, 2009; Lukas *et al.*, 2004; Meyer-Lucht *et al.*, 2010; Oppelt *et al.*, 2010; Schad *et al.*, 2004, 2005; Schensow *et al.*, 2007; Sommer, 2005; Tollenaere *et al.*, 2008).

1.4.3 Primate MHC studies

Most primate MHC studies have had an evolutionary focus and are used as models to study human diseases. In non-human primates, MHC class I genes have been identified and used to understand co-evolution with pathogens and selection of new populations of viruses (Vogel *et al.*, 1999). In apes and old world monkeys (OWM), homologues of the classical HLA-A and -B loci are polymorphic, with a high rate of non-synonymous substitutions in the PBR (ibid). Homologues of the non-classical I HLA-E and -F loci have also been identified in orang-utans and macaques. Recently, MHC-E orthologues were identified in gorillas, chimpanzees, bonobos and vervet monkeys (Knapp *et al.*, 1998). Class II regions DR, DQ and DP have been characterized in apes, old world and new world monkeys with the best characterizations made in the DR regions of chimpanzees, rhesus monkeys and cotton-top tamarins (Bontrop *et al.*, 1999). The evolutionary analyses indicate that these genes evolve much faster than class I genes (Bontrop *et al.*, 1999; Vogel *et al.*, 1999). Although most alleles of the DR region have human equivalents, studies suggest that rhesus monkeys have unprecedented polymorphism and this may reflect an alternative strategy of this species to cope with pathogens (Doxiadis *et al.*, 2000).

Studies on primate MHC variability are not only addressing evolutionary issues. Lukas *et al.* (2004) used DNA from faecal samples to characterize microsatellite loci as well as DRB exon 2 haplotypes in two populations of wild gorillas, and their results showed similar levels of variation at the MHC locus between the two species. A study of the DRB exon 2 of a subdivided mouse lemur (*Microcebus murinus*) population indicated that variation in MHC-allele frequencies in forest fragments was linked to parasite load. Futhermore, certain alleles, which differed in a few amino acids in the ABS from other alleles, were associated with parasite resistance or susceptibility (Schad *et al.*, 2004, 2005). In another study, a free-ranging fat-tailed dwarf lemur (*Cheirogaleus medius*) population was used as a model to investigate the role of neutral versus adaptive genetic variation in parasite resistance and to identify possible parasite-driven selection acting on the MHC under natural conditions (Schwensow *et al.*, 2007). In the same study they tested for associations between the MHC-DRB exon 2 supertype constitution and different measures of parasite burden. Although they found no associations between neutral overall individual genetic diversity and parasite load, evidence for a specific MHC supertype that was linked to infected individuals, a higher number of different nematode infections and high intensity of infection per individual was identified. Moreover, the study revealed that one rare MHC supertype was advantageous with respect to parasite burden (ibid).

MHC research in wild vertebrates could potentially reveal the effects of conditionally advantageous or deleterious alleles, discovered only in the presence of natural stress, such as spatially and temporally changes in climate, food availability, competition and parasitism. The studies of Shad *et al.* (2004, 2005) and Schwensow *et al.* (2007) suggest that MHC-variation might influence the long-term survival of small fragmented primate populations and indicate the functional importance of maintenance of MHC variability in declining or fragmented animal populations. Human impact on other vertebrate and parasite populations often causes a loss of genetic variation leading to short-term reduction of fitness and to an impaired ability to adapt to changing environments (Sommer, 2005). A major concern for demographically bottlenecked species is the maintenance of genetic diversity. Bottlenecks may reduce the amount of genetic variation which could be reflected in lower fitness, reduced potential for future adaptation and elevated extinction risk (Zhang *et al.*, 2006). Moreover, when direct intervention (reintroduction or translocation) has been planned to increase the size of a post-bottlenecked population (as is the case for orang-utans in the LKWS) the genetic variation of the source population must be evaluated to provide an indication for the long-term fitness of the reintroduced populations (ibid).

The mechanism of how diseases affect hosts can be investigated by monitoring host population genetic structure, particularly genes that are under selective pressure from parasites. For example, it has been suggested there might be a relationship between variations in the orang-utan's alpha-2 globin gene and malaria infection (Steiper *et al.*, 2006.) Due to their function and their suggested evolutionary mechanisms, MHC genes are ideal candidates for monitoring these kinds of relationships. The link between MHC and malaria is an often cited example of the influence of MHC genes on the course of a disease. It has been reported that certain MHC haplotypes seem to confer resistance to *Plasmodium falciparum* to children in West Africa (Hill *et al.*, 1991). In mice, the host's MHC genotype had a strong effect on *Plasmodium* parasitemia and on blood cell counts (Wedekind *et al.*, 2005). MHC Class I alleles have also been associated with increased resistance to malaria in wild bird populations (Bonneaud *et al.*, 2006: Westerdahl *et al.*, 2005). Yet few studies have examined variation in host genes in relation to susceptibility to infectious diseases in wild vertebrates. Previous studies have shown that specific MHC alleles could be associated with resistance or susceptibility to infectious diseases in the wild. Some examples are the resistance against strongyle nematodes in Soay sheep (*Ovis aries*) and the resistance of Atlantic salmon (*Salmo salar*) to *Aeromonas salmonicida* due to variations in MHC class II alleles (Langefors *et al.*, 2001; Paterson *et al.*, 1998). For conservation purposes, it is urgent to find out whether there are immunogenetic variants that may influence susceptibility to pathogens. There are still very few studies examining the genetic basis of resistance to naturally occurring pathogens in wild vertebrate populations or basal studies regarding the pathogens in themselves.

1.5 Parasites as a cause of wildlife declines

Habitat degradation and climatic conditions are crucial parameters in terms of distribution, transmission and developmental success of parasites. Wildlife diseases are powerful selective agents in natural populations (Altizer *et al.*, 2003) and their impact on primate populations is now recognized as a real extinction threat. For these reasons, it has been suggested that the ability to manage the spatial and genetic structure of host populations to minimize extinction risk by infectious diseases hinges upon better knowledge of host-parasite evolutionary dynamics in the wild (ibid). In the last two decades, research on wildlife emerging infectious diseases (EIDs) has shown the threat they pose to wild populations, especially if the animals are endangered (Chapman *et al.*, 2005a; Cleaveland *et al.*, 2001; Daszak *et al.*, 2000, 2001, 2004; Thompson *et al.*, 2009, 2010; Woodroffe, 1999). Examples of this are the mass mortality events and population declines in African primate fauna (Walsh *et al.*, 2003; Leendertz *et al.*, 2004; Leroy *et al.*, 2004). Control measures for wildlife EIDs have been attempted to prevent spread to humans or domesticated animals. However, it has been suggested that future research must adopt a multidisciplinary approach (veterinarians, medical workers, public health researchers, ecologists, conservation biologists, and others) to identify the underlying causes of EIDs and to control their spread (Daszak *et al.*, 2000; 2007).

Two major drivers of infectious disease emergence in wildlife have been identified, both of which are the result of human activities. The first is the "spill-over" of pathogens from domestic animals into wildlife populations and the other is the anthropogenic movement of pathogens into a new geographical location (Daszak *et al.*, 2001). The impact of other anthropogenic environmental changes on wildlife EIDs has not thoroughly been studied. Historically, wildlife diseases have been considered important only when agriculture, aquaculture or human health have been threatened. However, because of recent incidents of severe outbreaks of disease in endangered species, increasing veterinary involvement and advances in host-parasite population biology, the threat of wildlife disease is now taken more seriously (Daszak *et al.*, 2000; Pedersen, *et al.*, 2007; Smith, *et al.*, 2006, 2009; Thompson *et al.*, 2010; Woodroffe, 1999).

Although most pathogens have subtle and less easily quantifiable effects on their hosts, rather than disastrous impacts, their consequences are greatly exacerbated when hosts are stressed by other factors. It has been proposed that habitat alteration, destruction or fragmentation, or an increase in host density (e.g. in zoos or rehabilitation centres), are likely to impact on disease emergence (Lafferty & Gerber, 2002). Furthermore, the complex interactions between hosts, vectors and changing environments may have severe consequences for disease emergence

(Daszak *et al.*, 2001). Additionally, disease may cause extinction by killing hosts more rapidly than they can breed or by suppressing the size or growth rates of the populations, making them more vulnerable to stochastic factors (Woodroffe, 1999). Alteration of ecosystems can create conditions which facilitate the appearance and spread of new diseases. As an example, the creation of artificial water holes in Etosha National Park in Namibia resulted in repeated cases of Anthrax in large wild mammals (Williams *et al.*, 2002), and climate change might be a potential cause for the emergence of chytridiomycosis, responsible for amphibian population declines across Europe, America, Australia and New Zealand (Daszak *et al.*, 1999; Williams *et al.*, 2002). The interaction of wild and domestic species is another factor influencing EIDs. For instance, the extinction of Serengeti's wild dogs (*Lycaon pictus*) and black-footed ferrets (*Mustela nigripes*) were concurrent with epizootics of canine distemper in domestic dogs (Williams *et al.*, 1988; MacDonald, 1992). Movement of vectors can also contribute to the appearance of diseases. Hawaiian native bird species extinctions have been attributed to the introduction of avian malaria and mosquitoes to those islands (Atkinson *et al.*, 1995, 2000). Apparently, in the cases where populations were driven to extinction the size of the population was small and the diseases were caused by generalist pathogens with a wide host range (Woodroffe, 1999).

In the case of primates, there is clear evidence that effects of disease interact with habitat loss and other stress factors, with often catastrophic consequences. Populations of howler monkeys (*Alouatta palliata*) have been reduced by 50%, or even disappeared completely, as a consequence of yellow fever epidemics in Panama (Chapman *et al.*, 2005a; Nunn & Altizer, 2006). Similarly, an 85% decline in the size of a population of red howler monkeys (*Alouatta seniculus*) occurred due to an outbreak of an unidentified agent (Nunn & Altizer, 2006). *Mycobacterium bovis* infected a troop of free-ranging chacma baboons (*Papio ursinos*) with a prevalence of up to 50% and a rapid progress of the disease (Keet *et al.*, 2000). Both gorilla (*Gorilla gorilla*) and chimpanzee (*Pan troglodytes*) population declines in Central Africa have occurred due to Ebola and/or Anthrax outbreaks. Diseases, combined with the slow reproductive cycle of these apes, and hunting and poaching threats, may lead to the extinction of these hosts in western Central, and Equatorial, Africa (Leroy *et al.*, 2004; Walsh *et al.*, 2003; Leendertz *et al.*, 2004, 2006). Moreover, in general, as anthropogenic habitat change forces humans and primates into closer and more frequent contact, the risks of interspecific disease transmission might increase (Chapman *et al.*, 2005a). The main routes of transmission of human diseases to apes seem to be respiratory and faecal-oral (Woodford *et al.*, 2002). Chimpanzees of the Gombe National Park in Tanzania, have suffered from a polio-like virus and respiratory outbreaks. Although it was impossible to determine whether the disease was transmitted from humans, poliovirus was widespread in the local human population at that time, and so were colds,

influenza and pneumonia (Wallis & Lee, 1999; Nunn & Altizer, 2006). Gorillas and chimpanzees have also suffered from scabies epidemics. Again, the source of the outbreak was not confirmed but it is thought to have arisen from contact with humans, as this is a common disease in local populations (Wallis, 2000; Wallis & Lee, 1999). A measles outbreak in the Virunga Mountains, Rwanda, caused the death of six female gorillas before a vaccination program was established (Wallis & Lee, 1999; Nunn & Altizer, 2006). It is thought primates living in natural habitats are free of measles, but are highly susceptible to transmission from humans (Wallis & Lee, 1999). In olive baboons (*Papio hamadryas anubis*), yaws (a public health problem in Africa) was spread as a venereal disease, and *Schistosoma mansoni* showed the highest prevalence in populations that had greatest contact with people (Wallis & Lee, 1999; Wallis, 2000). Other studies have demonstrated a greater prevalence and richness of gastrointestinal parasites in red-tail monkeys (*Cercophithecus ascanius*) in logged compared to undisturbed forest (Chapman *et al.*, 2005b). It was also found that *Ascaris* spp. and *Giardia* spp. occurred in red colobus monkeys in forest fragments near human populations with high prevalence of these parasites; the same parasites were not found in monkeys dwelling in pristine areas where people and primates barely interacted (ibid).

There have been some attempts to reduce or avoid the disease risk in wildlife. Likewise, it has been noticed that despite the negative effects infectious diseases have over wildlife, there is a general lack of information about the pathogens that cause them (Daszak *et al.*, 2001; Nunn & Altizer, 2006). Increased knowledge of host-parasite interactions might be the key for the development of protocols to track pathogens implicated in wildlife declines and minimize disease risks (Daszak *et al.*, 2001; Lafferty & Gerber, 2002; Nunn & Altizer, 2006; Smith *et al.*, 2009; Wolfe *et al.*, 1998).

Lack of information on the occurrence and prevalence of parasites in wildlife populations have made it difficult to assess the effectiveness of past control measures and to design new measures (Smith *et al*, 2009; Thompson *et al.*, 2010, Woodroffe, 1999). Therefore, the monitoring of diseases and the storage and use of that information is essential. Surveys of wildlife parasite and pathogen biodiversity are an essential tool for identifying the agents of wildlife EIDs and predicting their future emergence (Daszak *et al.*, 2001). There is also a need to screen for diseases in wild and domestic species that may act as reservoir hosts (Cleaveland *et al.*, 2002, Thompson *et al.*, 2009, 2010). General research is also needed into the dynamics of pathogens infecting multiple hosts, and the consequences this has for disease control (Woodroffe, 1999, Cleaveland *et al.*, 2002). For instance, understanding the transmission mechanisms of different parasites might be fundamental for disease control, for example, some vector–borne diseases may be impossible to eliminate without controlling the vectors (Nunn & Altizer, 2006). In

summary, possibly the greatest need in the context of conservation and infectious disease, is to increase the knowledge about host-parasite interactions in natural systems.

To our knowledge, there have been no studies regarding parasite prevalence or abundance in long-tailed macaques or proboscis monkeys in the Kinabatangan. Furthermore, according to the Global Parasite Database (Nunn & Altizer, 2006), there are no reports of helminths in proboscis monkeys in Malaysia, and just a few reports of these parasites in long-tailed macaques but not for the island of Borneo, perhaps because they have not been investigated. As we will be using molecular scatology for microsatellites and MHC analyses, baseline information regarding the gastrointestinal parasite diversity of both primate species with also be gathered, taking full advantage of the available specimens.

1.6 Non-invasive sampling in conservation genetics

Genetic data provide a very useful tool to elucidate the dynamics and social organization of wild populations and also yield invaluable information for species conservation. Formerly, genetic studies were based on blood or tissue samples but obtaining such samples from living wild animals not only involves a risk for the animals, it is also time consuming and difficult (Lathuilliere *et al.*, 2001). Traditionally, sampling was destructive and led to the animal's death, or was non-destructive but the animal had to be captured nonetheless (Taberlet *et al.*, 1999). Advances in molecular biology now allow amplification of tiny amount of DNA and consequently non-invasive sampling techniques are being increasingly used as they produce genetic results comparable with those from blood or tissue samples (Bayes *et al.*, 2000; Lathuilliere *et al.*, 2001). In non-invasive sampling the source of the DNA is left behind by the animal and is an attractive method to field biologists because it allows genetic studies of free-ranging animals without having to catch, handle or even observe them (Taberlet *et al.*, 1999; Broquet *et al.*, 2007). The source of DNA can be shed hairs or feathers, faeces, urine, buccal cells from food wedges, skin and even bones and nails (Taberlet *et al.*, 1999, Piggott & Taylor, 2003, Goossens & Bruford, 2009). The analysis of non-invasive genetic samples can provide individual identification, relatedness estimates, pedigree reconstruction, sex identification, estimates of census and effective population size, and the level of genetic polymorphism within or between populations (Goosens & Bruford, 2009; Goossens *et al.*, 2003b; Mills *et al.*, 2000; Taberlet *et al.*, 1999). Non-invasive sampling of DNA has been proposed or carried out for bears, humpback whales, coyotes, and a suite of mid-level carnivores in North America, as well as primates, tigers, rhinoceros, elephants and pandas, just to mention a few (Arandjelovic *et al.*, 2009; Bhagavatula & Singh, 2006; Garnier *et al.*,

2001; Johnson *et al.*, 2007; Kohn & Wayne, 1997; Mills, 2007; Zhan *et al.*, 2007; for a more exhaustive review see Goossens & Bruford, 2009).

1.6.1 Molecular scatology and its challenges

Molecular scatology has proved to be a powerful technique to study the genetics of free-ranging mammal populations (Goossens *et al.*, 2003b; Kohn & Wayne, 1997; Taberlet *et al.*, 1999). It is a convenient method especially for the study of primates and other arboreal species, where the necessity of climbing trees to collect hairs and the risk of mixing samples from different individuals in a nest pose a risk for researchers and the study per se (Goossens & Bruford, 2009). Consequently, collecting faecal samples on the floor under nests seems to have an advantage over hair since they are often abundant and the individual can also sometimes be identified. Faeces contain cells shed from the intestinal lining, thus DNA from the host itself can be isolated and analyzed. The information faeces can provide is diverse: number of offspring produced by individuals and the relationship of individuals in a social group to each other, estimates of population size, approximate dimensions of territories, assessments of effective population size, extent of genetic variation, relationships of populations to each other and the degree of gene flow that occurs between them, and a more precise identification of food and pathogen species (Broquet, *et al.*, 2007; Kohn & Wayne, 1997; Piggot & Taylor, 2003; Taberlet *et al.*, 1999). Furthermore, faecal samples are not subject to CITES constraints (Goossens *et al.*, 2000). Despite its applications, several authors have underlined the risk of misinterpretations owing to the very low quantity and quality of DNA coming from faecal samples, which may yield spurious results (De Barba & Waits, 2009; Goossens *et al.*, 2000; Goossens *et al.*, 2003b; Lathuilliere *et al.*, 2001; Pompanon *et al.*, 2005; Taberlet *et al.*, 1996, 1999). Broquet *et al.* (2007) mention that less DNA (1ng) may be extracted per single hair than from faeces (e.g. 38.4 ng per extract on average) and that amplification success and genotyping errors have been found to be sensible to template DNA concentration. These errors are principally: (i) the revelation of only one allele in the case of heterozygous individuals (allelic dropout, ADO); and (ii) the revelation of additional allele(s) (false allele, FA) problematic in the case of homozygous individuals (Valière *et al.*, 2002). Still, estimates for mean per locus error rates using faeces have ranged from 0.05 to 0.29 compared to 0.14 to 0.35 in microsatellites amplified from hair extracts (Roon *et al.*, 2005). However, using appropriate measures very accurate results may be obtained (Arandjelovic *et al.*, 2009; De Barba & Waits, 2009; Goossens, 2000; Morin *et al.*, 2001; Roon *et al.*, 2005; Taberlet, 1996; Taberlet *et al.*, 1999).

Multiple methods have been used to extract faecal DNA. Currently the most commonly used method for extracting DNA from faecal samples is silica-binding extraction kits (Qiagen)

(Goossens & Bruford, 2009; Waits & Paetkau, 2005). A major limitation in this step is the quantity and quality of the extracted DNA. Usually the total amount of DNA is often in the picogram range, and it might be degraded or accompanied by PCR inhibitors (especially in faeces that contain plant remnants) (Broquet *et al.*, 2007; Little, 2003). Furthermore, this low quantity DNA may be damaged due to exposure to moisture, heat, and ultraviolet radiation encountered during non-invasive sample collection (McKelvey & Schwartz, 2004). Besides, contamination from humans (particularly for primate species) and cross contamination between samples are common during the extraction and amplification process (Goossens & Bruford, 2009; Taberlet *et al.*, 1999). However, the quality of the samples can be greatly improved by collecting them just after the animal leaves them behind. If the samples remain in the field several weeks before collection, then the DNA could become more degraded and more difficult to amplify (DeBarba & Waits, 2009). Contamination and degradation, however, can be minimized by carefully conduced collection and preservation techniques; contamination can also be minimized by focusing on laboratory cleanliness and good lab technique (Goossens & Bruford, 2009; Goosens *et al.*, 2003b; Taberlet *et al.*, 1999; Waits & Paetkau, 2005).

Although working with faecal samples can be very demanding, it has a number of advantages. It can increase the number of animals that can be sampled in secretive species, and thus make feasible the estimation of important population parameters. It also offers a better way of studying demographic characteristics of rare or hard-to-capture species, which include avoiding stress and contact with endangered populations. Finally, it is convenient in situations in which capturing the individuals of interest is not conceivable (for example in the case of a small endangered population, or in behavioural studies where capturing animals would disturb the system) (Goossens & Bruford, 2009; Mills, 2007; Taberlet *et al.*, 1999).

The study of endangered species requires approaches that minimize the handling of animals for acquisition of samples. Molecular scatology is ideal to study endangered species because being protected by law, invasive sampling is not allowed. Moreover, the use of faecal samples has proved to give excellent results in a number of species, including primates such as gorillas, chimpanzees, baboons and langurs (Bayes *et al.*, 2000; Bradley *et al.*, 2000; Launhardt *et al.*, 1998; Little, 2003; Morin *et al.*, 2001). All the primate genetic studies carried out in the LKWS have been done through non-invasive collection of samples with very successful results (Goossens *et al.*, 2005; 2006b). In particular, the mtDNA studies on proboscis and macaques were performed on DNA extracted from faecal samples (Jalil, 2007), hence this approach was selected to conduct the present study.

1.7 Hypotheses

Previous population genetic analyses of primates in the LKWS have focused on neutral (demographically mediated) genetic variation. In the case of orang-utans, mitochondrial and nuclear (microsatellites) DNA studies have clearly shown genetic differentiation between populations on either side of the river (Goossens *et al.*, 2005; Jalil *et al.*, 2008). Population differentiation was also observed for long-tailed macaques (Jalil, 2007), however this species is known to be a good swimmer so this pattern is probably a result of differences among social groups between populations. In contrast, proboscis monkeys' showed no structuring, which correspond very well to their dispersal behaviour and swimming ability (Jalil, 2007).

As the long-tailed macaque and proboscis monkey studies were carried out using a maternally inherited marker, bi-parentally inherited markers might be expected to reveal a different population structure, especially in the case of long-tailed macaques. If contemporary patterns of neutral variation (microsatellites) depend primarily on dispersal, in species with female biased dispersal, estimates of population structure derived from microsatellite data analyses are expected to be similar to that obtained from mtDNA analyses. Additionally, the genetic signatures of natural selection may be superimposed on the signatures of drift and/or gene flow (features that affect neutral variation). Therefore, a modified population structure pattern could be expected from adaptive genetic variation compared to neutral variation, especially neutral nuclear DNA (microsatellites). MHC genes are under localised selection while microsatellites evolve by the means of (mainly) genetic drift. If the primate populations in the Kinabatangan have undergone recent habitat-mediated contraction, the separation of populations might be too short to leave a signal at neutral loci so that differences between populations will only be detectable at genes under selection.

Habitat loss has been suggested to change the behaviour and abundance of wildlife which in turn affects parasite transmission and distribution (Chapman *et al.*, 2005a, b, 2006a, b; Gillespie *et al.*, 2005b; Hudson *et al.*, 2006; Nunn *et al.*, 2003). Furthermore, as human population density continues to increase, speeding the reduction and fragmentation of primate habitats, greater human-primate contact is inevitable and even higher rates of parasite transmission are likely between humans and monkeys and between monkey species (Gillespie, 2006; Goldberg *et al.*, 2008). When moving between forest patches and because of the proximity to human settlements and to domestic animals, primates may be exposed to a wider range of parasitic vectors and/or intermediate hosts (Trejo-Macías *et al.*, 2007). Therefore, parasite diversity and frequency of co-infections could depend on the primate host species. For instance, in the LKWS, although the proboscis monkey and the long-tailed macaque share habitat (several groups of primates were

observed to use the same sleeping trees day after day – Goossens, *et al.* 2003a), the macaques are known to come closer to human habitation and to domestic fauna. This behaviour might increase their chances of having higher parasite diversity and harbouring a mixed infection. Additionally, the fragmentation of the habitat might impact the parasite species richness, the proportion of individuals carrying a mixed infection and the prevalence of particular parasites.

All these conditions allow the following hypotheses to be tested:

I) In long-tailed macaques, which exhibit male- biased dispersal, the pattern of population structure is expected to be less pronounced than that obtained by previous mtDNA analysis. In contrast, the female-transfer behaviour of the proboscis monkey is expected to reveal a pattern of genetic structure similar to that reported by Jalil (2007) from mtDNA.

II) A more pronounced spatial genetic structure between river sides or forest fragments (for instance, higher F_{ST} value) is expected when analysing MHC (linked to positive selection), than when analysing microsatellites (linked to gene-flow).

III) Long-tailed macaques are expected to have higher parasite diversity and to harbour more mixed infections than proboscis monkeys.

IV) Both primate species are known swimmers, thus no effect is expected from the natural barrier presented by the Kinabatangan River. Contrarily, it is predicted that in areas of the sanctuary with higher rates of human activities, the parasite richness, the frequency of co-infections and the prevalence of particular parasites will be higher in both primate species, compared to areas with less human activity.

As the LKWS is bisected along its length by the Kinabatangan River, this site provided the opportunity to study the impact of this natural potential geographical barrier on the population structure of the two primates and their parasitic fauna. This project supplements the genetic datasets for proboscis monkeys and long-tailed macaques by genotyping microsatellites which in the former had not been previously attempted (Chapters Two and Three). Technical difficulties prevented comparison of neutral and adaptive genetic diversity and potential differences in spatial genetic structuring with the two types of markers. Nonetheless, MHC genes were identified in both species as a first step towards the resolution of these questions (Chapter Four). Moreover, this study has established baseline data regarding the endoparasite (helminth) diversity in the two primate species by measuring richness, co-infections and prevalence; these factors were then correlated to fragmentation (natural and anthropomorphic-mediated) of the habitat (Chapter Five). In the final Chapter (Six), the data from Chapters Three, Four and Five are assimilated and

discussed in the context of a conservation strategy for these primates in the LKWS. Each chapter is self contained, and Chapter Two has been published in Conservation Genetic Resources (Salgado-Lynn, *et al.*, 2010).

CHAPTER TWO

Microsatellite markers for the proboscis monkey (Nasalis larvatus)

Salgado-Lynn M[1], Stanton DWG[1], Sakong R[2], Cable J[1], Goossens B[1,3], Bruford MW[1]

[1]*School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK*
[2]*Red Ape Encounters, Kinabatangan Orangutan Conservation Project, PO Box 3109, 90734 Sandakan, Sabah, Malaysia*
[3]*Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia*

This chapter was published (August 2010) as a technical note in Conservation Genetics Resources: therefore it is presented in a slightly different format compared to the other data chapters.

Abstract

We describe eight polymorphic microsatellite loci for the proboscis monkey (*Nasalis larvatus*). These markers were tested with 33 samples, collected from Sabah and exhibited a mean of 6.25 alleles per locus and a mean expected heterozygosity of 0.674. All but one locus were in Hardy-Weinberg equilibrium, and no evidence for linkage disequilibrium was detected between any loci. Another 30 loci were isolated but remain to be fully examined. These markers should be useful for the future study of population genetic diversity and genetic structure in this emblematic species.

Classified as endangered by IUCN (2010) and listed in Appendix I of CITES (UNEP-WCMC, 2010), the proboscis monkey (*Nasalis larvatus* van Wurmb 1787) is endemic to the island of Borneo. Its distribution is restricted to lowland coastal and riverine forests, mangrove, and peat swamp. With a declining population, major threats include hunting, fire and, most importantly, anthropogenic habitat loss and fragmentation (Meijaard & Nijman, 2000; Sha *et al.*, 2008). Despite its uniqueness and conservation status, limited research on genetic variation in proboscis monkeys has been carried out due to a lack of reliable genetic markers and challenging sample collection (Jalil, 2007). Here, we describe the isolation and characterization of microsatellite markers which can be used for individual and population-level genetic analyses, suitable for both invasive and non-invasive samples, for the conservation of this species.

Muscle samples were opportunistically collected from two deceased proboscis monkeys at Lok Kawi Wildlife Park, and from two road killed animals, in Sabah, Malaysia; the former were used for the isolation of microsatellite loci. Faecal samples from another 29 wild individuals, also from Sabah, were collected as part of a population study and were used for characterizing the markers along with the tissue samples. Stool samples were stored in 70% ethanol, and muscle samples in a -70°C freezer. Faecal DNA was extracted via the DNA Stool Mini Kit (Qiagen GMBH, Germany) using a previously described protocol (Goossens *et al.*, 2005). Tissue samples were extracted with DNeasy Blood & Tissue Kit (Qiagen GMBH, Germany) following the recommendations of the manufacturer, with minor modifications during elution (namely, 5 min incubation at 70°C with buffer AE, which was also preheated at the same temperature). To verify the existence of primate DNA from faecal extracts, a partial mitochondrial Control Region fragment was amplified using species specific primers (Jalil, 2007) while DNA from the muscle samples was visualized in agarose gels (1.5%) and quantified by spectrofluorometry (Invitrogen's Quant-iT™ PicoGreen® Kit microtiter assay, Molecular Devices' SOFTmax Pro®).

Genomic libraries were constructed based on the protocol from Glenn and Schable (2005). DNA was digested overnight with RsaI (New England BioLabs) and the products were subsequently ligated to Super SNX24 linkers. Linker-ligated DNA was electrophoresed in a 1.5% agarose gel and fragments between 300 and 800 bp were electroeluted, precipitated with 3M NaOAc-ethanol and resuspended in TE Buffer (Bruford *et al.*, 1992). Fragments containing microsatellites were captured using biotinylated oligonucleotides (see mix 2 in Glenn & Schable, 2005), and the biotinylated probe-DNA complex was enriched by hybridization to streptavidin-coated magnetic beads (Dynabeads M-280, Invitrogen). Nonspecific DNA was removed by subsequent washes with SSC-SDS buffers as described in the same protocol, and recovery was performed by PCR using the forward SuperSNX-24 primer. Enriched libraries were constructed using a TA Cloning Kit according to the manufacturer's protocol (Invitrogen) and positive colonies were amplified using universal M13 forward and reverse primers (M13F: 5'-GTAAAACGACGGCCAG-3'; M13R: 5'-CAGGAAACA- GCTATGAC-3'). Fragments between 500 and 1200bp were sequenced using the BigDye terminator kit v1.1 (Applied Biosystems. Sequences were assembled and edited in Mega 4.0 (Tamura *et al.*, 2007) and visually checked for microsatellite repeats. Two libraries were constructed and a mean of 33% of sequenced clones yielded microsatellite motifs. Unique sequences with sufficient flanking DNA and at least five (trinucleotide) or ten (dinucleotide) repeat units were selected for primer design using Primer 3 in msatcommander (Faircloth, 2008). Ninety one primer pairs were designed with melting temperatures between 50 and 66°C and length of PCR products between 100 and 300 bp.

The tissue samples were used to optimize the PCR conditions for 46 unlabelled primer pairs. PCR reactions consisted of 1X Master Mix (QIAGEN, Multiplex Kit), 0.4μg/μl BSA (New England BioLabs), 0.2 μM of each primer, and 200 pg of template DNA to a total volume of 10 μl and were performed in a GeneAmp® PCR System 9700 (Applied Biosystems). The amplification conditions were as follows: 95°C for 15 min, 45 cycles at 94°C for 30 s, 48-63°C for 90 s, 72°C for 90 s and a final extension at 72 °C for 30 min. Thirty eight primers, which produced a single target band, were further optimized by amplifying 29 dung samples with a multi-tube approach (Taberlet *et al.*, 1996) using the above PCR protocol but with 2ul of DNA template and the optimized annealing temperatures (T_a) shown in Table 2.1.

Table 2.1. Primer and motif sequences of isolated microsatellite loci for the proboscis monkey. Sizes are based on the sequenced allele.

Locus	*Primers (F-forward; R-reverse)*	*Motif*	*T_a(°C)*	*Size (bp)*	*GenBank ID*
***NlB5**[a]*	F-CCCATCACCTCATGTAGTTACC R-CCTGAAATTTGCTAAGGGAGT	$(GT)_{10}$	57	107	HM588998
NlD10	F-TGTCCTTCTCCACTGCCTCT R-TGCAATTTCATTACACCAATGAT	$(CT)_{10}$	54	209	HM588999
NlD5	F-TGATTTTGCTCTCACCCTTG R-CCGATTCTCTGTTGGAGGAA	$(CA)_6AA(CA)_6TG(CA)_6$	54	171	HM589000
NlE10	F-CCATCACACCTGGCTGCTTA R-ATGCCTTGTTGGGAAGACAG	$(GT)_{16}$	58	169	HM589001
NlF1	F-GCCAATGTTGTAAACTCTATACCC R-TTTATCAACCTGGCCTTTGA	$(AC)_{10}AT(AC)TC(AC)_8$	52	176	HM589002
***NlG8**[a]*	F-GGAAATCCAAAGCCTACTGC R-CAGGAAATGTGAAATGGAGGA	$(GT)_{14}$	54	226	HM589003
NlP1A6	F-TCTCACTGGTAAAGAAATGTGGA R-CGGACTCTCTGGCTTTTCAG	$(AAC)_7$	54	158	HM589004
NlP1C3	F-CGACCCTCCAGGTTTAAGTG R-ACGCTTGTAATCCCACCTTG	$(GAT)_{10}$	54	229	HM589005
NlP1C5	F-AGGCCACTGAAGGCTGTCTA R-TGAGTCTAGCTTGGGCAACA	$(GT)_{15}$	54	204	HM589006
***NlP1C8**[a]*	F-CCAAATGGTTATTTTGCGAGA R-TTTTGGAAACACCAAAAATGG	$(GT)_{20}$	60	221	HM589007
***NlP1E9**[a]*	F-GCTGGCCTGCATACTCAAAT R-CAGACCAGTAGGGGGAGACA	$(GT)_{19}$	54	218	HM589008
***NlP1F2**[a]*	F-TGCAGTGAACCTAAACCTGCT R-CTCTGACTTGTGCCAGTGGA	$(AG)_{17}$	58	240	HM589009
NlP1F5	F-CCTATCACTTCAAGGGCATAAAA R-TGGCTTGGAGATGCATTTATT	$(GT)_{16}$	52	234	HM589010
***NlP1G7**[a]*	F-GGAGCTGGTGCTTCTACAGG R-GGCACCATAGCTTTCTATTCAA	$(GTT)_9$	52	160	HM589011
***NlP2B8**[a]*	F-GAGGTGGTCAGCTGGTCATAA R-GTGCACTGGCTCACTCATGT	$(GT)_{20}$	56	234	HM589012
NlP2B9	F-CGATTGAGTTCAGGTATCTTTTG R-TTCAATAATGATGGAAGAATACCG	$(GT)_{17}$	52	209	HM589013
NlP2C12	F-CCACAAAACACCATCTCCAA R-TGCTTCATGTCAAGGGATTG	$(AC)_{16}$	52	191	HM589014
***NlP2C5**[a]*	F-TCCTTTTGAATTGCCAAGTTTT R-AAGGCACCATGGTCTCAAAG	$(AC)_{21}$	58	184	HM589015
NlP2D6	F-AGGGGAAAACACATTTGCAG R-TTTTCCACTCCTCGTTTTGG	$(GT)_{17}$	54	160	HM589016
***NlP2E2**[a]*	F-TTGAGGCCTACCTGGTCAAC R-GCACTGAATTGCATCCAGAA	$(CT)_{20}$	60	247	HM589017
NlP2F3	F-CAGAACATTTTGCCCAACAG R-GTGGGCAGAAAAGAGAATCG	$(AC)_{21}$	58	191	HM589018
***NlP2F7**[a]*	F-CATTCAGACTCACTGGATTAAAAA	$(AG)_{17}$	56	150	HM589019

Locus	*Primers (F-forward; R-reverse)*	*Motif*	T_a(°C)	*Size (bp)*	*GenBank ID*
	R-AGATAGAGCCAGAACCTTTCCA				
NlP3A12[a]	F-CTGTGGCCAAACAGTTCATC	$(GT)_{17}$	50	246	HM589020
	R-CAGCAGTGGTTTTATTCATTTTTG				
NlP3B2	F-GCAATTTTGCTGAATTTGCTC	$(GT)_{23}$	58	165	HM589021
	R-GGCATCGAATTGAAAAGGAA				
NlP3B4[a]	F-TTCCAGCTATCAAAATAGTGGCTA	$(AC)_{18}$	54	185	HM589022
	R-GAAGTGGCTTGCCTTACAGC				
NlP3B6[a]	F-ATCATTTCTGGGCCTGTTTT	$(AC)_{22}$	58	237	HM589023
	R-CCTGCGGAACAAGAGTGAA				
NlP3C11[a]	F-TCCATCCCCTTTTTATGATACTT	$(AC)_{20}$	54	199	HM589024
	R-AGGTATGCAGCCAAGCAAAG				
NlP3E1[a]	F-ACTGGGCATCAGAGTCATCC	$(GAT)_{10}$	54	208	HM589025
	R-TCCATGCAATGGCACATAGT				
NlP3E7	F-GGAGAGGTGGCCTTTGAACT	$(GAT)_5$	52	151	HM589026
	R-TGTTCAGCAAACAATATAGAGCTAA				
NlP3E8[a]	F-CAAATGAAAAATGCCTCTAACAGA	$(ATCT)_{12}$	56	223	HM589027
	R-CAGAGCATGCAAGAAAGAGAGA				
NlP3G2	F-TCCCATGTTTATTGCAGCAC	$(AC)_{23}$	58	250	HM589028
	R-TCAGTGCCTGGCTGATTTTA				
NlP3H5[a]	F-CATTGTGAGAAAACTTGCTTCTG	$(AG)_{16}$	50	167	HM589029
	R-CCCATCAACTTCAGAATACACA				
NlP4B1	F-TTCCATGGTTTCCAGAGTCC	$(AC)_{16}$	58	158	HM589030
	R-AGAAATGGATGGGGCAGAG				
NlP4B2	F-TCAGGTGAATTGCTGGCATA	$(GT)_{15}$	54	199	HM589031
	R-GCATCCAAACTGGAAAGGAA				
NlP4C11	F-CTCCACAGTCCTGTGACCAA	$(GT)_{23}$	54	248	HM589032
	R-TGCAGAAAGCCAAAAGGATT				
NlP4C6	F-TGTTGAAAATTCTTGCATTTGTG	$(GTT)_7$	52	160	HM589033
	R-TCTTCCCCAAAACTGAGGAT				
NlP4E10[a]	F-CAAACCTGCATGTTCTCCAC	$(AC)_{20}$	56	218	HM589034
	R-GCTGGGAAACAATCAGTCCT				
NlP4E6	F-CAAGGAAGAATTGTGCCAAGA	$(GT)_{19}$	54	234	HM589035
	R-GGCATTCCCAAACCTCATAA				

[a] Denotes poor amplification in faecal DNA extracts

Twenty primers which successfully (≥80%) amplified in faecal samples were fluorescently labeled (5'- FAM, HEX, or NED) and assembled in 4 multiplexes. Although amplification of the tissue samples with the multiplex systems was successful (100%), consistent amplification of the dung samples with the multiplex approach was not possible. The limited volume of the faecal DNA extracts allowed thorough characterization of only eight markers (singleplex) with the 29 dung samples and the four tissue samples under the conditions previously described. For genotyping, PCR products were electrophoresed along with GeneScan ROX 350, or GS-400 HD LIZ, in a Prism 3700 Genetic Analyzer (Applied Biosystems) and fragment lengths were scored using GeneMapper® ID 3.2 (Applied Biosystems).

Table 2.2 Characteristics of 20 proboscis monkey microsatellite loci suitable for invasive and non-invasive genetic studies. Genetic diversity of eight microsatellite loci is based on genotypes of 33 individuals. A-Number of alleles per locus; "–" refers to markers tested only with four tissue samples.

Locus	*Label*	*A*	*Range (bp)*	H_O	H_E
Multiplex 1 (T_a 54°C)					
NlD10	HEX	6	177-189	0.848	0.581
NlP4C11	HEX	6	241-259	0.757	0.786
NlP1C3	FAM	1	230	-	-
NlD5	FAM	2	163-169	-	-
Multiplex 2 (T_a 54°C)					
NlP2D6	HEX	5	146-160	0.636	0.655
NlP1A6	NED	4	146-158	0.575	0.581
NlP1C5	FAM	7	177-205	0.696	0.685
NlP4B2	HEX	2	185-201	-	-
NlP4E6	FAM	3	237-241	-	-
Multiplex 3 (T_a 58°C)					
NlE10	NED	8	153-205	0.484	0.651
NlP2F3	FAM	5	175-187	0.575	0.662
NlP3B2	FAM	4	158-172	0.727	0.793
NlP4B1	HEX	3	152-156	-	-
NlP3G2	HEX	3	200-204	-	-
Multiplex 4 (T_a 52°C)					
NlP3E7	NED	1	152	-	-
NlP2B9	HEX	3	212-218	-	-
NlP4C6	HEX	2	156-162	-	-
NlP1F5	FAM	2	235-237	-	-
NlP2C12	FAM	4	186-194	-	-
NlF1	HEX	3	171-175	-	-

Allele diversity and size ranges of all 20 markers are included in Table 2.2, along with other details for the eight loci. Exact Hardy-Weinberg probabilities were assessed, and linkage disequilibrium was tested using GENEPOP version 4.0.10 (Raymond & Rousset, 1995; Rousset, 2008). Significance levels were adjusted using Bonferroni corrections for multiple testing ($P < 0.006$ in our dataset). All loci were in Hardy-Weinberg equilibrium except NlD10 ($P < 0.001$), and no evidence was found for linkage disequilibrium between any pair of loci. Observed and expected heterozygosities were calculated using ARLEQUIN version 3.5.1.2 (Excoffier & Lischer, 2010). The mean expected heterozygosity was 0.674 (range 0.581-0.793) and mean observed heterozygosity was 0.662 (range 0.484-0.848). The mean number of alleles was 6.25 (range 4-9) for the eight fecal-tested markers and 2.41 (range 1-4) for the other 12 loci. Deviation from Hardy-Weinberg equilibrium for some loci might be due to samples coming from three distant sites and the wildlife park individuals whose origin was unknown. It is also possible that balancing selection (i.e. heterozygote advantage) is acting on locus NlD10 or that this microsatellite is near a region under selection pressure. Large-scale testing on individuals from a single population is

needed for locus NlD10. The eight polymorphic microsatellite loci will be useful for the future study of individual identification, population genetic diversity and genetic structure in the proboscis monkey.

CHAPTER THREE
Genetic diversity and gene flow in two sympatric primate species of the Lower Kinabatangan Wildlife Sanctuary

***Salgado-Lynn M*[1], *Cable J*[1], *Bruford MW*[1], *Goossens B*[1,2]**

[1]*School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK*
[2]*Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia*

3.0 Abstract

The genetic diversity and the potential effect of the Kinabatangan River on the population structure of the proboscis monkey (*Nasalis larvatus*) and long-tailed macaques (*Macaca fascicularis*) in the Lower Kinabatangan Wildlife Sanctuary, northern Borneo, were examined. Cross amplification of 15 human microsatellite primers in 109 long-tailed macaque faecal samples revealed high levels of genetic diversity (mean H_E= 0.8) and gene flow (F_{ST}= 0.005). Sixty seven proboscis monkey stool samples were screened with eight species-specific microsatellite markers and the levels of genetic diversity and gene flow were also relatively high (mean H_E= 0.68, F_{ST}= 0.012). Bayesian clustering analyses revealed no influence of the Kinabatangan River on the population structure of either species, in accordance with the dispersal behaviour observed in both species. Significant departures from Hardy-Weinberg Equlibrium were detected particularly for long-tailed macaques (mean F_{IS}= 0.3) which can be explained by the presence of null alleles across the loci screened, which could not be eliminated by PCR optimisation. The results of this study can nonetheless be used as a baseline for conservation and management measures for the proboscis monkeys and long-tailed macaques of the LKWS.

3.1 Introduction

Genetic diversity plays a strategic role in the maintenance of adaptive evolutionary potential and the reproductive fitness of species (Allendorf & Luikart, 2007; Frankham *et al.*, 2002). Hence, for the last 20 years, the International Union for Conservation of Nature (IUCN) has included genetic diversity as a conservation priority in endangered species management plans. A threat to genetic variation can arise from restricted gene flow caused by habitat fragmentation. Moderate to high rates of gene flow among subpopulations prevent inbreeding depression, thereby preserving genetic variation. Hence, for the management of endangered species living in fragmented habitats it is essential to understand the role of gene flow, or migration of individuals and the subsequent transfer of genes among subpopulations (Allendorf & Luikart, 2007; Frankham *et al.*, 2002; Storfer, 1999). The Lower Kinabatangan Wildlife Sanctuary (LKWS), in Sabah, Malaysia,

is an example of a fragmented habitat harbouring endangered species. Renowned for its rich primate diversity, the sanctuary was created as a response to the logging and change of land use in the region (Pang, 2005; WWF – Asia/Pacific, 2005). Besides the anthropomorphic-mediated habitat fragmentation (ten forest fragments –Lots), the sanctuary is bisected lengthways by the Kinabatangan River, posing another potential (natural) barrier for the dispersal of wildlife, and consequently for gene flow.

Landscape genetics, an amalgamation of population genetics and landscape ecology, aims to provide information on how landscape and environmental factors influence gene flow and population structure (Manel *et al.*, 2003; Storfer *et al.*, 2007). In that context, the role of riverine barriers in speciation in the tropics, and especially with reference to primate diversity, has a long history of study and has been both heavily advocated and criticised over the past twenty years (Anthony *et al.*, 2007; Eriksson *et al.*, 2004; Gascon *et al.*, 2000; Gonder *et al.*, 1997, 2006; Salo *et al.*, 1986). Although many instances have now come to light which emphasise the role of riverine barriers in primate diversification, evidence is beginning to accumulate that rivers do not always play such a role, as some primate groups seem to be more affected than others (Ayres & Clutton-Brock, 1992; Clifford *et al.*, 2004; Collins & Dubach, 2000). In South America, Amazonian rivers are a boundary for capuchin and titi monkeys (Ayres & Clutton-Brock, 1992), tamarins (Vallinoto, *et al.*, 2006) and night monkeys (Couette, 2007). In Africa, rivers have played a role shaping the genetic diversity of mandrills (Telfer *et al.*, 2003), bonobos (Eriksson *et al.*, 2004), chimpanzees (Gonder *et al.*, 1997, 2006), and gorillas (Anthony *et al.*, 2007). An Asian example comes from the orang-utans, where rivers are a barrier and have influenced subspeciation of this ape (Jalil *et al.*, 2008; Warren *et al.*, 2001). Contrastingly, Clifford *et al.* (2004) did not find either the Sanaga or the Cross Rivers to delineate boundaries of western lowland gorillas. A similar result was obtained by Craul *et al.* (2008) where Madagascan rivers did not influence gene flow entirely for lepilemurs. Thus, the 'riverine-barrier hypothesis' (Salo *et al.*, 1986) cannot be generalized for primates, although they are likely to be an obstacle for gene flow in some species.

Long-tailed macaques (*Macaca fascicularis*) and proboscis monkeys (*Nasalis larvatus*) are two of the 10 primate species that occur together in sympatry in the LKWS. These species differ in their social structure, home ranges and conservation status. Macaques are organized in multimale – multifemale groups, under a hierarchical matriline and a dominant male (de Ruiter *et al.*, 1994; de Ruiter & Geffen, 1998, Engelhardt *et al.*, 2006). Females are philopatric while males migrate joining other groups, dispersing males have also been observed to swim across rivers in Sumatra (de Ruiter & Geffen, 1998), and the Kinabatangan (personal obs and Jalil, 2007). New groups may be formed by fission of a large group (implying female dispersal) (de Ruiter & Geffen,

1998). Home ranges depend on the quality of the forest, 25-50 ha in primary forest and up to 200 ha in secondary or degraded forest (de Ruiter & Geffen, 1998; Wolfheim, 1983).

The long-tailed macaque has been listed as one of the 100 most invasive alien species inhabiting a variety of forest habitats, preferring edge habitats and riverine areas, but can also be found in villages (i.e. disturbed habitat), often raiding crops and where they may be classified as a pest (van Schaik *et al.*, 1996; Abegg & Thierry, 2002). In contrast, the proboscis monkey is endemic to Borneo. They occur in two types of social group: harems consisting of a single dominant male and several females and all-male groups (Murai, 2004). However, larger bands of individuals comprising both types of social groups often co-habit with overlapping home ranges (Kawabe & Mano, 1972; Bennett & Sebastian, 1988). Female transfer between harems is frequent and is more common in subadult than adult females (Murai *et al.*, 2007). They are excellent swimmers and have been found swimming from riverside to riverside; if frightened while swimming they can dive for several minutes (Bennett & Gombek, 1993; Fleagle, 1998). Home range sizes vary depending on the availability of food, with estimates that fluctuate between 137 ha in Kalimantan to 900 ha in Sarawak (Bennet & Davies, 1994). Currently, the proboscis monkey is threatened by habitat destruction and hunting, and much of its former range has been reduced by logging (e.g. in the Kinabatangan), swamp reclamation, gold mining, shrimp farming and forest fires (Meijaard & Nijman, 2000; Sha *et al.*, 2008). Hunting is much in evidence in Sarawak and Kalimantan (Meijaard & Nijman, 2000). Its conservation status has changed over the past ten years from vulnerable to endangered according to IUCN (2010). The proboscis monkey is currently listed in Appendix I of CITES (UNEP-WCMC, 2010) and is protected by law throughout its range (Meijaard & Nijman, 2000; Sha *et al.*, 2008).

Comparative phylogeography analyses geographic patterns of genetic variation across species in order to find general patterns of evolutionary history within biogeographic regions. It also aims to reveal the evolutionary processes behind these patterns (Avise 2000, 2008; Hickerson *et al.*, 2010). The applications of comparative phylogeography provide a powerful tool for understanding evolutionary history and strengthening conservation efforts, but most studies to date have been based on mtDNA or chloroplast DNA (cpDNA) (Hickerson *et al.*, 2010). For example, Jalil (2007) studied the effect of the Kinabantangan River on the genetic structure of the long-tailed macaque and the proboscis monkey using mitochondrial control region DNA (mtDNA) data. The study revealed that Kinabatangan long-tailed macaques have a stable population and high level of genetic diversity. Evidence of some genetic differentiation was detected between long tail macaque populations on each side of the Kinabatangan River, which the author attributed to be the product of their social systems. Proboscis monkeys exhibit high haplotype diversity without any apparent geographical partitioning, probably because this species

is not as philopatric as the long-tailed macaque and also because they are excellent swimmers (Bennett & Gombek, 1993). Jalil (2007) acknowledged that although informative, the exclusive use of mtDNA in his long-tailed macaque and proboscis monkey study was limited. The use of markers such as mtDNA or cpDNA implies the analysis of only one locus which could be linked to selection, could present introgression, or might not be possible to identify its dispersion between populations as a consequence of behavioural or ecological differences between males and females or between species (Avise, 2008; Domínguez-Domínguez & Vázquez-Domínguez, 2009). For this reason, phylogeography has expanded its focus to nuclear markers with which to obtain more precise parameters from multi-locus data (Hickerson *et al.*, 2010).

In the current study, the genetic diversity, gene flow and, consequently, population structure, of the LKWS proboscis monkey and long-tailed macaque populations previously analysed by Jalil (2007) were re-examined using microsatellite data. As the previous study was based on a maternally inherited marker, bi-parentally inherited markers might be expected to reveal a different population structure, especially in the case of long-tailed macaques. In species with female biased dispersal, estimates of population structure of a population are expected to be similar to that obtained from mtDNA analyses. However, in species such as the long-tailed macaque, which exhibits male- biased dispersal, the pattern of population structure pattern is expected to be less pronounced than that obtained by mtDNA analysis. In contrast, the female-transfer behaviour of the proboscis monkey is expected to reveal a pattern of genetic structure similar to that previously reported by Jalil (2007) from mtDNA.

3.2 Methods

3.2.1 Collection and preservation of faecal samples

Sampling was divided in two seasons to cover different areas of the LKWS. The first sampling season was carried out between October 2007 and March 2008, focusing in Lots 1-4 of the sanctuary. The second sampling season covered Lots 5-10 between June and November 2008 (Figure 3.1). As long-tailed macaques and proboscis monkeys use riverine trees for sleeping sites, the Kinabatangan River was used as a transect to perform a census of the primates during their inactive periods (particularly at sunset and early night). A total of ~330 km was covered, equalling 660 km of riverbank including smaller tributaries. Expeditions lasted from five to ten days, with teams of 2-5 people. Faeces were collected at dawn to ensure freshness, and occasionally at dusk. GPS (Garmin eTrex Vista HCx) coordinates of collected samples were noted for subsequent localisation of groups and populations (Appendix One; detailed list of samples analysed in this study are found in Appendices Seventeen (long-tailed macaque) and

Eighteen (proboscis monkey)). Samples were stored in 50 ml Falcon tubes with 70% ethanol following the protocol described by Goossens *et al.* (2003b) to avoid contamination. Within the constraints of time and logistics, we attempted to sample as widely as possible within each population seen in the riparian forest; because individual recognition was not always possible, and being unhabituated primates, some individuals may have been sampled more than once (Table 3.1).

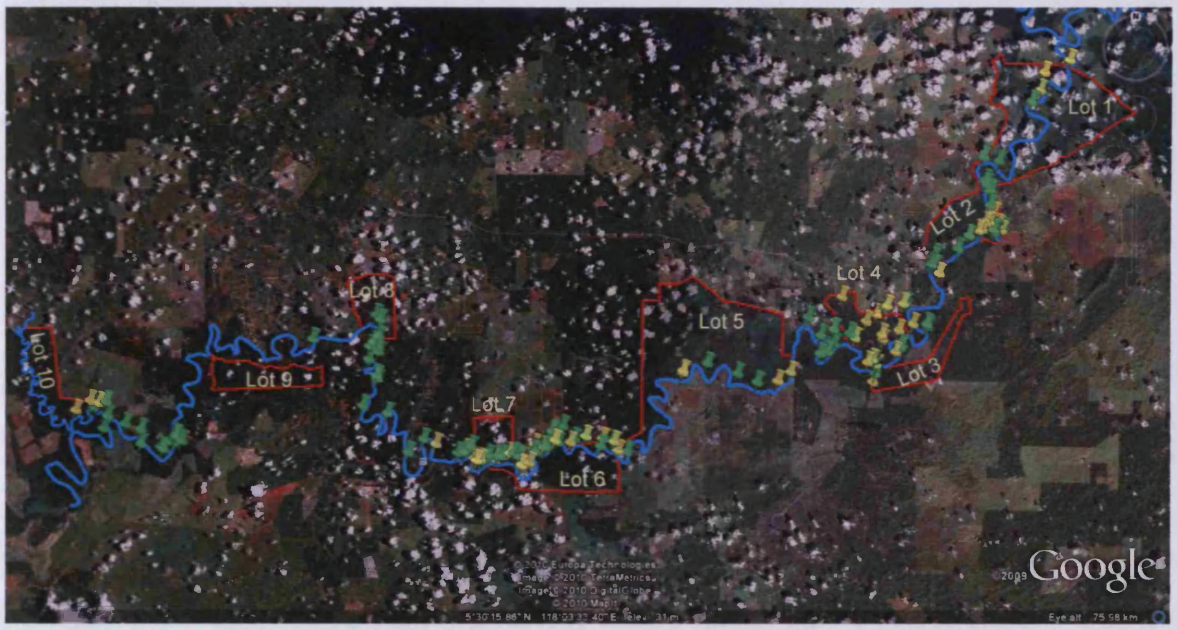


Figure 3.5. Stool sampling of long-tailed macaques and proboscis monkeys in the LKWS. Limits of the Lots conforming the sanctuary are delineated in red, the blue line represents the Kinabatangan River. Green marks correspond to long-tailed macaque samples and yellow marks correspond to proboscis monkey samples.

In addition to the samples collected in the LKWS, eight proboscis monkey stool samples were donated by Dr. Henry Bernard (Unit for Primate Studies Borneo, Institute of Tropical Biology and Conservation, Universiti Malaysia Sabah). These samples were collected along the Garama River in Klias (Northern Sabah), and were stored in 95% ethanol. LKWS samples for preliminary microsatellite screening (58 long-tailed macaque and 80 proboscis monkey DNA faecal extracts) were kindly provided by Dr. Faius Jalil (hereafter referred to as Jalil's samples). These samples varied in volume and DNA quantity, most of which had very low amounts when determined by mtDNA amplification (Jalil personal communication).

3.2.2 Blood and tissue samples

Four samples of DNA from B-lymphocytes of long-tailed macaques ('Pedro', 'Hippo', 'Tresbella' and 'Yabaa') were provided by INPRIMAT (EU FP5 ID: QLRI-CT-2002-01325) and used as

positive controls in PCR reactions. Tissue samples (muscle) from two proboscis monkeys were donated by Lok Kawi Wildlife Park. These originated from a deceased juvenile female (Lily) and a new born (Baby –gender unspecified), Lily's sample was collected a week after the death and stored in 95% ethanol and at 4°C, while Baby's sample was collected immediately after death and stored at -70°C. Skin and hair from both individuals were also donated and were subsequently stored at -70°C. Additionally, samples (muscle, liver and spleen) from two road-killed male proboscis monkeys (Male and Male2) were donated by Sepilok Orangutan Rehabilitation Centre (Sabah, Malaysia). All proboscis monkey tissue samples were obtained after a year and a half of initiating the project and they were stored in 95% ethanol and kept at -70°C.

3.2.3 DNA Extractions

DNA was extracted from each sample in duplicate (unless stated otherwise; see Appendix One) using the QIAamp DNA Stool Mini Kit (QIAGEN GMBH, Germany) following the manufacturer's protocol for "Isolation of DNA from stool for pathogen detection" (Table 3.1, Appendix One). This protocol was previously used by Goossens *et al.* (2000) with a modification in the last step, where DNA was eluted in 100 μl instead of 200 μl, and was subsequently concentrated to ca.70 μl (Jalil, 2007). Sample concentration was performed in a lyophilizer (Concentrator Eppendorf 5301) using the manufacturer's protocol "Aqueous liquids mode". Tissue samples were extracted with DNeasy Blood & Tissue Kit (QIAGEN GMBH, Germany) following the recommendations of the manufacturer, with minor modifications during elution (namely, 5 min incubation at 70°C with buffer AE, which was also preheated at the same temperature). These types of samples were used as positive controls in PCR reactions. To evaluate possible contamination, DNA was extracted from hair of the author (referred to hereafter as Human DNA) according to a protocol described by Jalil (2007).

Table 3.15. Number of collected and extracted stool samples of long-tailed macaques and proboscis monkeys of the LKWS. Samples belonged to populations on the north and south of the Kinabatangan River, conformed by different groups and Lots of the sanctuary.

		LTM			PM		
Riverbank	Lot	Groups	Collected	Extracted	Groups	Collected	Extracted
South	1	7	27	27	5	65	31
	3	17	80	43	11	199	76
	6	15	83	38	7	31	28
	9	11	46	27	-	-	-
	Outside	1	6	3	-	-	-
North	2	9	59	31	8	87	42
	4	11	33	31	17	89	27
	5	10	54	28	6	72	31
	7	8	39	27	14	85	39
	8	6	32	26	-	-	-
	10(A-C)	10	33	28	4	33	29
	Total	105	492	309	72	661	303

LTM – long-tailed macaque, PM- proboscis monkey

To verify the existence of primate DNA from faecal extracts, a partial mitochondrial control region fragment was amplified using species-specific primers (Table 3.2) (Jalil, 2007) and visualized in agarose gels (1.5%). All PCR reactions consisted of 1X Master Mix (QIAGEN, Multiplex Kit), 1X Q-Solution (QIAGEN, Multiplex Kit), 0.5 μM of each primer, and 2 μl of DNA, to a total volume of 10 μl. The cycling parameters comprised an initial denaturation step of 15 min, 95°C, followed by 45 cycles of 30 sec at 94°C, and 90 sec at 55°C (annealing) and 72°C (extension) in each temperature, and a final extension step of 10 min at 72°C.

Table 3.2. Species-specific mtDNA Control Region primers described by Jalil (2007).

Species	***Primer code***	***Primer Sequence***
Long-tailed macaque	*Mf-5'*	5'-GCA ACT ACT TTC TGC ACT-3'
	Mf-3'	5'-GAA CAA GGG ATT CCT AAG-3'
Proboscis monkey	*Nl-5'*	5'-CGT AAA CCA GAA ACG GAT-3'
	Nl-3'	5'-TAA TGG GAA TAT CCG TGC-3'

DNA from the muscle/hair samples was visualized in agarose gels (1.5%) and quantified by spectrofluorometry. The Quant-iT™ PicoGreen® (Invitrogen) assay was performed as stipulated by the manufacturer, except that all volumes were scaled 10-downfold to make the assay suitable for microtitration. Microtiter plates were read using a SpectraMAX GEMINI XS spectrofluorometer (Molecular Devices) and DNA concentrations were calculated using the SOFTmax® Pro software (Molecular Devices).

3.2.4 Microsatellite screening

3.2.4.1 Cross-amplification of human microsatellite markers in long-tailed macaques

Human primers (available in the laboratory) were used to test a total of 25 microsatellite loci in long-tailed macaques (Table 3.3). Initial PCR methods were based on Bonhomme *et al.* (2005) and Kikuchi *et al.* (2007) to screen nine loci using the QIAGEN Multiplex Kit and different primer concentrations depending on each protocol (Table 3.3). As a template, 2 µl of two Jalil's samples were used, and negative (H_2O only) and contamination/positive controls (Human DNA) were included for all loci. Thermal profile for PCR amplification was used as follows: 95°C for 15 min, followed by 36 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 90 s (54°C for D3S1768 and for D3S1768), extension at 72°C for 90 s, and a final extension at 72°C for 10 min. All reactions were carried out in a GeneAmp PCR System 9700 (Applied Biosystems).

Table 3.3 Human microsatellite primers tested in *Macaca fascicularis*.

Locus	Repeat motif	Reported allele size range[d]	
		Long-tailed macaque (*Macaca fascicularis*)	Human (*Homo sapiens*)
D1S207[a]	(CA)n	126-158	142-170
D1S548[b]	(GATA)n	191-215	148-172
D1S550	(GATA)n	ND	179
D3S1766[c]	(GATA)n	ND	208-232
D3S1768[a,b]	(GATA)n	170-250	186-206
D4S243[a,b]	(AGAT)n	182-282	173
D5S820[b]	(GATA)n	169-225	190-218
D5S1457	(GATA)n	108-131	151-159
D6S265[a]	(CA)n	113-137	83-183
D6S271	(CA)n	180-200	166-208
D6S291	(CA)n	ND	168-219
D6S2883	(CA)n	65,67, 236-268	203-303
D7S503	(CA)n	ND	170
D7S2204	(GATA)n	216-256	217-269
D8S1106	(GATA)n	130-174	131-151
D10S611[a]	(GATA)n	164-202	146-158
D11S925[a]	(CA)n	225-239	172-199

Locus	Repeat motif	Reported allele size range[d]	
		Long-tailed macaque (*Macaca fascicularis*)	Human (*Homo sapiens*)
D12S67	(GATA)n	ND	233-273
D12S391[c]	(GATA)n	ND	211-251
D13S765	(GATA)n	208-240	193
D14S255[c]	(CA)n	ND	197-207
D14S306	(GATA)n	165-200	190-210
D16S420	(CA)n	ND	179-201
DXS571	(CA)n	133-153	129-130
DXS6799	(TATC)n(ATCC)	171-195	241-261
DXS6810[b]	(GATA)n	171-210	208-223
MIB	(CA)n	185-215	226-256

ND-Not determined.
[a] Primers tested under Kikuchi *et al.* (2007) protocol (primer concentration = 0.33 μM).
[b] Primers tested under Bonhomme *et al.* (2005) protocol (primer concentration = 0.15 μM, except D1S5468 and D3S1768 which were used at 0.1 μM).
[c] Primers not tested in the genus *Macaca*.
[d] Bonhomme *et al.*, 2005; Clisson *et al.*, 2000; Hadfield *et al.*, 2001; Kikuchi *et al.*, 2007; Roeder *et al.*, 2009.

Additional optimisation of all primers was carried out using the QIAGEN Multiplex PCR Kit and with the multi-tubes PCR repetition approach (Taberlet *et al.*, 1996). All PCR reactions were performed in 10 μl total volume, containing 500 pg of template (INPRIMAT samples) or 2 μl of DNA (13 faecal extracts, including at least one per Lot of the LKWS), and 0.2 μM of each primer. PCR amplification was as described above but temperature gradients were performed during the annealing step in a Robocycler Gradient 96 thermocycler (Stratagene). Optimised annealing temperatures were verified in a GeneAmp PCR System 9700 (Applied Biosystems). All PCR products were electrophoresed in 3.0% agarose-0.5X TBE gels, and visualized using 1.5 ng/ml EtBr in a GelDoc-IT™ Imaging System (UVP). Loci which had not been previously described were genotyped to obtain approximate size ranges. Fluorescently labelled (5'- FAM, HEX, or NED), undescribed markers were electrophoresed along with GeneScan ROX 350, an internal size standard, in a Prism 3700 Genetic Analyzer (Applied Biosystems) and fragment lengths were scored using GeneMapper® ID version 3.2 (Applied Biosystems).

Seventeen microsatellite loci, consisting of both di- and tetra-nucleotide repeats, which produced a single target band and consistently amplified DNA from faeces, were assembled in five multiplexes (named M1 to M5: Table 3.4). Multiplexes M1 to M4 were assembled taking into

account the primers' annealing temperature, the allele sizes (from the original and this study), and the non-overlap of fragment sizes for loci labelled with the same fluorescent dyes. M5 was assembled by combining PCR products from singleplex amplification. Genotyping was performed as described above. After the screening process, two tetra-nucleotide loci (D10S611 and D5S1457) were removed from the study because of difficulties in amplification.

Table 3.4. Human microsatellite loci screened via multiplex in LKWS long-tailed macaques

Locus	Label	Range (bp)	Repeat motif	T_a
M1				
D8S1106	FAM	124-174	GATA	54°C
D7S2204	HEX	213-265	GATA	54°C
D12S67[a]	NED	138-216	GATA	54°C
DXS571	HEX	117-167	GATA	54°C
			CA	
M2				
D7S503[a]	NED	129-185	CA	56°C
D10S611	HEX	164-202	GATA	56°C
D5S1457[a]	FAM	121-140	GATA	56°C
D11S925	FAM	225-239	CA	56°C
D6S2883	HEX	107-139	CA	56°C
			CA	
M3				
D1S548	NED	145-209	GATA	60°C
D3S1768	HEX	185-233	GATA	60°C
D16S420[a]	FAM	198-244	CA	60°C
			CA	
M4			CA	
D1S207	HEX	125-158	CA	63°C
D3S1766[a]	FAM	209-265	GATA	63°C
D5S820	NED	176-200	GATA	63°C
M5				
D1S550[a]	FAM	138-182	GATA	51°C
D6S291[a]	HEX	199-219	CA	63°C

[a]–allele size ranges for *Macaca fascicularis* determined in this study

3.2.4.2 Cross-amplification of human microsatellite markers in proboscis monkeys

Human primers were used to screen 39 loci in 46 proboscis monkeys DNA faecal extracts (Table 3.5). Primer selection was based on the amplification success (polymorphic loci) in langurs (*Semnopithecus entellus*, *Trachypithecus francoisi*, and *Presbytis melalophos nobilis*) (Bonhomme *et al.*, 2005; Coote & Bruford, 1996; Launhardt *et al.*, 1998; Little, 2003; Roeder *et al.*, 2009), and primer availability in the laboratory. PCR optimisation was carried out for each locus separately in 10 μl reactions, testing different conditions: 1X Master Mix (QIAGEN Multiplex PCR Kit), ± 0.5X Q-Solution (included in Multiplex Kit), forward and reverse primers (0.2, 0.4 or 0.6 μM), and 1 or 2 μl of DNA extract. Initial amplifications were carried out in a Robocycler Gradient 96 thermocycler (Stratagene), testing two profiles: 1) initial denaturation at 95°C for 15

min, 40 cycles of 94°C for 30 sec, 50-61°C for 90 sec, 72°C for 90 sec, and a final extension of 72°C for 10 min; 2) 95°C for 15, 40 cycles of 94°C for 30 sec, 48-59°C for 120 sec, and a final extension of 72°C for 30 min. Further optimisations were carried out in a GeneAmp PCR System 9700 (Applied Biosystems), under the following profile: 95°C for 15, 40 cycles of 94°C for 30 sec, optimised annealing temperature for 105 sec, and a final extension of 60°C for 30 min. Negative (H_2O only) and contamination/positive controls (Human DNA) were included for all loci, and all PCR products were electrophoresed in 3 % agarose-0.5X TBE gels, and visualized using 1.5 ng/ml Ethidium Bromide in a GelDoc-IT™ Imaging System (UVP).

For initial genotyping, positive PCR products were run individually on an Applied Biosystems 3130xl Genetic Analyzer using a 50 cm capillary array with POP-7 polymer; the electropherograms were analyzed with PeakScanner V1.0 (Applied Biosystems).

Table 3.5. Human microsatellite primers tested in *Nasalis larvatus*

		Reported allele size range	
Locus	Repeat motif	Humans	Langurs
D1S207	CA	133-170	131-153
D1S533	GATA	193-225	ND
D1S548	CTAT	148-172	191-211
D1S550	GATA	179	ND
D2S1326	GATA	233-268	ND
D2S434	GATA	173	138-146
D3S1766	GATA	208-232	ND
D3S1768	GATA	186-206	162-166
D4S2365	GATA	284-304	282-290
D4S2366	GATA	120-140	98-106
D4S243	GATA	173	138-146
D5S1457	GATA	151-159	ND
D5S820	GATA	169-225	197-201
D6S271	CA	166-208	172-178
D6S287	CA	143-171	136-140
D6S2883	CA	203-303	183-193
D6S291	CA	168-219	149-151
D6S2972	CA	122-150	122-130
D6S311	CA	229-276	234-238
D7S2204	GATA	217-269	223-239
D7S503	CA	149-171	151-167
D8S1106	GATA	131-151	114-138
D10S611	GATA	146-158	ND

		Reported allele size range	
Locus	Repeat motif	Humans	Langurs
D11S871	CA	186	ND
D11S925	CA	182-212	182-212
D12S375	GATA	175-176	ND
D12S391	GATA	211-251	ND
D12S67	GATA	233-273	134-146
D13S159	CA	168-203	136-140
D13S317	GATA	175-199	ND
D13S765	GATA	193	198-202
D14S306	GATA	190-210	ND
D16S420	CA	195-201	164-184, 195-201
D17S791	CA	165-199	146-168, 118-130
DXS571	CA	129-130	119-125
DXS6799	(TATC)n(ATCC)	241-261	187-191
DXS6810	TCTA	208-223	226-295
MIB	CA	326-256	177-192
SCA-1	CAG	199-232	145-157

ND- Not Described

3.2.4.3 Microsatellite screening in proboscis monkeys

Full optimisation of any human-derived microsatellite marker could not be achieved for the proboscis monkeys, hence the need to develop species-specific primers, which are described in Chapter Two (Salgado-Lynn *et al.*, 2010). Eight polymorphic microsatellite loci, consisting of both di- and tri-nucleotide repeats, were amplified in singleplex and subsequently combined in multiplexes for final genotyping. PCR mixes consisted of 1X Master Mix (QIAGEN, Multiplex PCR Kit), 0.4μg/μl BSA (New England BioLabs), 0.2 μM of each primer, 200 pg (positive controls and Human DNA) or 2 μl (faecal extracts) of template DNA to a total volume of 10 μl and were performed in a GeneAmp® PCR System 9700 (Applied Biosystems). The amplification conditions were as follows: 95°C for 15 min, 45 cycles at 94°C for 30 s, T_a for 90 s, 72°C for 90 s and a final extension at 72°C for 30 min. The PCR products of the three multiplexes (M1, M2 and M3) (Table 3.6) were electrophoresed along internal size standards GeneScan ROX 350 or GS-400 HD LIZ, in a Prism 3700 Genetic Analyzer (Applied Biosystems) and fragment lengths were scored using GeneMapper® ID 3.2 (Applied Biosystems).

Table 3.6. Microsatellite loci screened via multiplex in LKWS proboscis monkeys

Locus	Label	Range (bp)	Repeat motif	T_a
M1				
NlP1A6	NED	146-158	$(AAC)_7$	54°C
NlP4C11	HEX	241-259	$(GT)_{23}$	54°C
NlP3B2	FAM	158-172	$(GT)_{23}$	58°C
M2				
NlE10	NED	153-205	$(GT)_{16}$	58°C
NlD10	HEX	177-189	$(CT)_{10}$	54°C
NlP1C5	FAM	177-205	$(GT)_{15}$	54°C
M3				
NlP2D6	HEX	146-160	$(GT)_{17}$	54°C
NlP2F3	FAM	175-187	$(AG)_{17}$	58°C

3.2.4.4 Genotyping criteria

Specific conditions were followed for genotyping. Given the initial number of extracted DNA samples (roughly 300 samples for each primate species), time and financial considerations prevented us from typing each individual as many times as recommended by Taberlet *et al.*, (1996) when genotyping nuclear loci of very low DNA samples (up to 7 positive PCR repeats to confirm homozygosity). In the case of long-tailed macaque samples, the simulation program GEMINI v.1.4.1 (Valière *et al.*, 2002) was used to determine the number of positive PCR repetitions needed to obtain a reliable genotype. To run GEMINI, allele frequencies for a subset of the macaque samples (40) were calculated using GENEPOP 4.0.10 (Raymond & Rousset, 1995; Rousset, 2008) and error rates were calculated with PEDANT v. 1.0 (Johnson & Haydon, 2007, 2009). PEDANT estimates allelic dropout (ADO) and false allele (FA) error rates in the absence of reference data by comparing duplicate genotypes based on the frequency and nature of mismatches. With the allele frequencies and the error rates, a consensus threshold test was performed with GEMINI. General parameters included 1000 simulation replicates, the heterozygosity rate was selected as not constrained within the population parameters, the sampling parameters included the use of the non-invasive model and a probability of capture uniform over individuals, genotyping parameters allowed double errors and not uniform error rates over loci, also the range of repetition number was set from 2 to 12. After calculating the consensus threshold, a PCR repetition test was performed using the same parameters but including the calculated consensus to the repetition parameters. This test determined that four PCR repetitions were sufficient per locus to achieve a high probability of identity (> 94%), and consensus genotypes were determined using the per locus consensus threshold calculated also by GEMINI, meaning that a genotype had to be observed a particular number of times out of four to

be confirmed (Table 3.7). Finally, for a sample to be included in the final analysis it had to have at least three positive PCR results for eight out of the 15 loci (Table 3.8).

Table 3.7. Genotyping consensus threshold for loci screened in *Macaca fascicularis* of the LKWS.

Locus	Consensus threshold	Locus	Consensus threshold
D11S925	2	D5S820	3
D12S67	3	D6S2883	2
D16S420	2	D6S291	3
D1S207	2	D7S2204	3
D1S548	3	D7S503	2
D1S550	3	D8S611	2
D3S1766	2	DXS571	2
D3S1768	2		

In the case of proboscis monkey samples, since multiplexing was not possible, the volume of the samples (ca. 70 μl/extract) was an additional constraint for a full multi-tube approach (Taberlet *et al.*, 1996). An individual was typed as heterozygous if both alleles appeared at least twice within four replicates, and a homozygous was typed if it appeared as such at least three times. As with the macaques, for a proboscis monkey sample to be included in subsequent analyses at least 3 positive PCR results were required in five out of the eight loci screened (Table 3.8).

Table 3.8. Total samples tested and included in the genetic analyses. Two populations at each side of the river (South and North) divided by lot of the LKWS for each primate species, and amplification success.

South	Lot 1		Lot 3		Lot 6		Lot 9		Outside				Total		
	T	S	T	S	T	S	T	S	T	S	T	S	T	S	% A
LTM	16	9	15	11	15	10	15	12	1	1			62	43	69.35
PM	15	6	27	8	20	8	-	-	-	-			62	22	35.48
North	Lot 2		Lot 4		Lot 5		Lot 7		Lot 8		Lot 10				
LTM	16	7	15	13	15	12	15	14	15	11	15	9	91	66	72.53
PM	15	1	15	13	15	12	15	10	-	-	14	9	74	45	59.46

LTM – long-tailed macaque, PM- proboscis monkey, T- total samples tested, S-total number of samples screened in this study, %A- PCR amplification success.

3.2.5. Estimation of error rates and null alleles

Genotyping errors, such as null alleles, allelic dropout, and scoring of stutter peaks, were assessed statistically using MICROCHECKER, version 2.2.3 (van Oosterhout *et al.*, 2004). Additionally, the probabilities of genotyping error for homozygotes and heterozygotes, as well as false allele and allelic dropout rates were calculated by the method described by Zhan *et al.* (2009). Null allele frequencies were additionally calculated using the software FREENA, which uses the algorithm by Dempster, Laird & Rubin (1977) and has previously been found to perform better than other null allele frequency estimators (Chapuis & Estoup, 2007). However, this software calculates null allele frequencies for all the locus/population combinations and does not take into account that heterozygote deficits may be due to other possible causes.

3.2.6. Genetic diversity analysis

GENETIX 4.05 (Belkhir *et al.*, 1996-2004) was used to perform all standard population genetic analyses: mean number of alleles per locus and per population, allele frequencies at each locus, and expected (H_E) and observed heterozygosity (H_O) per population. H_E and H_O were estimated per locus using ARLEQUIN version 3.5.1.2 (Excoffier & Lischer, 2010). Genotypic linkage disequilibrium (LD) was also estimated using ARLEQUIN using 10000 permutations. Significance levels were adjusted using Bonferroni corrections for multiple testing ($P < 0.01$). Heterozygote deficiency was tested, as compared to Hardy-Weinberg equilibrium for each locus. Deviation from Hardy-Weinberg equilibrium were tested calculating Weir & Cockerham's estimate of F_{IS} (Weir & Cockerham, 1984) for each locus and also globally, using GENETIX with 10000 permutations. Significant positive values of F_{IS} indicate heterozygote deficiency, and significant negative values indicate heterozygote excess.

3.2.7. Population structure analysis

Genetic differentiation and gene flow among populations were estimated using the F_{ST} analogue (theta) of Weir and Cockerham (1984) implemented by GENETIX. The same program was used to explore patterns of genetic differentiation between individuals sampled on the north and south of the Kinabatangan River using Factorial Correspondence Analysis (FCA) based on allele frequencies (Belkhir *et al.*, 1996-2004). Additionally, as suggested by Pearse & Crandall (2004), two different Bayesian clustering methods were used to investigate the spatial genetic structure of the long-tailed macaque and the proboscis monkey in the LKWS. First, the software STRUCTURE (version 2.3.1; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009; Pritchard *et al.*, 2000) was used because the variety of modelling options available make it well suited to the

detection of various patterns of population genetic structure and it incorporates null alleles present in the dataset (Carlsson, 2008; Evanno *et al.* 2005; Falush *et al.*, 2007; Latch *et al.* 2006). The second approach incorporates spatial information through the inclusion of coordinates for each sample into the model and is implemented in GENELAND 3.1.5 (Guillot *et al.*, 2005a, b; Guillot, 2008) an extension of the program R (version 2.10.1; R Development Core Team, 2010).

To infer the number of genetically differentiated clusters (subpopulations, K), STRUCTURE was run for a range of K values between one and five using 10 independent runs of 1'000,000 iterations (plus a burn-in of 100,000 iterations) for each K value. Runs were performed with the "admixture model" and the "correlated allele frequency" model without prior information and with an initial alpha value set to 1.0. Null alleles were considered as recessive to all other alleles, where the non-amplified samples were coded as homozygotes for the recessive allele (null allele). The maximum K value was chosen as the number of sampled subpopulations (north and south riverbanks) plus three, as suggested by Evanno *et al.* (2005). The most probable K value was estimated by using the model choice criterion implemented in STRUCTURE that is the maximal value estimate of posterior probability of the data for a given K,Pr(X|K) (Pritchard *et al.*, 2000). Individuals were assigned to one cluster if their proportion of membership (q_i) to that cluster was equal to or larger than 0.600, considering that if more than half the genome of an individual is assigned to the same genetic group then this individual can be assigned to this group with reasonable confidence; the individuals with maximum inferred ancestry < 0.6 were not assigned to any group (Coulon *et al.*, 2008; Quéméré *et al.*, 2009).

To infer the number of K in GENELAND, as suggested in Guillot *et al.* (2005a), a preliminary series of 10 runs of 100,000 MCMC iterations were performed to infer the most probable number of genetic groups (K), where K was allowed to vary. The number of subpopulations was set from 1 to 12, using a matrix of genotypes, spatial coordinates for each individual, and 1000 stored MCMC iterations (100,000 iterations, thinning 100). Allele frequencies were drawn from independent Dirichlet distributions (Pritchard *et al.* 2000) as this model has been shown to perform better than the alternative model (F-model; Guillot *et al.* 2005a). The amount of uncertainty to spatial coordinates was set at 0.3 and the maximum rate of the Poisson process was set to 100, a value close to the number of individuals in the datasets as suggested by Guillot *et al.* (2005a). This parameter controls the number of polygons in the geographical area under study, and the value used herein corresponds to strongly fragmented partitions and weak dependence on the spatial organization of populations. In the spatially explicit GENELAND model, subpopulations are assumed to be spatially organized through the Poisson–Voronoi tessellation (Dupanloup *et al.* 2002); the maximum number of nuclei within this tessellation was

set to 300 (3 times the maximum rate as suggested by Guillot *et al.* 2005a). Next, we inferred the number of subpopulations from the modal K of these 10 runs, and ran MCMC 5 more times with K fixed to this number and other parameters unchanged. For each of the 5 runs with fixed K, the posterior probability of subpopulation membership was computed for each pixel of the spatial domain (100 x 100 pixels), using a burn-in of 100 iterations. Finally, the most probable number of clusters (K) was found comparing the histograms across the 5 runs.

3.3 Results

3.3.1. Genetic diversity in long-tailed macaques of the LKWS

A total of 109 long-tailed macaque individuals from the north (n=66) and south of the Kinabatangan River (n= 43) were genotyped using 15 microsatellite loci (Appendix Two). Genetic diversity was measured as the mean number of alleles per locus (MNA), and observed (H_O) and expected (H_E) heterozygosities. In total, 202 different alleles were observed with a MNA of 13.5 (SD=3.2), ranging from eight (D5S820) to 21 (D7S507) (Figure 3.2). The level of polymorphism per population was also high, with a mean number of alleles of 11.2, being 11.6 for the north bank and 10.8 for the south bank (Table 3.9). The frequencies of the alleles generally showed multimodal distributions with more than two common alleles and a range of other alleles at low frequencies, except for the locus D12S67, which had a single common allele with high mean frequency of 78 % over both populations. Allele frequency distribution by locus and population is shown in Appendix Three.

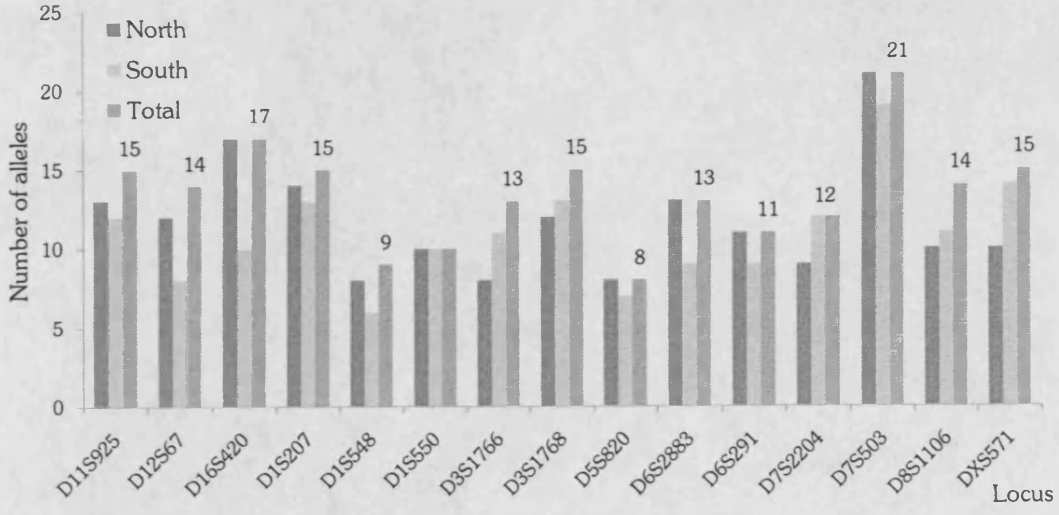


Figure 3.2. Total number of alleles per locus across both populations (to the north and south of the Kinabatangan River) of long-tailed macaques.

The mean H_E and H_O per population were high: at 0.8 (H_E) for both subpopulations and H_O values of 0.59 and 0.53 in the north and south populations respectively (Table 3.9). Across all loci, H_E varied from 0.41 (D12S67) to 0.9 (D7S503), and H_O varied from 0.23 (D12S67) to 0.86 (D6S2883). High F_{IS} values were also observed for most loci, ranging from 0.001 (D6S2883, north) to 0.592 (D3S1766, south). Most of these loci exhibited highly significant deviation from Hardy-Weinberg equilibrium (HWE) proportions for both populations (Table 3.9).

Table 3.9. Average number of alleles across populations (N_a), observed (H_O) and expected (H_E) heterozygosities and departures from Hardy-Weinberg proportions (F_{IS}) for sampled populations and all loci of long-tailed macaques.

		Population					Population		
Locus		North	South	N_a	Locus		North	South	N_a
D11S925	H_E	0.85	0.84		D5S820	H_E	0.80	0.80	
	H_O	0.51	0.38	13		H_O	0.53	0.39	11
	F_{IS}	0.409	0.553			F_{IS}	0.346	0.528	
		***	***				***	***	
D12S67	H_E	0.41	0.34		D6S2883	H_E	0.86	0.85	
	H_O	0.23	0.19	10		H_O	0.86	0.63	11
	F_{IS}	0.454	0.446			F_{IS}	0.001	0.266	
		***	***				NS	***	
D16S420	H_E	0.83	0.72		D6S291	H_E	0.84	0.83	
	H_O	0.52	0.46	14		H_O	0.67	0.77	10
	F_{IS}	0.381	0.378			F_{IS}	0.203	0.087	
		***	***				***	NS	
D1S207	H_E	0.86	0.86		D7S2204	H_E	0.82	0.83	
	H_O	0.77	0.70	14		H_O	0.54	0.61	11
	F_{IS}	0.116	0.203			F_{IS}	0.353	0.287	
		**	***				***	***	
D1S548	H_E	0.63	0.69		D7S503	H_E	0.90	0.86	
	H_O	0.40	0.44	7		H_O	0.75	0.56	20
	F_{IS}	0.369	0.380			F_{IS}	0.175	0.357	
		***	**				***	***	
D1S550	H_E	0.74	0.74		D8S1106	H_E	0.86	0.86	
	H_O	0.44	0.57	10		H_O	0.74	0.64	11
	F_{IS}	0.41	0.244			F_{IS}	0.152	0.266	
		***	**				**	**	
D3S1766	H_E	0.83	0.86		DXS571	H_E	0.85	0.83	
	H_O	0.52	0.36	10		H_O	0.65	0.62	12
	F_{IS}	0.376	0.592			F_{IS}	0.25	0.268	
		***	***				***	***	
D3S1768	H_E	0.87	0.86						
	H_O	0.69	0.63	13					
	F_{IS}	0.369	0.38						
		***	***						

Table 3.9. Continued

	Total	North	South
	H_E	0.80	0.80
	(SD)	0.13	0.13
	H_O	0.59	0.53
	(SD)	0.17	0.15
	F_{IS}	0.269	0.338
		***	***
	MNA	12	11

MNA – mean number of alelles per locus.
NS- non significant, ** - P <0.01, *** - P <0.001.

Significant linkage disequilibrium ($P<$ 0.05) was found between some loci and in both populations (Appendix Four). Pairs of loci comprising D7S2204, D7S503, D16S420, D3S1766 and D5S820 exhibited the most significant LD values, ranging from 5.5 to 6.5 across populations (SD ranges 0 - 2.12).

DNA degradation, low DNA concentrations and primer-site mutations may also result in the incorrect assignment of microsatellite genotypes, biasing population genetic analyses. Assessment of genotyping errors using MICROCHECKER (van Oosterhout *et al.*, 2004) gave no evidence for large allelic dropout but notified scoring errors due to stuttering in two loci (DXS571 and D11S925) of the north population and one locus (D1S550) of the south population. The presence of null alleles due to a general excess of homozygotes for most allele size classes was also notified. Both populations were probably in Hardy Weinberg equilibrium but 14 loci of each population showed evidence of null alleles (excepting D6S28883 on the north and D6S291 on the south). Genotyping errors calculated by the method of Zhan *et al.* (2009) gave an allelic dropout (ADO) error rate of 0.25, and a false allele (FA) error rate of 0.1, however the mean probability of genotyping error was 0.05. Null allele frequencies calculated with Dempster *et al.* (1977) algorithm using FREENA were in the range 0.115–0.811, this program calculated null allele frequencies for all locus/populations that presented heterozygote deficits and no locus/population combinations had null frequencies equal to zero. This excess of homozygotes very likely explains the high F_{IS} values, departures from HWE, and some of the LD between loci. Although the calculated null allele frequencies (see Appendix five) were generally non negligible, analysis of population structure was nonetheless carried out with the 15 loci, with adjustments made to account for the presence of null alleles (Carlsson, 2008; Falush *et al.*, 2007).

3.3.2. Population structure in long-tailed macaques of the LKWS

Despite the high F_{IS} values, little structure was observed from the Factorial Correspondance Analysis (FCA) plot of individual microsatellite genotypes (Figure 3.3). Using GENETIX, F_{ST} was

equal to 0.005 ($P > 0.01$) with a number of migrants per generation (Nm) of 43, indicating high levels of gene flow. However, the Bayesian clustering performed using STRUCTURE considering the null alleles as recessive, estimated a log likelihood of the data, ln Pr(X | K), maximal at K = 2 (Table 3.10). When K increased (K = 3-5), the results showed a similar pattern as with K = 2 (Figure 3.4).

Table 3.10. Inference for the number of populations (K) of long-tailed macaques with STRUCTURE. The posterior probability of the number of populations was maximum with K = 2

K	1	2	3	4	5
ln Pr(X | K)	-6340.3	***-6120.0***	-6141.9	-6411.8	-6202.6

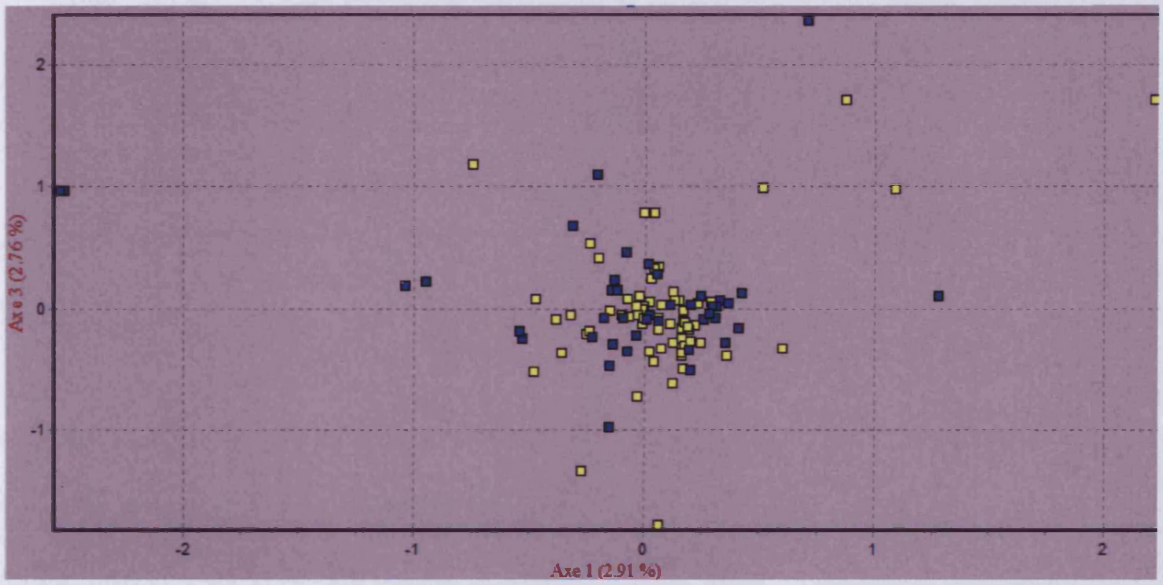


Figure 3.3 Factorial Correspondance Analysis (FCA) showing the relationship among multilocus genotypes of individual long-tailed macaques from the north (yellow) and south (blue) of the Kinabatangan River in the LKWS.

One cluster (II) grouped together the population sampled on the north bank of the Kinabatangan River (q_I= 0.646). In contrast, the population to the south of the river comprised individuals from both clusters and, as a whole, this population could not be assigned to any cluster (q_I= 0.434 and q_{II}= 0.566), indicating admixture. The proportions of membership of each sampled population in the two inferred clusters are shown in Table 3.11. The unassigned individuals were assigned to both clusters (I and II) with probability lower than 0.600, indicating that they are admixed (Randi *et al.*, 2003). Potential migrants observed within the population sampled to the north of the river are the number of individuals assigned with a probability larger than 0.600 in the opposite cluster (62% of the individuals of that population were assigned to cluster II, while 19 individuals averaging q_I= 0.751 were assigned to cluster I). These potential migrants belonged mainly to Lot 5 (42%) and Lot 7 (37%). However, when each Lot was analysed in STRUCTURE

as an independent population, two clusters were still recognized by the program but no population met the condition of $q_i > 0.600$ to be assigned to either cluster (Appendix Six).

Table 3.11. Bayesian clustering analysis in long-tailed macaques performed with STRUCTURE. The table shows the proportion of membership (q) of each predefined sampled population in each of 2 inferred clusters. The number (in parentheses) and percentage of total individuals assigned (including migrants) are indicated. Proportions of membership of potential migrants are indicated with their original cluster (I) (see text for details). n = population size.

Population (n)	Cluster		Potential Migrants	Unassigned individuals	% of total assigned
	qI	qII			
North(66)	0.354(0)	0.646(41)	0.751(19)I	6	90.91
South(43)	0.434(12)	0.566(20)	-	11	74.42

In the 10 runs performed with GENELAND to estimate K, the posterior density and the log-likelihood levels reached a plateau well before the end of the MCMC runs, indicating that they had reached convergence. In seven runs the modal number of genetic groups (K) was one, whereas the other three runs showed two clusters. Therefore, subsequent runs were performed with K fixed to 1. The 5 runs selected on the basis of their posterior also reached convergence. In all cases, the 109 individuals were assigned to one spatially defined population (Figure 3.5).

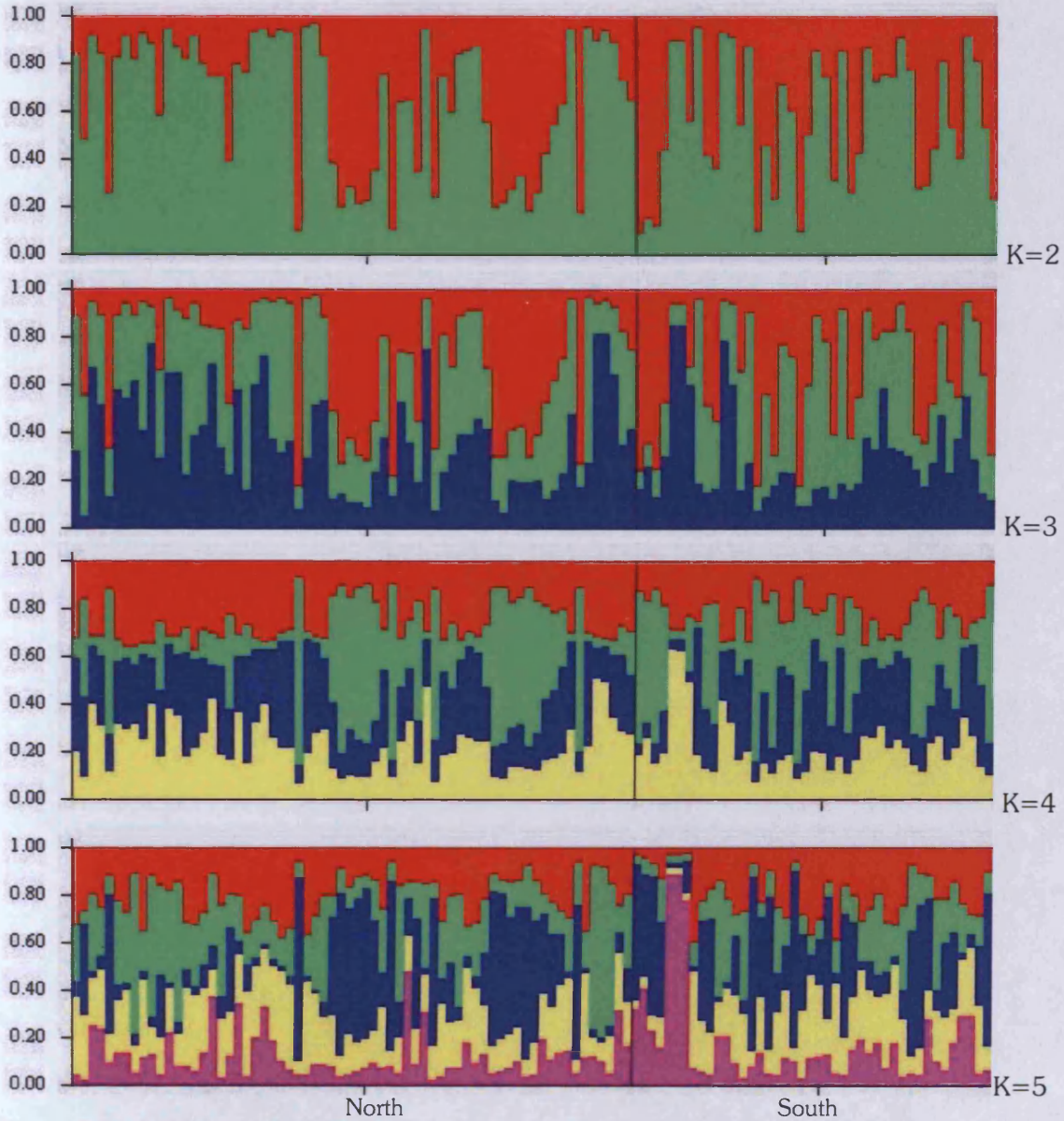


Figure 3.4. Long-tailed macaque clustering results (K = 2) for both sampled sites, according to STRUCTURE analysis. Each individual is represented as a vertical line partitioned into coloured segments. Sampled sites are separated by black vertical line and labelled below the figure.

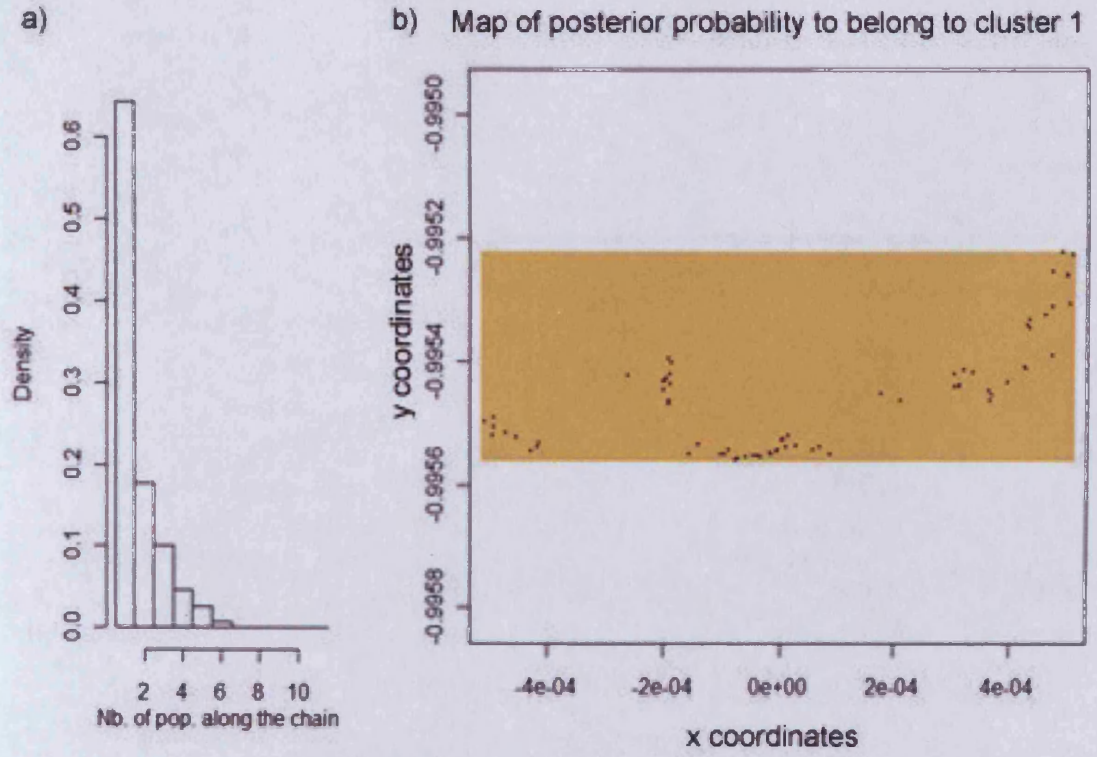


Figure 3.5. Spatial genetic clustering analysis of long-tailed macaques with GENELAND. a) Histogram showing the posterior density distribution of the number of clusters estimated from GENELAND analysis. b) The contour map shows the posterior probability for all long-tailed macaque individuals to belong to one cluster. Contour maps represent polygons that approximate the true pattern of population spread across space, each of them belongs to one the K population determined by GENELAND and is represented with a different colour (the yellow area in this case).

3.3.3. Genetic diversity in proboscis monkeys of the LKWS

A total of 67 proboscis monkey individuals from the north (n= 45) and south of the Kinabatangan River (n= 22) were genotyped using eight microsatellite loci (Appendix Seven). Seventy different alleles were observed in the whole sample with a mean number of alleles per locus of 8.8 (SD= 2.4), ranging from seven (D10, P1A6, P2F3) to 14 (E10) (Figure 3.6). The level of polymorphism per population was high, with a mean number of alleles of 6.5, being 6.2 for the north bank and 6.7 for the south bank (Table 3.12). The frequencies of the alleles generally showed bimodal distributions with two common alleles and a range of other alleles at lower frequencies. Allele frequency distribution by locus and population is shown in Appendix Eight.

The mean H_E and H_O per population were relatively high, with 0.68 (H_E) for both subpopulations and observed values of 0.65 and 0.70 in the north and south populations respectively (Table 3.12). Across all loci, H_E was high and varied from 0.55 (NlD10) to 0.82 (NlE10), and H_O varied from 0.26 (NlE10) to 0.82 (NlD10). High F_{IS} values were also observed for most of the loci, ranging from -0.323 (NlD10, north) to 0.693 (NlE10, south), although this was non-significant for

the northern population (Table 3.12). All but NIP2D6 exhibited significant deviation from Hardy-Weinberg equilibrium (HWE) proportions for both populations (Table 3.12).

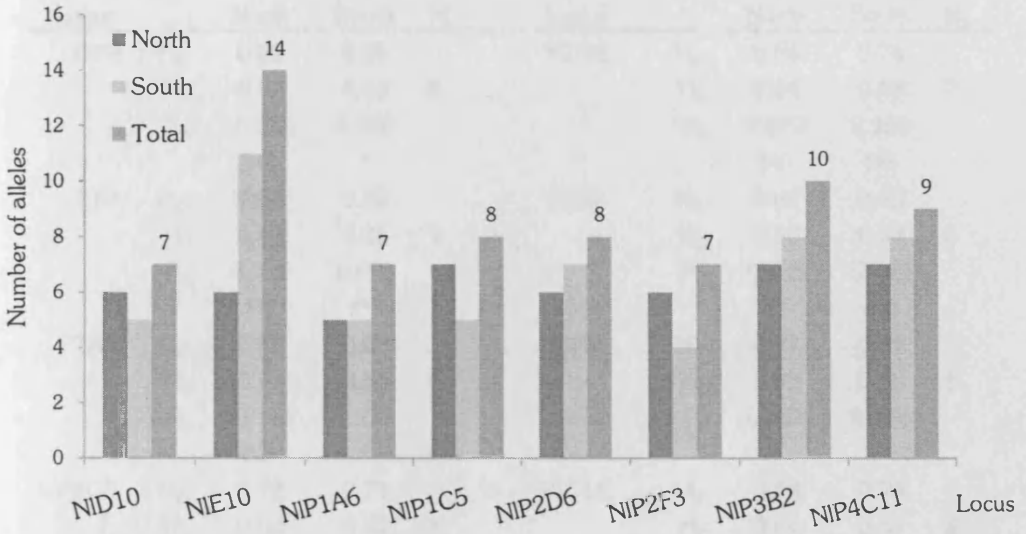


Figure 3.6. Total number of alleles per locus across both populations of proboscis monkeys

Significant linkage disequilibrium ($P< 0.05$) was found between some loci and in both populations (Appendix Nine). Within the northern population, loci NIP1C5 and NIP2D6 were linked ($P=0.001$), while locus NIE10 was the most common among linked pairs (four) in the southern population.

Assessment of genotyping errors using MICROCHECKER (van Oosterhout *et al.*, 2004) gave no evidence for large allelic dropout. Scoring errors due to stuttering and the presence of null alleles due to a general excess of homozygotes was notified for the locus NIP2F3 of the north population. The presence of null alleles was also notified for the locus NIP4C11 in the population to the south of the river and for locus NIE10 in both populations. Both populations were found to be possibly in Hardy Weinberg equilibrium. Genotyping errors (calculated by Zhan *et al.* (2009)'s method) gave an ADO error rate of 0.24, and a FA error rate of 0.07, however the mean probability of genotyping error was 0.04. Null allele frequencies calculated FREENA (Chapuis & Estoup, 2007) were in the range 0.04–0.737 (Appendix Five).

Table 3.12. Average number of alleles across populations (N_a), observed (H_O) and expected (H_E) heterozygosities and departures from Hardy-Weinberg proportions (F_{IS}) for sampled populations and all loci of proboscis monkey.

		Population					Population		
Locus		North	South	N_a	Locus		North	South	N_a
D10	H_E.	0.62	0.55		P2D6	H_E.	0.64	0.74	
	H_O	0.82	0.40	6		H_O	0.64	0.68	7
	F_{IS}	-0.323	0.306			F_{IS}	0.020	0.105	
		NS	*				NS	NS	
E10	H_E.	0.63	0.82		P2F3	H_E.	0.64	0.68	
	H_O	0.43	0.26	9		H_O	0.50	0.63	5
	F_{IS}	0.338	0.693			F_{IS}	0.235	0.092	
		***	***				**	NS	
P1A6	H_E.	0.58	0.69		P3B2	H_E.	0.71	0.70	
	H_O	0.69	0.55	5		H_O	0.62	0.60	8
	F_{IS}	-0.175	0.222			F_{IS}	0.141	0.176	
		NS	*				*	*	
P1C5	H_E.	0.72	0.71		P4C11	H_E.	0.68	0.74	
	H_O	0.62	0.59	6		H_O	0.69	0.47	8
	F_{IS}	0.146	0.200			F_{IS}	-0.011	0.388	
		*	*				NS	***	

	Total	North	South
	H_E.	0.65	0.70
	(SD)	0.05	0.08
	H_O	0.63	0.52
	(SD)	0.12	0.14
	F_{IS}	0.053	0.282
		NS	***
	MNA	6	7

MNA- mean number of alleles per locus.
NS- non significant, ** - $P< 0.01$, *** - $P< 0.001$.

3.3.4. Population structure in proboscis monkeys of the LKWS

Despite the significant F_{IS} values, little structure was observed from the FCA plot of individual microsatellite genotypes (Figure 3.7). Twenty migrants per generation were detected and F_{ST} was equal to 0.012 ($P> 0.01$) indicating high levels of gene flow. As with the macaque populations, the Bayesian clustering performed using STRUCTURE considering the null alleles as recessive, estimated a log likelihood of the data, ln Pr(X | K), maximal at K = 2 (Table 3.13). When K increased (K = 3-5), the results showed a similar pattern as with K = 2 (Figure 3.8).

Table 3.13. Proboscis monkey inferred number of populations (K) by STRUCTURE. The posterior probability of the number of populations was maximal at K = 2.

K	1	2	3	4	5
ln Pr(X | K)	-1655.3	***-1620.3***	-1625.6	-1784.3	-1814.6

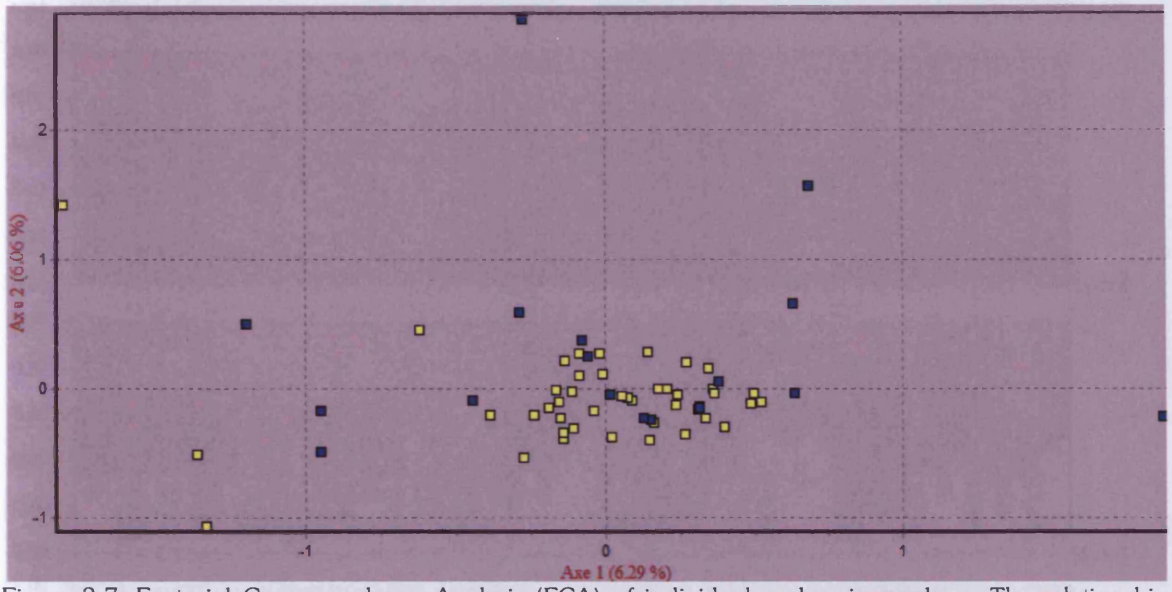


Figure 3.7. Factorial Correspondance Analysis (FCA) of individual proboscis monkeys. The relationship among multilocus genotypes of north (yellow) and south (blue) individuals is shown.

Following the same parameters as with the long-tailed macaques, one cluster (I) grouped together the population sampled on the south bank of the Kinabatangan River (q_I= 0.647). The population to the north of the river seems to be admixed since it was comprised of individuals from both clusters and, as a whole, the population could not be assigned to any one cluster (q_I= 0.41 and q_{II}= 0.59). The proportions of membership of each sampled population in the two inferred clusters are shown in Table 3.14. Within the population sampled to the south of the river, seven potential migrants were identified (average q_{II}= 0.768) and which mainly belong to Lot 6 (57%). However, when each Lot was ran in STRUCTURE as an independent population, two clusters were still recognized by the program but only Lot 1 (q_{II}), Lot 4 (q_I) and Lot 10 (q_I) met the condition of q_i> 0.600 and could be assigned to either cluster (Appendix Ten).

Table 3.14. Bayesian clustering analysis in proboscis monkeys performed with STRUCTURE. The table shows the proportion of membership (q) of each predefined sampled population in each of the two inferred clusters. The number in parenthesis and percentage of total individuals assigned (including migrants) are indicated. Proportions of membership of potential migrants are indicated with their original cluster (II) (see text for details). n = population size.

Population (n)	Cluster		Potential Migrants	Unassigned individuals	% of total assigned
	qI	qII			
North(45)	0.41(13)	0.59(25)	-	7	84.44
South(22)	0.647(15)	0.353(0)	0.768(7)II	0	100

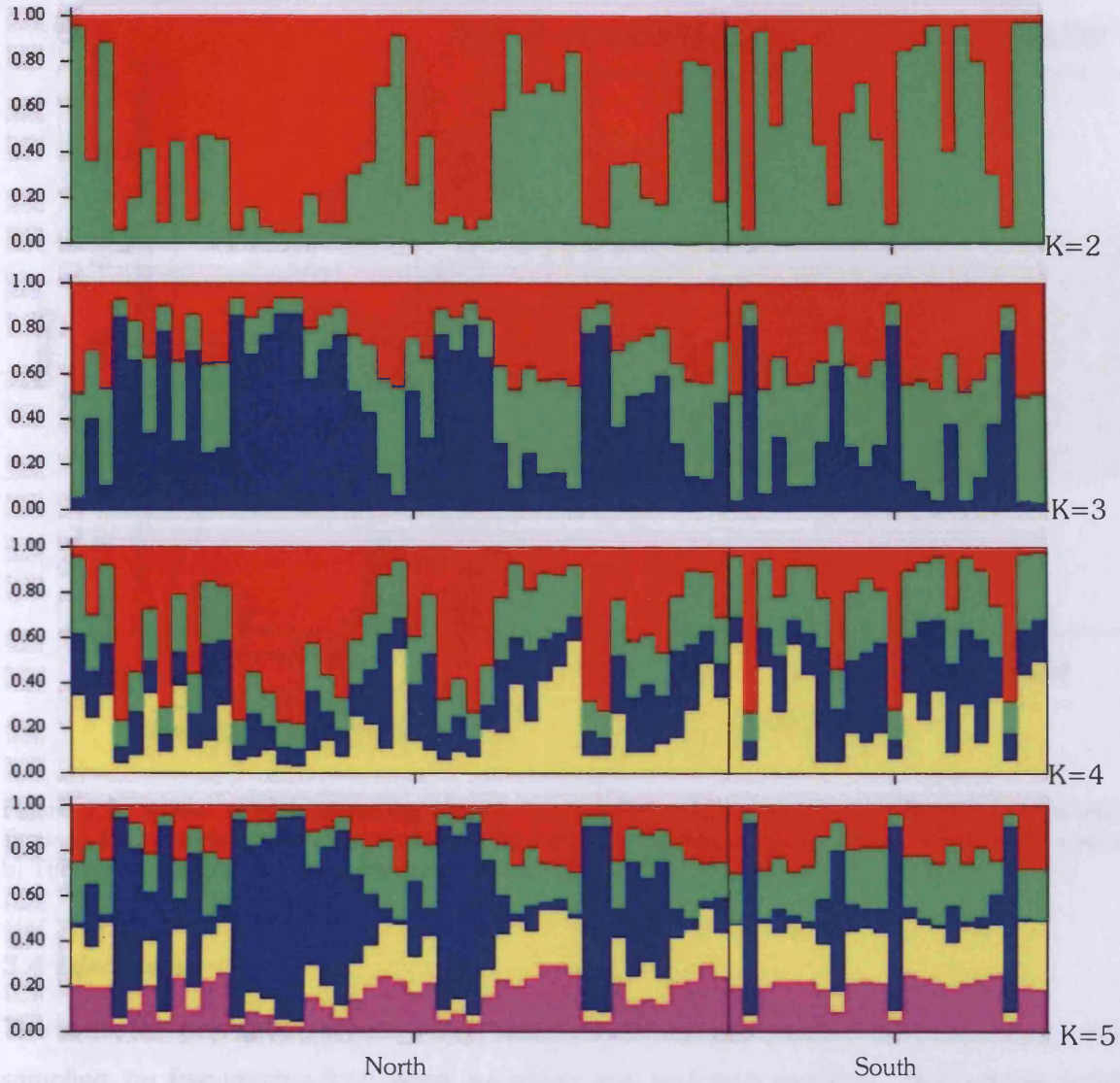


Figure 3.8. Proboscis monkey clustering results (K=2) for both sampled sites, according to STRUCTURE analysis. Each individual is represented as a vertical line partitioned into coloured segments. Sampled sites are separated by a black vertical line and labelled below the figure.

In the 10 runs performed with GENELAND, the posterior density and the log-likelihood levels quickly reached a plateau, indicating convergence. In eight runs, the modal number of genetic groups (K) was one, whereas the other two runs showed two clusters. Therefore, subsequent runs were performed with K fixed to 1. The 5 runs selected on the basis of their posterior probability reached also convergence. In all cases, the 67 individuals were assigned to one spatially defined population (Figure 3.9).

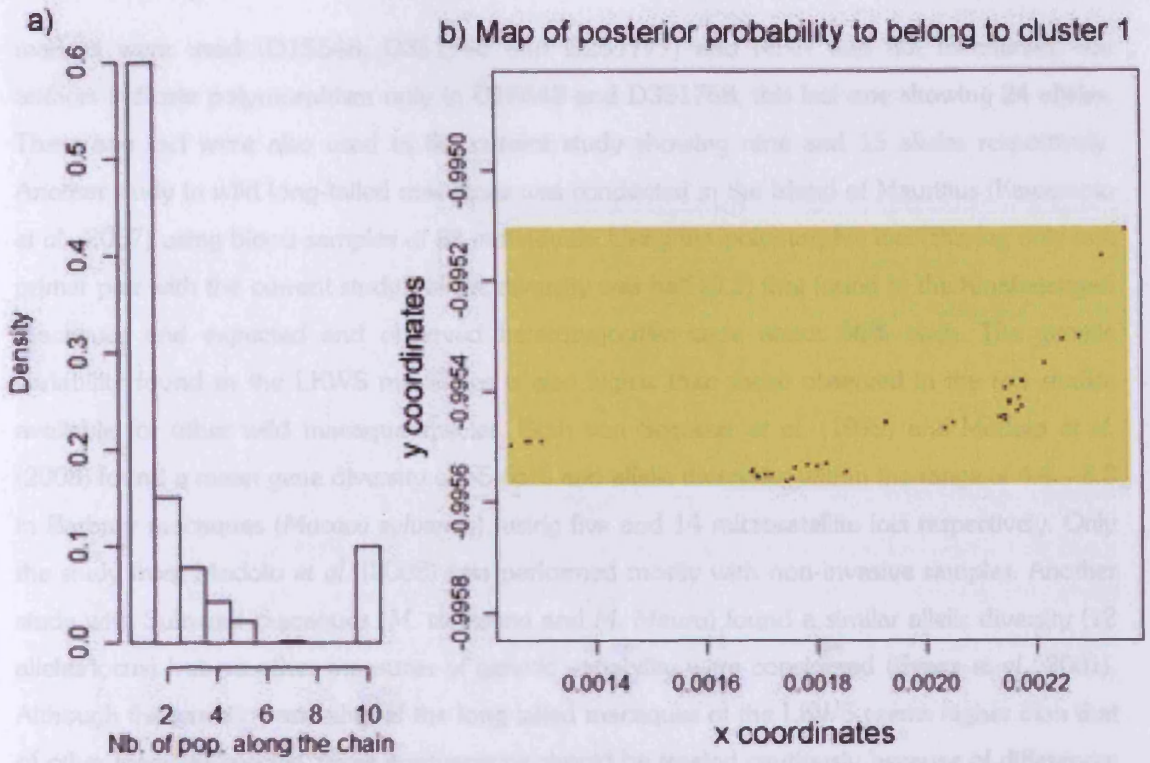


Figure 3.9. Spatial genetic clustering analysis of proboscis monkeys with GENELAND. a) Histogram showing the posterior density distribution of the number of clusters estimated from GENELAND analysis. b) The contour map shows the posterior probability to belong to one cluster (depicted in yellow here).

3.4 Discussion

This is the first population genetic study, based on microsatellite markers and using non-invasive sampling, on free-ranging long-tailed macaques and proboscis monkeys. Both species present high levels of genetic diversity and gene flow despite the presence of a potential natural barrier: the Kinabatangan River. Bayesian clustering analyses showed population admixture for both species, which is congruent with their ecology. However, null alleles and rates of allelic dropout above 20% were detected within both datasets. Since genotyping errors can bias population genetic studies (Dewoody *et al*, 2006, Hoffman & Amos, 2005; Roon *et al.*, 2005), potential causes for the estimated errors in this study are also discussed.

Long-tailed macaques, sampled from both sides of the Kinabatangan River within the Lower Kinabatangan Wildlife Sanctuary, appear to maintain a high level of genetic variability (MNA= 13.5, SD= 3.2 and mean H_E= 0.8; Table 3.9). These diversity values are higher than those observed in the few studies available for other wild long-tailed macaque populations. For example, Perwitasari-Farajallah *et al.* (2010) collected blood samples of 55 long-tailed macaque individuals from seven groups of Tinjil Island, Indonesia, and reported moderate genetic variability (H_O= 0.485, H_E was not reported). However, only three human-derived microsatellite

markers were used (D1S548, D3S1768 and D2S1777) and MNA was not mentioned; the authors indicate polymorphism only in D1S548 and D3S1768, this last one showing 24 alleles. These two loci were also used in the current study showing nine and 15 alleles respectively. Another study in wild long-tailed macaques was conducted in the Island of Mauritius (Kawamoto *et al.*, 2007) using blood samples of 82 individuals. Using ten polymorphic loci (sharing only one primer pair with the current study), allelic diversity was half (6.3) that found in the Kinabatangan macaques and expected and observed heterozygosities were about 66% each. The genetic variability found in the LKWS macaques is also higher than those observed in the few studies available for other wild macaque species. Both von Segesser *et al.* (1999) and Modolo *et al.* (2008) found a mean gene diversity of 55-65% and allelic diversities within the range of 4.4 – 8.3 in Barbary macaques (*Macaca sylvanus*), using five and 14 microsatellite loci respectively. Only the study from Modolo *et al.* (2008) was performed mostly with non-invasive samples. Another study with Sulawesi macaques (*M. tonkeana* and *M. Maura*) found a similar allelic diversity (12 alleles/locus) but no other measures of genetic variability were considered (Evans *et al.*, 2001). Although the genetic variability of the long-tailed macaques of the LKWS seems higher than that of other localities/species, these comparisons should be treated cautiously because of differences in the sampling design influencing the extent and resolution of the data (Storfer, 2007).

High levels of genetic diversity were also found in proboscis monkeys of the LKWS (MNA= 8.8, SD= 2.4 and mean H_E= 0.68; Table 3.12) despite the low estimated number of individuals (c.a. 1500) and the recognised decreasing trend in their population size (Goossens *et al.*, 2003a; Sha *et al.*, 2008). Within the "odd-nosed colobines" (Kirkpatrick, 2007) only the Yunnan snub-nosed monkey (*Rhinopithecus bieti*) has been subjected to a non-invasive population genetic study on wild individuals (n= 203) (Liu *et al.*, 2009). Using ten microsatellite loci, they report an overall mean H_O and H_E of 0.614 and 0.703 respectively, and allelic richness across loci being 7.5. Pan *et al.* (2005) also performed a population genetic study on an endangered odd-nosed colobine, the Sichuan golden monkey (*Rhinophithecus roxellana*) but relying on invasively collected samples (n=32). Mean expected and observed heterozygosities were of 0.64 and 0.613 respectively, and allelic diversity was of 3.9 across 14 loci. In a study with approx. 150 wild Hanuman langurs (*Semnopithecus entellus*), five microsatellites yielded similar results (MNA= 4.8, H_E= 5.9, H_O= 0.63) (Launhardt *et al.*, 1998), also using faecal material. Hence, the results obtained for the proboscis monkey here are similar to those reported for other free-ranging Asian colobines.

This high genetic diversity for both primate species is congruent with Jalil's (2007) mtDNA study and similar research on orang-utans' within the same area (Goossens *et al.*, 2005). Between 1950 and 2000, a third to one half of the original forest area has been cleared in Sabah

(McMorrow & Talip, 2001), and this habitat loss continues. As mentioned by Goossens *et al.* (2005) this trend could have forced primate groups into the Kinabatangan forests from other areas (adjacent forest), concentrating the surviving individuals in the forest patches along river, explaining the high values of genetic diversity (Jalil, 2007; current study).

Linkage Disequilibrium (LD) was found in some of the loci for both primate species. LD can be due to a variety of factors, including physical linkage, admixture and demographic fluctuation. Migration and admixture among two or more populations can generate LD, for example after recent introgression of novel haplotypes into a population, recombination may not have had time to break down LD (Hedrick, 1985, Slatkin, 2008). Changes in population size, particularly an extreme reduction in size (a population bottleneck), can also increase LD (Slatkin, 2008). Epistatic interactions between loci can also maintain LD, but this explanation seems less likely with supposedly neutral microsatellites markers. Hence, the significant LD found in this study might suggest admixture between groups of populations and/or population structure (Pfaff *et al.*, 2001; Pritchard *et al.*, 2010; Slatkin, 2008).

As the Kinabatangan River is only 200 m wide (Goossens *et al.*, 2005), it was not expected to be a barrier for both long-tailed macaques and proboscis monkeys. Indeed, high levels of gene flow were detected for the two primate species. F_{ST} values for populations to the north and south of the Kinabantangan River were of 0.005 and 0.012 for the long-tailed macaques and proboscis monkeys respectively ($P>$ 0.01). The number of migrants (Nm) calculated by GENETIX was above 15 also for both species. According to Frankham *et al.* (2002), an F_{ST} above about 0.15 is considered to be an indication of significant differentiation among fragments and more than 10 immigrants per generation are needed to prevent differentiation. The F_{ST} and Nm values were supported by the FCAs in GENETIX (Figs. 3.3 and 3.7) where no clear partition of populations was observed. The outputs from GENETIX were congruent with the results from GENELAND spatial where only one cluster was detected for each of the primate species (Figs. 3.5 and 3.9) inferring that the populations to the north and south of the Kinabatangan River are admixed. These results were predicted due to the known mobility of both species. Dispersal of long-tailed macaque males across rivers has been reported in Indonesia (de Ruiter, 1994; de Ruiter & Geffen, 1998) and also witnessed in the Kinabatangan (Jalil personal communication and personal observation). Proboscis monkeys are excellent swimmers (Fleagle, 1998) and have been found on many occasions swimming from riverside to riverside, also if frightened while swimming they can dive for several minutes (Bennett & Gombek, 1993). The proboscis monkey is indeed the most aquatic of all primates with several unique adaptations, including interdigital webbing on their feet and upturned nostrils (Davies & Oates, 1994; Fleagle, 1998). In addition to male emigration, female and infant transfer has also been confirmed by observed changes in group

composition (Matsuda 2008; Murai *et al.*, 2007). Anecdotally, a female with a baby was seen recently (August 2010) crossing the Kinabatangan River near Lots 5 and 6 of the LKWS (Gilmoore Belongon personal communication).

It was thus surprising to find apparent evidence of partitioning with STRUCTURE for both long-tailed macaques and proboscis monkeys (Tables 3.11 and 3.14, and Figs. 3.4 and 3.8). However, Pritchard *et al.* (2010) recommends re-evaluation of the results whenever (a) the proportion of the sample assigned to each cluster is roughly symmetric (~ 1/K in each inferred cluster) and, (b) there is no clear biological interpretation for the assignments. Both of these points do apply to the datasets of long-tailed macaques and proboscis monkeys since the clusters inferred by the program did not clearly delimit populations to the north and south of the river. Yet, the assignment of clusters was perhaps a reflection of human-mediated habitat fragmentation (Appendices Six and Ten). Genetic variation will be shared for some period of time between populations which have recently separated, even in the absence of gene exchange (Hey, 2006; Waples & Gaggiotti, 2006). Therefore, genetic differences due to a migration-drift balance will render a F_{ST} value which, by itself, cannot be distinguished from that yielded by an accumulation of genetic changes over time in completely isolated populations (Waples & Gaggiotti, 2006). These two scenarios could be distinguished by a non-equilibrium method such as the Isolation and Migration (IM) model developed by Hey and Nielsen (2004). The IM model allows variation in populations sizes over time and they are not assumed to be in migration-drift equilibrium (Hey & Nielsen, 2004; 2007). The current study did not evaluate this method due to time constraints, nevertheless it would be important to explore it further.

The current microsatellite results from STRUCTURE regarding the long-tailed macaques are somewhat comparable to those of Jalil (2007) where analysis of molecular variance from mtDNA data indicated restricted gene-flow between populations on the north and south side of the Kinabatangan River. However, Jalil (2007) stressed that a component of this genetic structure probably also arises due to the high level of genetic partitioning between demes associated with the social system of the long-tailed macaque. (de Ruiter *et al.*, 1994; de Ruiter & Geffen, 1996). For example, restriction enzyme analysis revealed that toque macaques in Polonnaruwa, Sri Lanka, exhibit two highly divergent haplotypes which occupied adjacent habitats (Hoelzer *et al.*, 1994). These two haplogroups were distinct despite a lack of known geographic barriers, but the distribution of these two haplotypes is consistent with known history of group fission (ibid). In contrast, the proboscis monkey results from STRUCTURE do not compare with the findings of Jalil (2007). Three distinct mitochondrial lineages were detected but there was no support for geographic partitioning between these lineages and no indication of any form of restricted gene flow across the Kinabatangan River. These results are more in agreement with the results of gene

flow and GENELAND clustering for proboscis monkeys obtained in this study. In addition, a possibility exist that distinct primate populations occurred at both sides of the Kinabatangan River. Animals often select habitats based on suitability (i.e. food availability and predation risk) (Weisser, 2001) and until 60 years ago, the forests at both sides of the Kinabatangan River should have offered enough resources, thus making dispersal across the river unnecessary or very limited. Deforestation implies restricted foraging areas and increased densities, conditions which might increase mobility between riversides despite crocodile predation risk. Thus the possibility that the clusters detected by STRUCTURE are a reflection of previous population structuring cannot be discarded.

With an apparent lack of population structuring it was unexpected to have positive significant values of F_{IS} across every loci and sampled populations of long-tailed macaques (Table 3.9), indicating homozygotes excess. Positive F_{IS} values were also found in the proboscis monkey dataset, but they were not statistically significant for all loci affected (Table. 3.12). According to Allendorf & Luikart (2007) and Dewoody *et al.* (2006) demographic or non-random mating system processes, such as a Wahlund effect or inbreeding, are expected to result in excess homozygosity at all loci, whereas errors due to stuttering, large-allele dropout and null alleles should affect only a subset of loci. In a study of kit foxes, Ralls *et al.* (2001) found positive F_{IS} values indicating a significant deficit of heterozygotes, similar to what was found for the long-tailed macaques. These levels were posteriorly explained by an inadvertent sampling bias as foxes living on adjacent home ranges tended to be more closely related than foxes that did not, largely because females on such ranges were often closely related. Ralls *et al.* (2001) suggested that this can be viewed as sampling across subpopulations at a very fine scale, thus creating a heterozygote deficiency due to the Wahlund effect. This explanation might also be plausible for the macaques, where all samples analysed came from one or a few individuals pertaining to different social groups whose ranges might have overlapped. With phylopatric females and the creation of new social groups by fission, the scenario presented by the macaques is akin to the one described for the foxes. A study including more individuals for each social group could resolve this incognita. Nevertheless, the homozygote excess can also be explained by genotyping errors and presence of null alleles common problems when working with faecal samples and heterologous primers.

Non-invasive genetic sampling (NGS) of wildlife populations is a powerful tool for assessing demography, gene flow, and population structure (Goossens & Bruford, 2009). However, a major issue relevant to the application of this technique is the reliability of the genotypes obtained and the amplification success of the markers by PCR. DNA extracted from faeces is usually of poor quality (possibly degraded or accompanied by PCR inhibitors, especially in faeces

that contains plant, bacterial and protozoan remnants, characteristic of folivorous primates such as the proboscis monkey) and quantity (often in the picogram range) (Broquet *et al.*, 2007; Launhardt, 1998; Monteiro *et al.*, 1997, Morin *et al.*, 2001; Vallet *et al*, 2008, Taberlet *et al.*, 1999). Furthermore, this low quantity DNA may be damaged due to exposure to moisture, heat and ultraviolet radiation encountered during collection (McKelvey & Schwartz, 2004) or be negatively influenced by the age of the sample and the season in which it is collected (Piggott, 2004).

All the samples tested in this study gave positive PCR amplification for mtDNA using species specific primers, a method regularly used by our group to verify DNA extraction success. However, we were able to reliably amplify nuclear DNA (at least 3 positive PCRs) from only 71% of the long-tailed macaque samples, and 47% of the proboscis monkey samples. Interestingly, most of the samples which gave amplification problems came from individuals sampled south of the Kinabatangan River (v.g. 35% of success for proboscis monkey samples from that side of the river vs. 61% from the north; $\chi^2= 7.68$, df= 1, $P<0.01$). As a comparison, several recent non-invasive studies obtained lower success rates, e.g., 57% in gorilla (Bergl & Vigilant, 2007) and 38% for the giant panda (He *et al.*, 2008), but our results were lower than that for golden-crowned sifakas (94%, Quéméré *et al.*,2009) where the extraction protocol of Vallet *et al.* (2007) was used. This protocol is based on the use of CTAB an extensively used compound for plant DNA extractions (Bhattacharjee *et al.*, 2009; Murray & Thompson, 1980; Porebsky *et al.*, 1997). Vallet *et al.* (2007) mentions that "acquired experience from a species is not easily reproducible to another species because not all protocols are equally efficient in all species". In a similar way, this might also be the case between the different DNA types amplified in this study. The extraction protocol used in this study was the same as the one used by Jalil (2007), who mentioned the difficulties of amplifying mtDNA and hence the need of concentrating the DNA after elution (personal communication). This step of the protocol might be suitable for mitochondrial genes, where the high copy numbers can compensate for the also concentrated levels of PCR inhibitors, but it might be less suitable for single copy nuclear genes. Perhaps an extraction method like Vallet *et al.* (2007) would have yielded better amplification results, especially for proboscis monkeys since their diet consists mostly of leaves (188 plant species, Matsuda, 2008). Attempts were made to re-extract DNA from a few (6) DNA samples, using CTAB/PCI and CTAB/QIAGEN DNeasy Plant Kit, with apparently better results for a mixed Stool/CTAB/Plant QIAGEN Kits but not enough samples could be re-tested using this method (data not shown). It would be interesting to test this method systematically and compare the results, especially between the two river sides since differences in dietary components might be inferred. On the other hand, the difference in amplification success could also be the result of the

nature of the markers used (species-specific vs. heterologous primers). Nevertheless the difference of markers could be more related to genotyping errors (ADO and FA) and the presence of null alleles since the microsatellite primers used for proboscis monkeys were species-specific and yet PCR amplification success was actually lower than that of the macaques'.

The genotyping error rates (overall 16-33% and 15-28% for the dropout allele for macaques and proboscis respectively) were comparable to those reported in other studies using faeces (Bhagavatula & Singh, 2006; Morin *et al.* 2001; Hansen *et al.* 2008; Quéméré *et al.*, 2009), but also higher than some others (Mondol *et al.*, 2009, Zhan *et al.*, 2009). To minimize these stochastic errors and improve the confidence in the individual genotypes, Taberlet *et al.* (1996) proposed a multi-tubes procedure consisting of several independent PCR amplifications for each sample and locus. However, because this approach is very expensive, time-consuming and requires higher template volume, it might not always be feasible (Valière *et al.*, 2002). Hence, different approaches have been suggested recently (Frantz *et al.*, 2003, Valière *et al.*, 2002) and applied with success (Hansen *et al.*, 2008; Piggott *et al.*, 2006; Quéméré *et al.*, 2009; Valière *et al.*, 2007). In the current study, the simulation method (GEMINI) of Valière *et al.* (2002) was used for the macaque dataset to determine the number of PCR repetitions needed to obtain an accurate genotype. Although 94% of genotyping accuracy was calculated by GEMINI, error rates were still high (see above). This could be due to the "pre-error rates" calculated by PEDANT (Johnson & Haydon, 2007) since GEMINI requires error rates for the input files prior to the determination of PCR repeats to obtain an accurate genotype. A pilot study using the original Taberlet *et al.* (1996) approach to determine the "pre-error rates" might have given a different result regarding the error rates and hence the number of PCR repetitions needed to obtain an accurate genotype. Comparatively, neither the Taberlet *et al.* (1996) nor the Valière *et al.* (2002) methods were used for the proboscis monkey dataset, due to constrains of samples volume, and still the error rates were similar to those estimated for the macaques, whereas the estimation of null alleles differed between species.

The screening of human primers with particular amplification conditions has shown that many loci are also informative in several non-human primates, including gorilla, chimpanzee, orang-utan, macaque and langur (Roeder *et al.*, 2009). However, there are known difficulties with cross-amplification between divergent taxonomic species (Chapuis & Estoup, 2007; Clisson, *et al.*, 2000). Null alleles result from polymorphism (substitution and indel mutations) in the annealing sites of microsatellite locus primers and they have been found in a wide range of taxa (Chapuis & Estoup, 2007; Dakin & Avise, 2004). The current results from MICROCHECKER suggest that null alleles exist in every loci screened on the long-tailed macaques. This result was unexpected as 14 out of the 15 loci screened in this study had been used either in long-tailed

macaques or in other species of the same genus confidently (Tables 3.3 and 3.4). However, the divergence time between hominoids and cercopithecids is 25 mya (Kumar & Hedger, 1998; Page & Goodman, 2001) hence it is highly probable that modifications on the molecular structure of regions surrounding the microsatellites occurred (although they were not sequenced as part of this study). A notable case of null alleles is that reported by Deucher *et al.* (2010), where a case of non-maternity was due to a single transversion in the primer annealing site and was solved by redesigning the primers. Additionally, ten percent of the studies revised by Dakin and Avise (2004) report the redesign of primers in order to solve the problem of null alleles. This is congruent with our proboscis monkey results since after testing almost 40 human derived microsatellite primers (Table 3.5) and obtaining spurious results the problem was solved when the markers were redesigned. Only three of the eight loci screened in the proboscis monkey showed evidence for null alleles, but this could be an effect of allelic dropout rather than true null alleles and this explanation cannot be overruled for the macaques either. In any case, the impact of null alleles was attempted to be minimized by taking them into consideration when performing the Bayesian clustering analyses (STRUCTURE), as previously recommended and applied (see metods section; Carlsson, 2008; Chapuis & Estoup, 2007; Seabra *et al.*, 2009).

Altogether, the results of the current study suggest that the level of genetic diversity in proboscis monkeys and long-tailed macaques is relatively high and that the Kinabatangan River is not a barrier for gene flow in the LKWS. This last result contrasts with data on orang-utans in the same area (Goossens *et al.*, 2005; Jalil *et al.*, 2008). Future studies should investigate the role played by human barriers such as oil palm plantations, riparian villages, or roads in the development of genetic differentiation between remaining forest patches. For instance, the Sandakan-Lahad Datu bridge (as shown on Fig. 3.1) may provide a significant barrier to current and future gene flow, especially for proboscis monkeys since Jalil (2007) was able to collect proboscis monkey samples throughout the area between the road and Lot 10 whereas we (only four years later) did not find any individual up to Lot 10. Future studies might also benefit from the use of a different DNA extraction method, such as the one of Vallet *et al.* (2007) or the mixed approach (QIAGEN Stool Kit/CTAB/QIAGEN Plant Kit) that we started testing. Additionally, the use of species-specific markers could resolve the problem of null alleles/allelic dropout; a new set of 499 microsatellite markers is now available for *Macaca fascicularis* (see Higashino *et al.*, 2009). Furthermore, a redesign of the sampling strategy, trying to cover more individuals within social groups and bearing in mind the possibility of neighbour mating (Fogelqvist *et al.*, 2010; Schwartz & McKelvey, 2008), might also contribute to a better detection of population substructuring. Despite all these problems, molecular studies are useful to identify threats to genetic diversity in order to

determine appropriate intervention strategies to maintain gene flow and diversity. The results of this study can be used as a baseline for conservation and management measures for the proboscis monkeys and long-tailed macaques of the LKWS.

CHAPTER FOUR

Characterisation and diversity of Mhc-DRB sequences in proboscis monkeys (Nasalis larvatus) and long-tailed macaques (Macaca fascicularis) from Malaysia

Salgado-Lynn M[1], **Goossens B**[1,2], **Cable J**[1], **Bruford MW**[1]

[1]*School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK*
[2]*Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia*

4.0 Abstract

Knowledge of a population's genetic diversity often provides valuable information for conservation strategies and management. In vertebrates, growing evidence suggests that genetic diversity is particularly important at the Major Histocompatibility Complex (MHC) because its gene products play an important role in immune function. Thus, identification of MHC genes can provide a core component for genetic studies examining fitness and adaptation to fragmented ecosystems. In this preliminary study, PCR, sequencing and haplotype reconstruction identified *Mhc-DRB* sequences from 15 wild proboscis monkeys (*Nasalis larvatus*) and 36 long-tailed macaques (*Macaca fascicularis*) from Sabah, Malaysia. Using generic DRB primers, five proboscis monkey *–DRB* (*Nala-DRB)* and 47 long-tailed macaque sequences (*Mafa-DRB)* were identified. This represents the first MHC sequences reported for proboscis monkeys, and only one of the sequences detected in the macaques had been previously described. In neighbour-joining phylogenetic trees, the *Nala-DRB* sequences form a monophyletic group likely belonging to the *Mhc-DRB5* locus, while the *Mafa-DRB* sequences represent four *–DRB* loci and at least six different lineages. Evidence for positive selection was found for the long-tailed macaque sequences but not in the proboscis monkey, possibly due to the limited number of sequences analysed. This study represents a first step towards understanding the effects of habitat fragmentation on fitness of long-tailed macaques and proboscis monkeys of the Lower Kinabatangan Wildlife Sanctuary.

4.1 Introduction

Genetic diversity is important for the maintenance of population viability and essential for preserving the evolutionary and adaptive potential of species (Holderegger *et al.*, 2006). Traditionally such diversity has been routinely measured using quasi-neutral molecular markers, particularly mitochondrial DNA (mtDNA), microsatellites or single nucleotide polymorphisms

(SNPs; Holderegger *et al.*, 2006; Kohn *et al.*, 2006). Such data provides valuable information for demographic analysis, and allows us to examine factors such as population size fluctuations, dispersal patterns, genetic structure (and hence gene flow), and to classify individuals by relatedness and paternity analyses (examples in Allendorf & Luikart, 2007 and Frankham *et al.*, 2002). However, variation at neutral loci cannot provide direct information on selective processes involving the interaction of individuals with their environment, or on their capacity for future adaptive change (Meyers & Bull, 2002). Both issues are of particular relevance for conservation therefore the use of adaptive molecular markers has been advocated (Crandall *et al.*, 2000; DeSalle & Amato, 2004; Vernesi *et al.*, 2008). In addition, research on a variety of taxa in different contexts and situations has revealed that evolution can occur over rapid timescales, often within decades (summarised in Stockwell *et al.*, 2003). In some cases, the time span between the separation of populations might even be too short to leave a signal at neutral loci so that differences between populations are only detectable at genes under selection (Cohen, 2002), such as those of the Major Histocompatibility Complex (MHC) (Sommer, 2005).

The MHC is central to the vertebrate immune system, being a multigene family that encodes key receptor molecules that recognise and bind foreign peptides for presentation to specialist immune cells and subsequent initiation of an immune response (Piertney & Oliver, 2006). The MHC has key biological functions with respect to resistance to infectious diseases, mate choice, kin recognition and reproductive success (Bernatchez & Landry 2003; Piertney & Oliver 2006). The genes of the MHC represent one of the most well-established systems available in vertebrates to investigate how natural selection can promote local adaptation at the DNA level (Bernatchez & Landry, 2003). Two major subfamilies, denoted class I and class II, which differ in levels of expression, structure of the encoded protein and function, control the immunological self/non-self recognition (Hughes, 2008; Penn & Ilmonen, 2005). MHC class I genes are expressed on virtually all nucleated somatic cells and their products are essential for immune protection from intracellular pathogens. MHC class II genes are only expressed on specific antigen-presenting cells such as B cells and macrophages, and their coding proteins bind and present peptides mainly stemming from extracellular parasites (Bernatchez & Landry, 2003, Piertney & Oliver, 2006; Knapp, 2005a; Penn & Ilmonen, 2005; Hughes, 2008).

The two MHC classes have the highest known variation of any functional vertebrate genes (Robinson *et al.*, 2003). In the codons encoding the peptide-binding region (PBR) of both class I and class II MHC molecules, a highly unusual pattern of nucleotide substitution is observed (Garamszegi *et al.*, 2009; Hughes & Nei, 1988, 1989). The ratio of non-synonymous (amino acid altering) to synonymous (silent) nucleotide substitutions per non-synonymous and synonymous site (d_N/d_S) can be used to detect selection at any gene region (Hughes & Yeager, 1998).

Synonymous mutations do not affect amino-acid composition and are therefore effectively neutral, non-synonymous mutations, however, that do alter amino acid composition, are more likely to be under selection (reviewed by Piertney & Oliver, 2006). If selection favours diversity at some genes, advantageous non-synonymous mutations will be retained and a high ratio of non-synonymous to synonymous substitutions will be observed (Hughes & Nei, 1988). The PBR codons are under various types of selection (e.g. Bernatchez & Landry 2003; van Oosterhout, 2009) which increases genetic diversity at MHC loci and enables a wider recognition of parasites (Klein, 1986). Functional MHC loci are characterised by an extensive repertoire of alleles in virtually all natural populations investigated so far (Garamszegi *et al.*, 2009). This variation suggests that there has been evolutionary pressure for organisms to combat a wide range of immunological challenges (Hughes, 2008; Hughes & Nei, 1998).

In humans (where the MHC is referred to as the human leucocyte antigen, HLA) the MHC II molecule is a heterodimer consisting of transmembrane chains α and β coded by distinct genes. These genes include three classical regions designated *-DR*, *-DQ* and *-DP*, each containing multiple A or B loci (Bontrop, 2006). In particular, the second exon of the *HLA-DRB* encodes the α1 /β1 domains, comprising four functional loci (*-DRB1*, *3*, *4*, and *5*), that form the functionally important antigen-binding sites (ABS) or PBR. In addition to the functional loci, five pseudogene loci have been identified (*-DRB2*, *6*, *7*, *8* and *9*) (Hughes, 2008; Hughes & Yeager, 1998; Hughes & Nei, 1989; Marsh *et al.*, 2005). Each *–DRB* locus is further divided into phylogenetic lineages or clusters based on nucleotide and amino acid sequence motifs (i.e. *HLA-DRB1*01*, *-DRB1*02*, etc.) and, to date (September 2010), a total of 902 *HLA-DRB* sequences have been described (IMGT/HLA Sequence Database; see Robinson *et al.*, 2003). The *Mhc-DRB* gene also appears to be polymorphic in all non-human primate species studied thus far (Bontrop, 2006). For instance, many of the *-DRB* alleles of *Macaca* sp. belong to loci/lineages that are shared between humans and macaques: namely, *-DRB1*, *-DRB3*, *-DRB4*, and *-DRB5*, as well as *DRB6*, with the latter appearing to be a pseudogene in all primate species studied. In addition, loci/lineages for which no human equivalent is known are present in macaques. These are named *-DRB*W*, and various *-DRB*W* loci/lineages have been defined (IPD-MHC; see Robinson *et al.*, 2005, 2010).

Humans and rhesus macaques have been the primary focus of primate MHC research, thus information on *Mhc-DRB* of other Old World monkeys is limited. Most of the information is restricted to other macaque species (*M. arctoides*, *M. fascicularis*, *M. fuscata*, *M. nemestrina* and *M. silenus*), two baboon species (*Papio hamadryas* and *P. ursinus*), a drill (*Mandrillus sphinx*) and the vervet monkey (*Chlorocebus aethiops*) (see IPD-MCH database, Robinson *et al.*, 2005, 2010). As pointed out by Lukas *et al.* (2004), most of this information was based on studies using

captive individuals. *-DRB* information on wild non-human primates is limited (but see Lukas *et al.*, 2004 and O'Connor *et al.*, 2007) perhaps due to problems associated with obtaining suitable samples from feral animals. Taking blood samples is impossible without disturbing behaviour and potentially compromises the welfare of wild primates and other endangered animals. However, this has led to the development of techniques utilizing DNA obtained from non-invasively collected samples such as faeces or hair (reviewed in Goossens & Bruford, 2009). In the current study, DNA obtained from faecal samples was used to characterize variation at the *-DRB* loci of the MHC class II complex in two free-ranging primates: the proboscis monkey (*Nasalis larvatus*) and the long-tailed macaque (*Macaca fascicularis*).

Long-tailed macaques and proboscis monkeys are sympatric primates of the Lower Kinabatangan Wildlife Sanctuary (LKWS), in Sabah, Malaysia. While the former inhabits a variety of habitats and is sometimes classified as a pest (van Schaik *et al.*, 1996; Abegg & Thierry, 2002), proboscis monkeys are endemic to Borneo and are classified as endangered by IUCN (IUCN, 2010; Meijaard & Nijman, 2000; Sha *et al.*, 2008). Population genetic studies using mtDNA data revealed high levels of genetic diversity in both species (Jalil, 2007). Due to the quasi-neutral nature of this marker, little can be inferred about genetic components of fitness for these two species. While molecular studies have shown that long-tailed macaque *Mafa-DRB* loci are highly variable, there is no information on the MHC variability of proboscis monkeys. To date, 171 *Mafa-DRB* sequences have been identified (Blancher *et al.*, 2006; Doxiadis *et al.*, 2006, 2010; Kriener *et al.*, 2000; Leuchte *et al.*, 2004; O'Connor *et al.*, 2007) and assigned locus and lineage designations based on their similarity to human sequences. O'Connor *et al.* (2007) reported sequences from wild *M. fascicularis*, identifying 12 new *Mafa-DRB* sequences.

This preliminary study aimed to (1) identify and characterize *Mhc-DRB* alleles in proboscis monkeys, (2) determine variation within the *Mhc-Mafa-DRB* loci in long-tailed macaques of the LKWS and (3) to consider the ratio of synonymous (d_S) to non-synonymous (d_N) nucleotide substitutions within and outside the antigen binding site (ABS) for evidence of selection processes.

4.2 Methods

4.2.1 Study site

In 2005, the state government of Sabah gazetted almost 27 000 ha of forest in the Lower Kinabatangan flood plain (5°20' – 5°45' N, 117°40' – 118°30' E) as a wildlife sanctuary under the Wildlife Conservation Enactment 1997. The Lower Kinabatangan Wildlife Sanctuary

comprises 10 sectors or Lots (Lots 1–10, with Lot 10 divided into 10A–C) linking seven patches of protected forests (Virgin Jungle Forest Reserves, VJFR) totalling about 15 000 ha. They are connected with 10 000 ha of state and private forests at various stages of degradation (Ancrenaz *et al.*, 2004).

4.2.2 Collection and preservation of faecal samples

Sampling covered all Lots of the LKWS in two seasons (October 2007-March 2008, June-November 2008). The behaviour of proboscis monkeys and long-tailed macaques allows the Kinabatangan River to be used as a transect to perform censuses of the primates during their inactive periods (particularly at sunset and early night). A total of ~660 km of riverbank including smaller tributaries were covered in expeditions that lasted from five to ten days, with teams of 2 to 5 people. Faeces were collected at dawn to ensure freshness, and occasionally at dusk. GPS (Garmin eTrex Vista HCx) coordinates of collected samples were noted for subsequent localisation of groups and populations (Appendix One). Samples were stored in 50 ml Falcon tubes with 70% ethanol following the protocol described by Goossens *et al.* (2003b) to avoid contamination. Within the constraints of time and logistics, we attempted to sample as widely as possible within each population seen in the riparian forest; because individual recognition was not always possible, and being unhabituated primates, some individuals may have been sampled more than once.

In addition to the LKWS samples, eight proboscis monkey stool samples were donated by Dr. Henry Bernard (Unit for Primate Studies Borneo, Institute of Tropical Biology and Conservation, Universiti Malaysia Sabah). These samples were collected along the Garama River in Klias (Northern Sabah), and were stored in 95% ethanol.

4.2.3 Blood and tissue samples

Four samples of DNA from B-lymphocytes of long-tailed macaques were provided by INPRIMAT (EU FP5 ID: QLRI-CT-2002-01325) and used as positive controls in PCR reactions. Tissue samples (muscle), skin and hair from a deceased juvenile female and a new born proboscis monkeys were donated by Lok Kawi Wildlife Park. Additionally, samples (muscle, liver and spleen) from two road-killed male proboscis monkeys were donated by Sepilok Orangutan Rehabilitation Centre (Sabah, Malaysia). All proboscis monkey tissue samples were stored in 95% ethanol and kept at -70°C.

4.2.4 DNA Extractions

DNA was extracted using the QIAamp DNA Stool Mini Kit (QIAGEN GMBH, Germany) following a protocol previously used by Goossens *et al.* (2000) with a modification in the last step, where DNA was eluted in 100 μl instead of 200 μl, and was subsequently concentrated to ca.70 μl (Jalil, 2007). Sample concentration was performed in a lyophilizer (Concentrator Eppendorf 5301) using the manufacturer's protocol "Aqueous liquids mode". Tissue samples were extracted with DNeasy Blood & Tissue Kit (QIAGEN) following the recommendations of the manufacturer, with minor modifications during elution (namely, 5 min incubation at 70°C with buffer AE, which was also preheated at the same temperature). These types of samples were used as positive controls in PCR reactions. To evaluate possible contamination, DNA was extracted from hair of the author (referred to hereafter as Human DNA) according to a protocol describe by Jalil (2007).

Successful DNA extractions from faeces were assessed by PCR amplification of a partial mitochondrial control region fragment using species specific primers under a previously described protocol (Jalil, 2007). DNA from the muscle/hair samples was quantified by microtiter spectrofluorometry (Quant-iT™ PicoGreen®, Invitrogen), and DNA concentrations were calculated using the SOFTmax® Pro software (Molecular Devices). In addition, all PCR products were visualized in agarose gels (1.5%).

Although over 300 samples were available for each primate species after DNA extraction, only 44 long-tailed macaque and 33 proboscis monkey samples were selected to characterize DRB alleles. This decision was made given the problems to consistently amplify microsatellite loci (15 and eight loci for long-tailed macaque and proboscis monkey respectively; Chapter Three) from the same samples. Only samples that produced three positive PCR reactions with consistent microsatellite genotypes were used (Table 4.1).

Table 4.1. Summary of primate stool samples selected to characterize *Mhc-DRB* alleles. The number of tested (T) and sequenced (S) samples are shown per primate and location.

	South Kinabatangan				North Kinabatangan							
Primate	L1	L3	L6	L9	L2	L4	L5	L7	L8	L10 (A-C)	Klias	Total
	T/S	T/S	T/S	T/S	T/S	T/S	T/S	T/S	T/S	T/S	T/S	T/S
LTM	3/1	6/5	6/6	8/8	4/3	4/4	3/3	3/3	4/4	3/3	0	44/40
PM	3/0	7/5	5/4	0	2/1	5/4	3/3	2/1	0	2/1	4/2	33/21

L- Lot inside the LKWS, to the north or south of the Kinabatangan River.
LTM – long-tailed macaque, PM – proboscis monkey.

4.2.5 Mhc-DRB *amplification*

PCR optimisation was performed with the DNA extracted from the INPRIMAT and the proboscis monkey tissue samples. Reactions were made to a total volume of 15 µl, containing 1ng of template, with final concentrations of 1.5 mM $MgCl_2$ (Promega), 0.25 mM of each deoxyribonucleotide triphosphates (dNTPs), 1X GoTaq® Colorless Flexi Buffer (Promega), 0.02 U/µl GoTaq® Hot Start Polymerase (Promega) and 0.5 µM of each oligonucleotide primer of the pairs (1) MDRB5 and 3'DRBseq, for long-tailed macaque samples, and (2) 5'DRB(*SalI*) and 3'DRB(*XbaI*), for proboscis monkey samples (Table 4.2). The first primer pair has successfully amplified *Mhc-DRB* alleles in *Macaca fascicularis* (see Leuchte *et al.*, 2004) and the other has been used to characterize alleles in *M. mulatta* (see Doxiadis *et al.*, 2003). This second pair was selected to screen the proboscis monkey samples as it has successfully amplified langur samples (Doxiadis personal communication, unpublished). The amplification conditions were as follows: 95°C for 2 min, 35 cycles at 94°C for 30 s, 50-65°C for 60 s, 72°C for 60 s and a final extension at 72°C for 10 min, and were performed in a Veriti® 96-Well Thermal Cycler (Applied Biosystems). After PCR, 5 µl of the product was electrophoresed in 2.0% agarose-0.5X TBE gels, and visualized using 1.5 ng/ml EtBr in a GelDoc-IT™ Imaging System (UVP) to confirm successful amplification of DNA samples.

Table 4.2 Oligonucleotide primers used to amplify and sequence *Mhc-DRB* alleles, with optimized annealing temperatures (TA).

Name	Sequence (5' - 3')	Optimized T_A
MDRB5	GCCTCGAGTGTCCCCCCAGCACGTTTC	62°C
3'DRBseq	TGTAAAACGACGGCCAGTCACCTCGCCGCTGCACTGT	
5'DRB(*SalI*)	CCGGTCGACTGTCCCCCCAGCACGTTTC	58°C
3'DRB(*XbaI*)	TCTAGATCACCTCGCCGCTGCACTGT	
DRBseq[a]	TCGAGTGTCCCCCCAGCA	n/a
M13 forward (-20)[b]	GTAAAACGACGGCCAG	n/a

[a] Leuchte *et al.*, 2004; [b] McGinnis *et al.*, 1995

Optimal PCR conditions for faecal samples were as described above, save for the optimized annealing temperatures (Table 4.2) and with an increase in the cycling parameter to 45. In addition, in an attempt to minimize allelic dropout, an approach similar to Taberlet *et al.*'s (1996) was used. In brief: at least two independent PCR reactions were performed for each tested sample followed by two independent forward and reverse copies from a single PCR product.

4.2.6 Mhc-DRB *sequencing*

Forty long-tailed macaque and 21 proboscis monkey samples produced a detectable amplicon of the appropriate size (approx. 300 bp; Table 4.1). Unattached primers, primer-dimers and

unincorporated dNTPs were removed prior to sequencing with an enzymatic reaction. In brief: seven µl of PCR product were mixed with 0.25 U of Shrimp Alkaline Phosphatase (SAP) and 2.5 U of *Exo*I to a total volume of 10 µl, and incubated at 37°C for 1 h, followed by inactivation of the enzymes at 80°C for 15 min. Cleaned products were sequenced using either a forward (DRBseq for macaques or 5'DRB(*Sal*I) for proboscis) or a reverse primer (M13 (-20) for macaques and 3'DRB(*Xba*I)) (Table 4.2). Sequencing reactions were performed in a final volume of 8 µl, containing 2µl of purified PCR product, 2.5µl Better Base (Web Scientific), 0.5µl BigDye Terminator V.1.1 (Applied Biosystems), and 0.2µM of primer. PCR was then carried out following an initial denaturation for 5 min at 96°C followed by 25 cycles at 96°C for 15 s, 50°C for 10 s and 2 min of extension at 60°C. Products were subsequently precipitated by adding 12 µl ultrapure H_2O, 5 µl of 125 mM EDTA and 60 µl of 100% ethanol. The reactions were mixed by gentle inversion, left to stand for 1 hr, and then centrifuged for 45 min at 2000 g. The supernatant was decanted and tubes were centrifuged upside down for 1 min at 200 rpm to eliminate ethanol residuals. Sixty µl of 70% ethanol were added to each PCR tube followed by centrifugation for 20 min at 2000 g. Ethanol residuals were eliminated as in the previous step plus tubes were also left standing open for an hour or until no odour was detected. PCR products were run individually on an Applied Biosystems 3130xl Genetic Analyzer using a 50 cm capillary array with POP-7 polymer.

4.2.7 Sequence analysis and haplotype reconstruction

Electropherograms were analyzed using SEQUENCHER 4.9 (Genecodes Corp), and corrected by eye. As a biparentally inherited gene, the forward and reverse sequences of every individual were treated as independent sequences and contigs were created to detect unique alleles. Contigs were converted into consensus sequences and all unique alleles were exported in FASTA format in separate files for each primate species. Every sequence was compared to those reported in GenBank®, using the Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi), to verify they corresponded to *Mhc-DRB* genes.

Sequences with ambiguities were resolved using a Bayesian statistical method for reconstructing haplotypes. To identify phase ambiguities, a preliminary sequence alignment was performed in MEGA 5.0 Beta (Tamura *et al.*, 2007) using Clustal W and default parameters. Once the ambiguities were localised, reconstruction of alleles was performed in PHASE 2.1.1 (Stephens *et al.*, 2001; Stephens and Donnelly, 2003). As suggested by Stephens *et al.* (2004), to obtain reliable results the algorithm was applied five independent times with different seeds for the random number generator (784, 1328, 2509, 2831, and 3844, generated from 1 to 5000 by

RANDOM.ORG, www.random.org). The number of MCMC iterations was set to 1000 with a thinning of 100 and a burn-in of 1000. The haplotype frequency estimates and the goodness-of-fit measures were checked for consistency across runs. With consistency achieved, final haplotypes were predicted on a second step, where PHASE was set to automatically run three independent times, the results corresponded to the run with the best average goodness-of-fit measure. Finally, repeated sequences were detected by DAMBE 5.0.47 (Xia, 2000; Xia & Xie, 2001).

4.2.8 Construction of phylogenetic trees

Phylogenetic trees were constructed by neighbour-joining/Jukes-Cantor method using MEGA 5.0 Beta (Kumar *et al.*, 2007), to infer evolutionary history and distances respectively. Published –*DRB* sequences were obtained from the Immuno Polymorphism Database (IPD, see Robinson *et al.*, 2005, 2010). Proboscis monkey sequences were compared by Clustal W alignment to selected –*DRB* published sequences different taxa (Table 4.3). Long-tailed macaque sequences were aligned with all the published *Mafa-DRB* sequences (171 to July 2010), some selected sequences from other macaque species, and a few sequences from old world monkeys (Table 4.3). The analyses were based on 265 nucleotides, and all ambiguous positions were removed for each sequence pair. Bootstrap analyses using 1000 replications were performed to determine the repeatability of sequence-clustering patterns.

Table 4.16. Summary of published –*DRB* sequences that were aligned to putative proboscis monkey (PM) and long-tailed macaque (LTM) –*DRB* sequences.

-DRB	Primate Species/Common name	Aligned with
NWM		
Pipi-	*Pithecia pithecia*/White-faced Saki	PM
Saoe-	*Saguinus oedipus*/Cotton-top Tamarin	PM
OWM		
Chae-	*Chlorocebus aethiops*/Vervet monkey	PM
Mafa-	*Macaca fascicularis*/Long-tailed macaque	PM/LTM
Mamu-	*Macaca mulatta*/ Rhesus macaque	PM/LTM
Maar-	*Macaca arctoides*/Stump-tailed macaque	PM/LTM
Mane-	*Macaca nemestrina*/Pig-tailed macaque	PM/LTM
Masi-	*Macaca silenus*/Lion-tailed macaque	PM/LTM
Mafu-	*Macaca fuscata*/Japanese macaque	PM/LTM
Masp-	*Mandrillus sphinx*/Mandrill	PM/LTM

-*DRB*	Primate Species/Common name	Aligned with
Paur-	*Papio ursinus*/Chacma baboon	PM/LTM
Paha-	*Papio hamadryas*/Hamadryas baboon	PM/LTM
Apes		
Popy-	*Pongo pygmaeus*/Orang-utan	PM
Papa-	*Pan paniscus*/Bonobo	PM
Patr-	*Pan troglodytes*/Chimpanzee	PM
Gogo-	*Gorilla gorilla*/Gorilla	PM

NWM-new world monkeys, OWM-old world monkeys

4.2.9 Statistical analyses

Patterns of nucleotide substitution were also analysed with MEGA 5.0 Beta (Tamura *et al.*, 2007). Seventeen variable amino acid positions, presumed to represent the ABS (Brown *et al.*, 1988, 1993), were assigned as for rhesus macaques (positions 11, 13, 23, 26, 28, 30 and 37 of the β sheet and 57, 61, 67, 70, 71, 73, 74, 77, 78 and 86 of the α helix, Slierendregt *et al.*, 1992). Rates of non-synonymous (d_N) and synonymous (d_S) substitutions within the ABS and non-ABS were calculated according to Nei and Gojobori (1986), with the Jukes and Cantor (1969) correction. Significance levels were determined using a Z test ($P< 0.05$) in MEGA 5.0 Beta.

4.3 Results

4.3.1 Amplification, sequencing and identification of proboscis monkey –DRB *alleles*

Mamu-DRB primers, 5'DRB(SalI) and 3'DRB(XbaI), successfully amplified *Nala-DRB* (*Nasalis larvatus*) alleles. Twenty one samples produced a PCR product of the expected size, however sequencing reactions were only successful for 15 (two from Klias and the remainder from the LKWS). Five different *Nala-DRB* (*Nala-DRB*Allele1* to 5) sequences, 265 nucleotides in length, were identified by direct sequencing in two heterozygote and 13 homozygote individuals (full sequences in Appendix Eleven). None of these alleles was found in published sequences from the aligned taxa; predicted nucleotide and amino acid sequences are presented in Figures. 4.1 and 4.2.

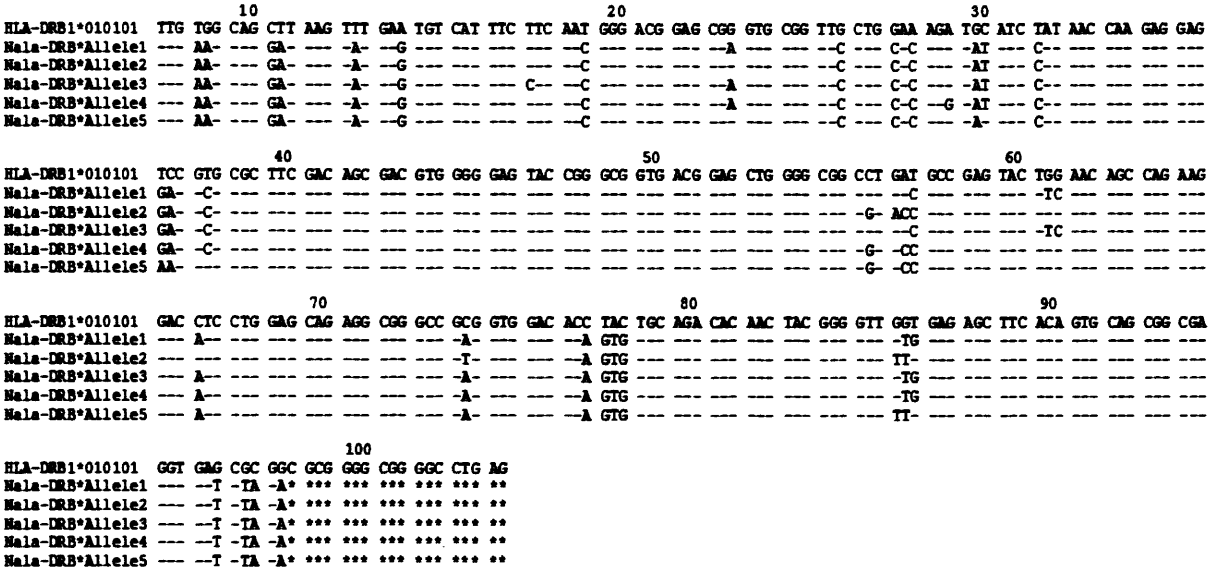


Figure 4.6. Nucleotide sequence of proboscis monkey DRB (exon 2). Codon numbering is shown above the consensus of –DRB sequences. Identity to consensus is shown by dashes (-) and asterisks (*) denote unsequenced nucleotides).

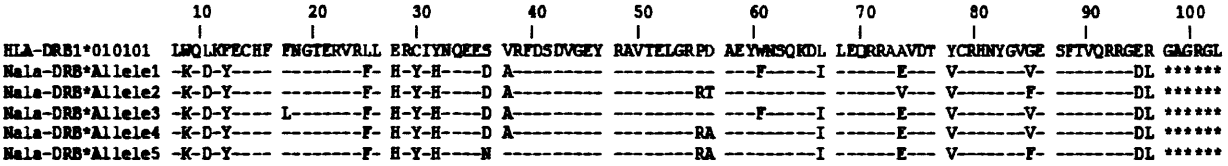


Figure 4.7. Predicted amino acid sequences of proboscis monkey DRB (exon 2). Numbering of amino acids is shown above the consensus of *–DRB* sequences. Red typeface indicates inferred ABS sites. Identity to consensus is shown by dashes (-) while unsequenced regions are denoted by asterisks (*).

Phylogenetic analysis of the *Nala-DRB* sequences was performed to identify the potential locus and lineage of each sequence. The neighbour-joining tree included the new sequences and selected *Mhc-DRB* sequences from New World (NWM), Old World monkeys (OWM) and apes. The five *Nala-DRB* sequences seem to be monophyletic and clustered within the *–DRB5* locus (Fig. 4.3).

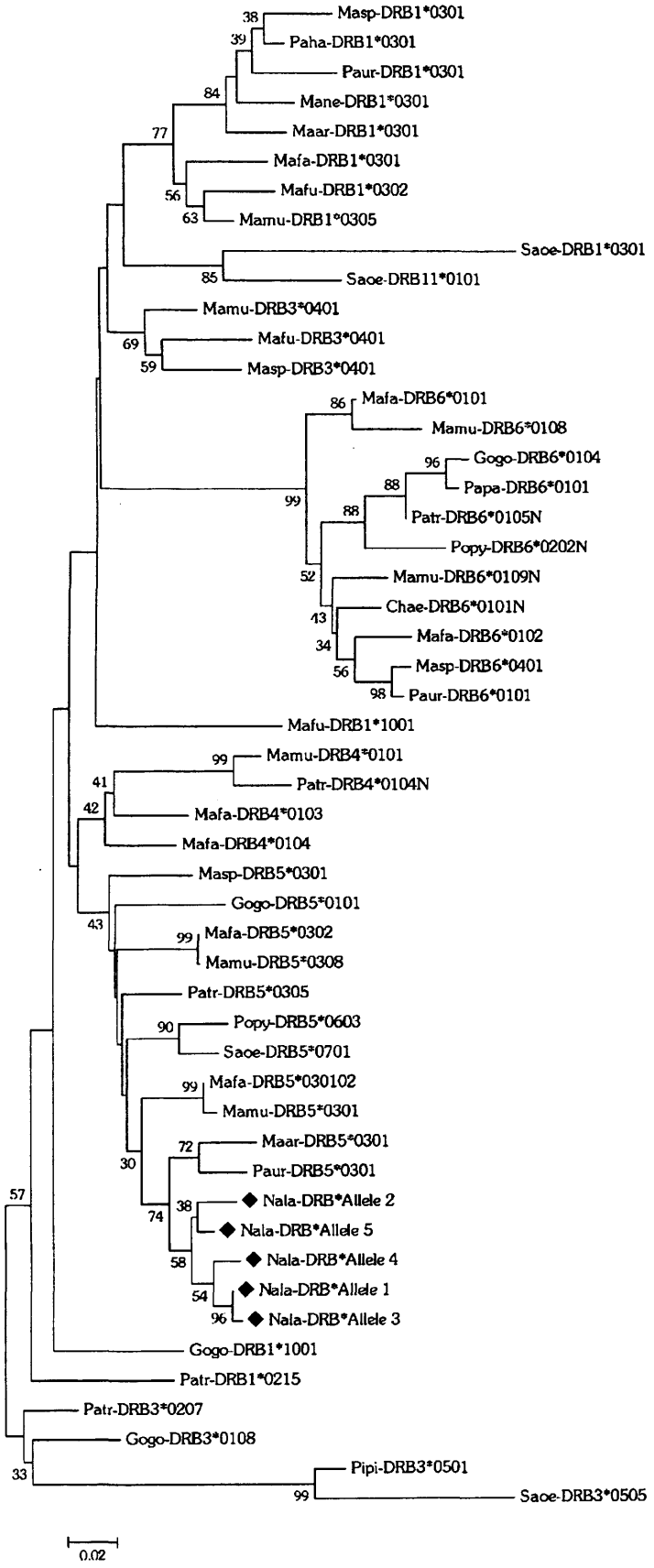


Figure 4.3. Phylogenetic tree of five *Nala-DRB* sequences (◆) and other selected taxa (see Table 4.3 for details). The tree configuration was derived from 51 nucleotide sequences using the neighbour-joining and Jukes-Cantor methods in MEGA 5.0 Beta. Numbers on the branches refer to confidence probability values.

4.3.2 Amplification, sequencing and identification of Mafa–DRB *alleles*

Exon 2 *Mafa-DRB* sequences were amplified using MDRB5 and 3'DRBseq primers. Successful sequencing reactions occurred for 36 out of 40 individuals screened. Direct sequencing identified 13 different sequences, 265 nucleotides in length, corresponding to 13 homozygote individuals. Haplotype reconstruction by PHASE 2.1.1 identified 36 more sequences in the remaining 23 heterozygote individuals. Subsequent analysis on DAMBE 5.0.47 revealed a total of 47 different sequences in the 36 individuals (Table 4.4; full sequences in Appendix Twelve). When aligned with the published *Mhc-Mafa-DRB* sequences, only one positive match was found, corresponding to the allele *Mhc-Mafa-DRB4*0110*. Predicted nucleotide and amino acid sequences are presented in Appendix Thirteen and Figure 4.4 respectively. Phylogenetic analysis of the LKWS *Mafa-DRB* sequences was performed to identify potential locus and lineage of each sequence. Neighbour-joining trees included the new sequences and the published *Mafa-DRB* sequences plus other selected *Macaca* spp. and old world monkey sequences. Figure 4.5 includes the 47 sequences found in the LKWS long-tailed macaques and some of the published *Mafa-DRB* sequences; an extended tree including all *Mafa-DRB* sequences and the other taxa is presented in Appendix Fourteen. LKWS *Mafa-DRB* representatives of *-DRB1*, *-DRB3*, *-DRB4* and *-DRB*W* loci and lineages were recognised. Thirty sequences clustered together with the lineages *-DRB1*03* (n= 5), *-DRB1*04* (n= 10), *-DRB1*07* (n= 3), *-DRB3*04* (n= 3), *-DRB*W31* (n= 5), *-DRB*W601* (n= 2) and *-DRB*W63* (n= 2). Other single sequences clustered with other *-DRB*W* sequences.

Table 4.4. Summary of LKWS *Mafa-DRB* sequences and typing techniques to identify each gene

Name	Technique	Name	Technique	Name	Technique	Name	Technique
DRB*Allele1	S/P	DRB*Allele20	S/P	DRB*Allele31	S/P	DRB*Allele42	S
DRB*Allele10	S/P	DRB*Allele21	S/P	DRB*Allele32	S/P	DRB*Allele43	S/P
DRB*Allele11	S/P	DRB*Allele22	S/P	DRB*Allele33	S/P	DRB*Allele44	S/P
DRB*Allele12	S/P	DRB*Allele23	S/P	DRB*Allele34	S	DRB*Allele45	S
DRB*Allele13	S/P	DRB*Allele24	S/P	DRB*Allele35	S/P	DRB*Allele46	S/P
DRB*Allele14	S/P	DRB*Allele25	S/P	DRB*Allele36	S	DRB*Allele5	S/P
DRB*Allele15	S/P	DRB*Allele26	S	DRB*Allele37	S/P	DRB*Allele6	S/P
DRB*Allele16	S/P	DRB*Allele27	S/P	DRB*Allele38	S/P	DRB*Allele7	S/P
DRB*Allele17	S	DRB*Allele28	S/P	DRB*Allele39	S	DRB*Allele8	S
DRB*Allele18	S	DRB*Allele29	S	DRB*Allele4	S/P	DRB*Allele9	S
DRB*Allele19	S	DRB*Allele3	S/P	DRB*Allele40	S/P	DRB4*0110	S
DRB*Allele2	S/P	DRB*Allele30	S/P	DRB*Allele41	S/P		

S- direct sequencing; S/P – direct sequencing plus haplotype reconstruction with PHASE 2.1.1

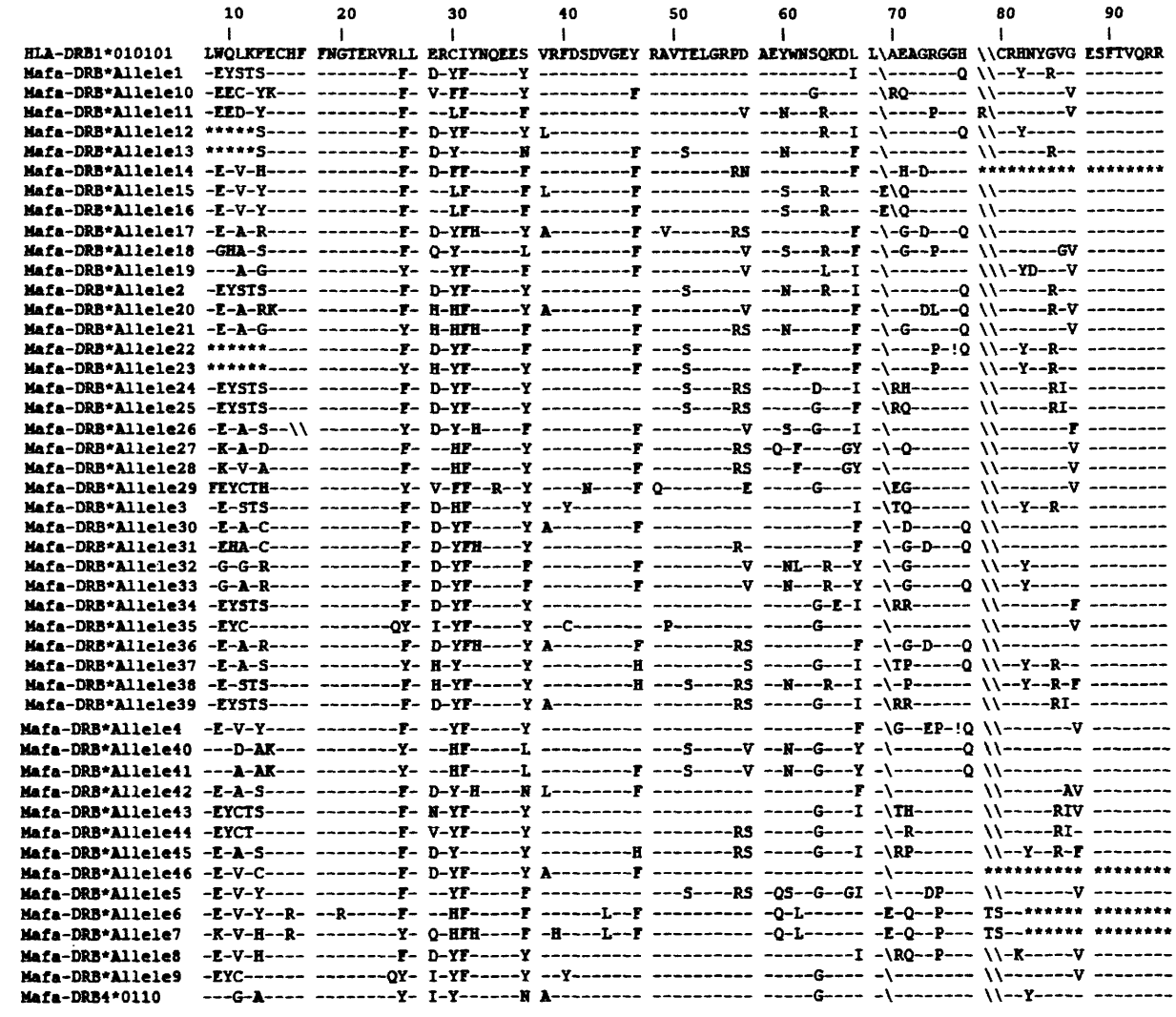


Figure 4.4. Predicted amino acid sequences of long-tailed macaque DRB (exon 2). Numbering of amino acids is shown above the consensus of *–DRB* sequences. Red typeface indicates inferred ABS sites. Identity to consensus is shown by dashes (-), indels are indicated by inverse slash (\), a stop codon is indicated by an exclamation mark (!), and unsequenced regions are denoted by asterisks (*).

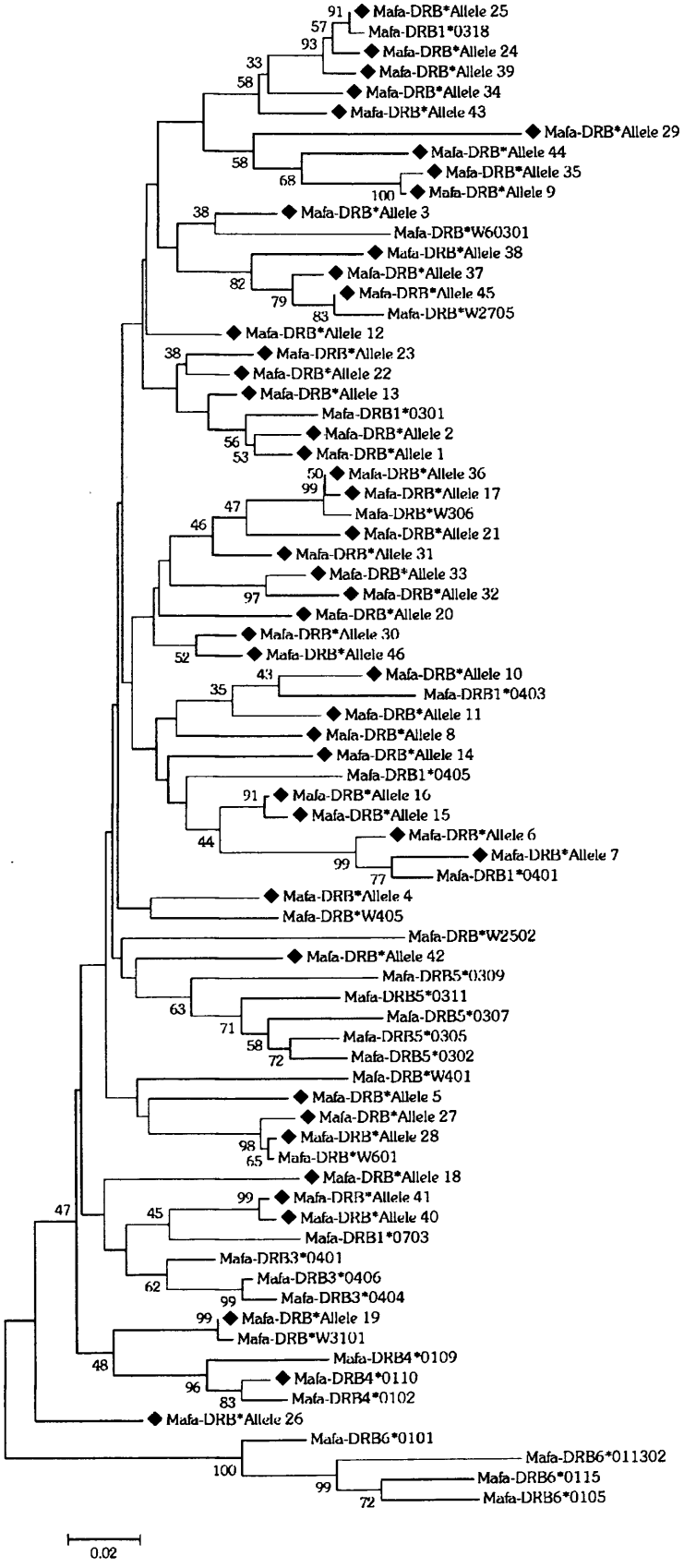


Figure 4.5. Phylogenetic tree of 47 LKWS *Mafa-* (◆) and other published *Mafa-DRB* sequences. The tree configuration was derived from 75 nucleotide sequences using the neighbour-joining and Jukes-Cantor methods in MEGA 5.0 Beta. Numbers on the branches refer to confidence probability values.

4.3.3 Patterns of nucleotide substitution

Most nucleotide substitutions were confined to residues representing the antigen binding site (ABS). In the long-tailed macaque, the variable residues (three or more) were at positions 11, 13, 26, 28, 30 and 37 of the β sheet and 57, 61, 67, 70, 71, 73, 74, 77, 78 and 86 of the α helix, while for proboscis monkey the variability was confined to a single position (57) (Fig. 4.6).

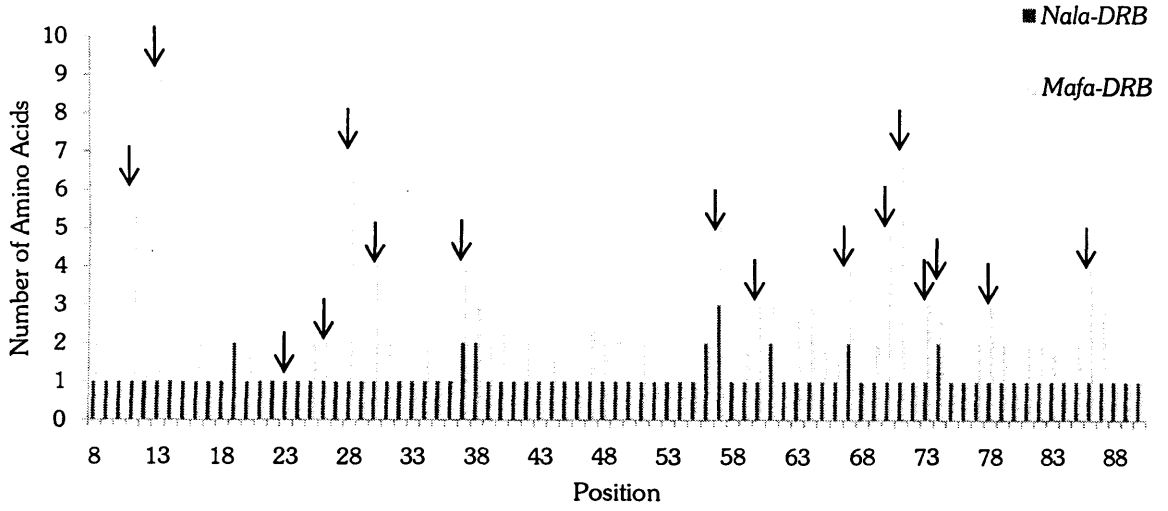


Figure 4.6. Amino acid variability plot for proboscis monkey (*Nala-*) and long-tailed macaque (*Mafa-*) *DRB* alleles. The most variable sites in the rhesus macaque MHC (Slierendregt *et al.*, 1992) are indicated by arrows.

Estimated rates of non-synonymous and synonymous substitutions within the ABS and non-ABS were also examined. In the macaques, the ABS (Table 4.5) contained significantly more non-synonymous changes than synonymous changes (Z test= 2.02, P= 0.02). However, the region outside the ABS contained fewer non-synonymous than synonymous substitutions, and this difference was not significant. For the proboscis monkey, non-synonymous changes either outside or within the ABS were not significantly higher than synonymous substitutions (Table 4.5).

Table 4.5. Estimated rates of non-synonymous (d_N) and synonymous (d_S) substitutions for protein binding (ABS) and non-protein binding (Non-ABS) regions for *DRB* exon 2 of long-tailed macaque (LTM) and proboscis monkey (PM).

Primate	Positions	# of codons	d_N	d_S	d_N/d_S	P
LTM	ABS	17	0.24±0.03	0.14±0.05	1.66	<0.05
	Non-ABS	73	0.21±0.01	0.28±0.04	0.76	ns
	Overall	90	0.21±0.01	0.25±0.03	0.86	ns
PM	ABS	17	0.2±0.03	0.13±0.08	1.5	ns
	Non-ABS	73	0.22±0.01	0.01±0.01	18.3	ns
	Overall	90	0.22±0.01	0.03±0.01	6.6	ns

P – significance level, ns-not significant

4.4 Discussion

In this preliminary study, the genetic variability of the functionally important MHC gene *-DRB* (exon 2) was investigated in an endemic primate (the proboscis monkey) and in a more widely distributed primate (the long-tailed macaque). This is the first report of MHC sequences for proboscis monkeys. Novel *Mhc-DRB* sequences were also detected in long-tailed macaques despite the nature of the sample (faeces) and the small number of typing techniques used. Furthermore, evidence of positive selection was found in the macaque sequences. The procedures employed here could be enhanced and applied to examine future population genetics in these primate species.

As with the study on gorilla MHC by Lukas *et al.* (2004), the current project also relied on non-invasive samples to assess the *Mhc-DRB* variability in proboscis monkeys and long-tailed macaques. As anticipated, the use of DNA templates extracted from faeces made the acquisition of MHC variation challenging, especially for the proboscis monkey. Five *Mhc-Nala-DRB* sequences were obtained from 15 individuals (45% of the tested samples). Failure of PCR reactions could have been due to the presence of plant inhibitors in the DNA extractions or low DNA concentration (see Chapter Three; Broquet *et al.*, 2007). In a study of long-tailed macaque *-DRB* genes, Blancher *et al.* (2006) noted that generic primers are not all equally suitable for every *-DRB* gene so that amplification of some exon 2 genes (or alleles) can be favoured. If such a situation pertains in macaques, it is likely to occur in other taxa (Babik, 2010). For instance, phylogenetic analysis of the *Nala-DRB* samples (Fig. 4.3) suggests that the five detected sequences belong exclusively to the *-DRB5* locus. While it may be possible that the proboscis monkey possesses only this *-DBR* locus, representatives from missing loci/lineages may be uncovered through further study of additional individuals from the LKWS and/or different populations in Borneo. Furthermore, MHC typing techniques such as cloning and sequencing might distinguish more alleles in these samples (Knapp, 2005b). Unfortunately, finite DNA availability did not allow the testing of such methodologies, the use of combining generic *-DRB* primers, or the development of species-specific primers.

The limited number of *-DRB* sequences detected in the proboscis monkeys might account for the apparent low amino acid variation in the positions involved in antigen recognition (ABS). Brown *et al.* (1988, 1993) identified 17 positions as most variable (i.e. four or more amino acid changes) and were assumed to make direct contact with peptides. The variable residues were at positions 11, 13, 23, 26, 28, 30 and 37 of the β sheet and 57, 61, 67, 70, 71, 73, 74, 77, 78 and 86 of the α helix. In the proboscis monkey the highest number of amino acids in any given position was three, and this only occurred in position 57 which forms part of the ABS (Fig. 4.6). According to

Hughes and Nei (1989), non-synonymous substitutions, or changes that alter encoded amino acids within the ABS should be more frequent than synonymous (non-altering) changes in functioning *Mhc-DRB* alleles. Concomitantly, a greater rate of non-synonymous substitution would be indicative of positive selection. In this study, no evidence of positive selection was found for *Nala-DRB* alleles (Table 4.5). Again, the analysis of more proboscis monkey sequences should reveal whether the amino acid polymorphism in the ABS is higher and consequently, whether there is indeed no positive selection acting over the proboscis monkey *Mhc-DRB* loci.

In contrast to the proboscis monkey data, 47 *Mafa-DRB* sequences were obtained from 36 long-tailed macaque samples, representing 80% of the tested samples. These sequences were identified either through direct sequencing or with a combined approach using an algorithm for haplotype reconstruction (PHASE 2.1.1, Stephens *et al.*, 2001; Stephens and Donnelly, 2003). The *Mhc-DRB* variability of the LKWS macaques seems to be slightly higher than that reported in other *Mafa-DRB* studies. For instance in 253 individuals from Mauritius and Indonesia, Blancher *et al.* (2006) reports 50 sequences, 28 previously described. However, those individuals were captive-bred and whether the original animals came from different populations within each country is unclear. Therefore, the variability observed by Blancher *et al.* (2006) could be slightly overrepresented. Doxiadis *et al.*'s (2010), found 118 alleles (28 novel) in 230 captive individuals, of which 162 had a Malaysian/Indonesian origin (two to six generations back), but the origin of the remaining animals was unknown. In a study using samples of over 100 wild Mauritian long-tailed macaques, O'Connor *et al.* (2007) found 34 alleles, 12 being unrecorded. Although the study from O'Connor *et al.* (2007) was the first to evaluate *Mafa-DRB* variability in wild animals, there is no indication that the individuals belonged to a single population. Hence, the current study is likely to be the first to assess *Mhc-DRB* variability in a single population of long-tailed macaques.

The 47 *Mafa-DRB* alleles found in the current study appear to belong to different loci and lineages (Fig 4.5 and Appendix Fourteen). Using the same generalist primers (MDRB5 and 3'DRBSeq), Leuchte *et al.* (2004) found representatives of 17 allelic lineages in the 33 sequences found. Therefore, it was expected to find different loci/lineages in the Kinabatangan samples. However, no evidence of sequences belonging to the *-DRB5* locus was found. As discussed above, the generic primers might not have been suitable to detect this particular locus. Variability in the long-tailed macaque sequences followed the assumptions by Brown (1988, 1993) regarding the amino acid variation in the ABS. These positions are directly comparable to those described for rhesus macaques by Slierendregt *et al.* (1992) (Fig. 4.6). In addition, the analyses of nucleotide substitution support Hughes and Nei's (1989) selection model (Table 4.5). As expected, the *Mhc-DRB* ABS contained significantly more non-synonymous than synonymous

substitutions (Z test, $P< 0.05$). Therefore, positive selection may be a major force in maintaining sequence diversity in LKWS long-tailed macaque *-DRB* genes.

Overall, the results from this preliminary study suggest higher levels of *Mhc-DRB* diversity in long-tailed macaques than in proboscis monkeys. However, different types and number of samples, typing techniques and primer combinations might have rendered different results. For instance, a common feature in previous long-tailed macaque studies is the presence of one to six alleles per individual. In the current study, only one or two *Mhc-DRB* alleles per individual were detected in both the long-tailed macaques and the proboscis monkeys. As stated by Knapp (2005c) identification of MHC alleles in any species is a complicated undertaking, no matter what type of sample is collected, and determination of haplotypes and homozygosity can be extremely problematic when using non-invasive samples. Unlike the current study, none of the previous macaque studies had been performed using non-invasive samples. A common problem when working with non-invasive samples is allelic dropout (ADO) (reviewed by Goossens & Bruford, 2009), which might explain the lack of alleles per individual despite the multi-tube approach (methods).

An alternative explanation for low intra-individual MHC variation is the typing technique used. Primate *Mhc–DRB* alleles are in general expected to be 270 bp long (Doxiadis, per. communication). In this study it was noted that some of the samples analysed presented an amplicon of such size when analysed in agarose gels, but the signal of the electropherograms was corrupted by noise. Such noise was almost as strong as the signal, this strongly suggests the presence of more than two alleles in the same sample. Due to lack of sequence resolution, those samples were not included in the analyses. Future studies should take this into account and for these kind of samples apply a different typing technique. Cloning and sequencing are essentially the 'gold standard' for MHC allele identification in novel species. However, this approach might be impractical when many individuals must be typed (Knapp, 2005b). Conformation-based mutation-detection techniques such as denaturing gradient gel electrophoresis (DGGE) and single-stranded conformation polymorphism (SSCP) have also been traditionally used for MHC typing, but these methodologies require significant amounts of optimization before maximal separation of various alleles can be achieved reliably (Babik, 2010; Knapp, 2005b). This can be a limitation when working with finite samples such as those obtained non-invasively. Recent studies on MHC variation suggest that sequence-based typing may also be applied because it is possible to correctly reconstruct haplotypes from diploid chromatograms containing heterozygous positions using the Bayesian inference method implemented in PHASE (Stephens *et al.*, 2001; Stephens and Donnelly, 2003; reviewed in: Bos *et al.*, 2007, Babik, 2010; Garrick *et al.*, 2010). While not widely used yet, examples of MHC reconstruction by PHASE can be found for

humans, artyodactiles, birds and fish (Lee *et al.*, 2006; Mona *et al.*, 2008; Silva *et al.*, 2009; Turner *et al.*, 2009). In this study, PHASE was useful to clarify base uncertainties, with a mean confidence of 90%. However, the issue of having more than two alleles per sample still needs to be addressed and perhaps a more adequate approach will include cloning or a conformation-based mutation-detection technique.

Although the methodologies presented in this study have to be refined, non-invasive samples could provide suitable material for the assessment of MHC variation in proboscis monkeys and long-tailed macaques. Human impact (e.g. habitat fragmentation, degradation, isolation, urbanisation, pollution) often causes a loss of genetic variation leading to short-term reduction of fitness components and to an impaired ability to adapt to changing environments, which in turn influences evolutionary outcomes (Sommer, 2005; Young & Clarke, 2000). The results from this study represent a first step towards understanding the effects of habitat fragmentation on the fitness of long-tailed macaques and proboscis monkeys of the Lower Kinabatangan Wildlife Sanctuary. As the LKWS is a highly fragmented habitat, the study of MHC diversity will be a useful indicator of the fitness of these two primate species and will ultimately aid their conservation and management.

CHAPTER FIVE

Parasite richness and prevalence in two primate species of the Lower Kinabatangan Wildlife Sanctuary: effects of habitat fragmentation

Salgado-Lynn M[1], **Stanton DWG**[1], **Sakong R**[2], **Goossens B**[1,3], **Bruford MW**[1], **Cable J**[1]

[1]*School of Biosciences, Cardiff University, Museum Avenue, Cardiff CF10 3AX, UK*
[2] *Red Ape Encounters, Kinabatangan Orangutan Conservation Project, PO Box 3109, 90734 Sandakan, Sabah, Malaysia*
[3] *Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia*

5.0 Abstract

A non-invasive survey of the gastrointestinal parasite fauna of free-ranging long-tailed macaques (*Macaca fascicularis*) and proboscis monkeys (*Nasalis larvatus*) in the Lower Kinabatangan Wildlife Sanctuary revealed 14 taxa. These included three trematodes (an unidentified dicrocoeliid, *Clonorchis* sp. and *Fasciola* sp.), two cestodes (*Taenia* sp. and *Dipylidium*-like morph), one unidentified acanthocephalan, seven nematodes (*Strongyloides* sp., *Trichuris* sp., *Anatrichosoma* sp., *Ascaris* sp., strongylids, oxyurids and an oxyurid-like morph) and an unidentifiable parasite. Parasite richness was higher in proboscis monkeys, and prevalence of particular parasites differed between the primates. In particular, *Trichuris* sp. was 2.2 times more prevalent in proboscis monkeys while strongylids were twice as common in long-tailed macaques. Potential effects of natural and anthropomorphic mediated habitat fragmentation on parasite species richness, proportion of individuals with mixed infections and the prevalence of particular parasites between the two primate species were explored. Natural fragmentation (the Kinabatangan River) did not affect parasite distribution. Although it was expected that areas with high rates of human – non-human primate contact would have a positive correlation with parasite diversity, prevalence of particular parasites, or with the proportion of individuals presenting a co-infection, this was not always the case as host-parasite dynamics are likely to be affected by complex interactions between environmental, and host demographic, behavioural and genetic factors. Nonetheless, the results of this study provide a baseline for future work regarding parasite-host ecology in primates of the LKWS.

5.1 Introduction

Parasites represent one of the most successful life forms on the planet and as such are an important component of the biological diversity of tropical forests (Nunn & Altizer, 2006; Trejo-Macías *et al.*, 2007). Research on parasitic fauna can potentially add a new dimension to the

understanding of ecological interactions, host distribution patterns and to the complex history of regions and habitats (Pérez-Ponce De León & García-Prieto, 2001). Parasite infections can critically influence endangered species conservation as they can have major impacts on host abundance and evolution (May, 1988; Nunn *et al.*, 2003; Pedersen *et al.*, 2007; Scott, 1988; Smith *et al.*, 2009). They have been linked increasingly with dramatic local and global declines of wildlife species, including lions, black-footed ferrets, Hawaiian forest birds and many amphibian species (e.g., Daszak *et al.*, 2000; Dobson & Grenfell, 1995; Packer *et al.*, 1999). Therefore, it is essential to obtain accurate data on parasite diversity and abundance at local levels in order to understand the role of infectious agents in wildlife endangerment, declines and extinctions (Smith *et al.*, 2009; Thompson *et al.*, 2010). Furthermore, the study of parasites in wild primate populations provides knowledge for evaluating the health and the infection risk in populations, and identifies general principles governing parasite occurrence which is critical for managing vulnerable wildlife populations and mitigating risks to human health (Chapman *et al.*, 2006a; Gillespie *et al.*, 2005a).

Primates are vulnerable to the effects of many parasitic infections because they often live in close social groups that facilitate parasite transmission (Stoner, 1996). Long term behavioural field observations and studies of endemic parasites in wild primate populations have provided direct and indirect evidence that infectious diseases can cause or contribute to death in primate hosts, such as vervet monkeys, chacma baboons and chimpanzees (for a more detailed list of hosts and infectious agents see Table 1.1 in Nunn & Altizer, 2006). Some of the most striking evidence of infectious disease comes from population declines associated with epidemics in African apes (Leendertz *et al.*, 2004; Leroy *et al.*, 2004; Walsh *et al.*, 2003). In many cases parasites may increase host susceptibility to predation or decrease the competitive fitness of the individual (Scott, 1998), and might be more prevalent in populations living in human modified habitats (Nunn *et al.*, 2003). Whereas the gastrointestinal parasite fauna of wild populations of African apes, baboon, and neotropical howler monkeys is comparatively well studied (reviewed by Gillespie 2006), the parasites in other primate taxa remain poorly known. Based on the IUCN Red List, Southeast Asia harbours a relatively large number of threatened primates (Cowlishaw & Dunbar, 2000). However, the Global Mammal Parasite Database, reveals a major gap in our knowledge of parasites from threatened primates in this area (Nunn & Altizer, 2005, 2006).

Long-tailed macaques (*Macaca fascicularis*) and proboscis monkeys (*Nasalis larvatus*) are two of the ten primate species that live in sympatry in the Lower Kinabatangan Wildlife Sanctuary (LKWS), in Sabah, Malaysia. The long-tailed macaque is listed as one of the 100 most invasive alien species inhabiting a variety of forest habitats, preferring edge habitats and riverine areas, but can also be found in villages (i.e. disturbed habitat), often raiding crops and where they may

be classified as a pest (van Schaik *et al.*, 1996; Abegg & Thierry, 2002). Endemic to Borneo, proboscis monkeys were considered to prefer habitats within riverine areas, peat swamps, mangroves and nipa dominated mangrove forests (Kawabe & Mano, 1972; Meijaard & Nijman, 2000). However, more recent data have shown that proboscis monkeys are more widely distributed, occurring throughout Borneo (Meijaard & Nijman, 2000). Classified as endangered by IUCN (2010), this primate is threatened by habitat destruction and hunting, and much of its former range has been reduced due to logging (e.g. in Kinabatangan), swamp reclamation, gold mining, shrimp farming and forest fires (Meijaard & Nijman, 2000; Sha *et al.*, 2008). Despite the implications for human health (potential zoonoses in the case of the macaques) and the conservation status of the proboscis monkey, little is known of the parasitic fauna of these two species. According to the last update of the Global Mammal Parasite Database (August 2008; www.mammalparasites.org), there was not a single report regarding proboscis monkey parasites, while the reports for *Macaca fascicularis* in the Eurasian region include three groups of bacteria, seven viruses, and nine species of protozoa and helminths (Nunn & Altizer, 2005). However, these numbers are likely to be an underestimate since there is at least one report of a proboscis monkey helminth (Hasegawa *et al.*, 2003) which is not included in the mentioned database but is included in the Natural History Museum Host-Parasite Database (2010; www.nhm.ac.uk/research-curation/resarch/projects/host-parasites/database/). There are no reports on the parasitic fauna of long-tailed macaques and proboscis monkeys of the Kinabatangan region, where the habitat has been subjected to different degrees of disturbance.

Habitat destruction, human population growth and tourism contribute to increased pressure on wildlife (Hahn *et al.*, 2003). Most primates today live in anthropogenically disturbed habitat mosaics of farmland, human settlements, forest fragments and isolated protected areas (Chapman *et al.*, 2005a). Such a landscape lowers the viability of primate populations in tropical forests (Chapman & Peres, 2001; Mittermeier *et al.*, 2009). Habitat loss can also change the behaviour and abundance of wildlife which in turn affects parasite transmission and distribution since they are influenced by host ranging patterns, density, intraspecific and interspecific contact rates, and host diet (Chapman *et al.*, 2005a, b, 2006a, b; Gillespie *et al.*, 2005b; Hudson *et al.*, 2006; Nunn *et al.*, 2003). Furthermore, as human population density continues to increase, speeding the reduction and fragmentation of primate habitats, greater human-primate contact is inevitable and even higher rates of parasite transmission, between primates and between primates and humans, are likely (Gillespie, 2006; Goldberg *et al.*, 2008). For instance, arboreal monkeys in small fragmented habitats may have to come to the ground to move from one forest patch to another with the possibility of acquiring new parasites. Additionally, when moving between forest patches and because of the proximity to human settlements and to domestic

animals, they may also be exposed to a wider range of parasitic vectors and/or intermediate hosts (Trejo-Macías *et al.*, 2007). Although some primate species, such as yellow and olive baboons (*Papio cynocephalus* and *P. anubis*), can persist or even temporarily increase in the face of human encroachment, routine contact with people, their livestock, pets and refuse may introduce new diseases into a previously naive population (Hahn *et al.*, 2003). Additionaly, generalist parasites that can infect multiple host species, including domesticated animals, can be relatively benign in one host species but may depress the density of other hosts for which they are more pathogenic (Altizer *et al.*, 2007, Smith *et al.*, 2009). In general, the effects of fragmentation on the dynamics of pathogen transmission between primates and other species, including humans, remain largely unexplored (Goldberg *et al.*, 2008).

Because of increasing human encroachment into previously pristine forests and the potential for disease transmission between human and non-human primate populations, further detailed investigations of primate ecological parasitology are warranted, especially at sites for which we currently have no information regarding parasite diversity and prevalence (Muehlenbein, 2005). Understanding how forest fragmentation and associated land-use changes affect parasite transmission among primates, humans and domestic animals is critical for designing rational intervention strategies to conserve wild primates, as well safeguarding human and animal health (Goldberg *et al.*, 2008). In addition, considering the evolutionary and ecological linkages between primates and their parasites (Stuart & Strier, 1995), parasites could be viewed as indicator species, potentially alerting us to imminent threats to primate conservation (Gillespie, 2006).

The aim of this study was to non-invasively survey the gastrointestinal parasite fauna of long-tailed macaques and proboscis monkeys in the LKWS. Parasite diversity and prevalence, presence of mixed infections and levels of environmental contamination (measured as parasite eggs per gram of faeces) were compared between the two species of primates. Compared to proboscis monkeys, long-tailed macaques have been observed to come closer to human habitation in the LKWS. Thus, it is expected for the macaques to be infected with a higher parasite diversity and a larger proportion of individuals are likely to carry mixed infections compared to proboscis monkeys. Additionally, effects of habitat fragmentation on parasite species richness, proportion of individuals with mixed infections and the prevalence of particular parasites were explored. Both primate species are known to be efficient swimmers, thus no effect is expected in terms of parasite diversity, prevalence or infection risk arising from the natural barrier of the Kinabatangan River. In contrast, these factors might be influenced by human-mediated habitat fragmentation. It is predicted that areas with higher rates of human contact and

activity will result in primates having higher parasite richness and increased frequency of mixed infections, as well as a higher prevalence of particular parasite species.

5.2 Methods

5.2.1 Study site

Located in eastern Sabah, Malaysia, the Lower Kinabatangan flood plain (5°20' – 5°45' N, 117°40' – 118°30' E) comprises different habitat types and a warm and humid climate. Temperatures range between 21°–34°C and annual precipitation averages 3000 mm (Ancrenaz *et al.*, 2004). Low-stature forests and grasslands occur in backswamp areas while riparian and mixed lowland dipterocarp forests are found in drier zones located along the banks of the rivers and higher terraces: seasonally flooded forest, nipa palms, and mangrove are also common in the area (Azmi, 1998). However, since the mid 1950s (but especially during the past 25 years), most of the dry lowland forest has been subjected to large-scale commercial timber exploitation or has been cleared for oil-palm development, resulting in a highly fragmented forest structure (McMorrow & Talip, 2001).

In 2005, the state government of Sabah gazetted almost 27 000 ha of these forests as a wildlife sanctuary under the Wildlife Conservation Enactment 1997, with the ultimate aim of creating a corridor for wildlife along the Lower Kinabatangan flood plain, between the remaining virgin forest reserves. The Lower Kinabatangan Wildlife Sanctuary comprises 10 sectors or Lots (Lots 1–10, with Lot 10 divided into 10A–C) linking seven patches of protected forests (Virgin Jungle Forest Reserves, VJFR) totalling about 15 000 ha, and they are connected with 10 000 ha of state and private forests in various stages of degradation (Ancrenaz *et al.*, 2004). Three villages (Abai, Sukau and Bilit) and most touristic activities (more than ten tour lodges) are located in Lots 1-4 (hereafter referred to as Area 1). In contrast, Lots 5-7 (Area 2) are quieter in terms of human activities (only one village – Batu Putih, one community based organization – Mescot, and one research centre – Danau Girang Field Centre). In addition, a major bridge crosses the Kinabatangan River, creating another potential barrier for wildlife dispersal, and separates Lots 8-10 (Area 3) from the rest of the sanctuary. This area includes several small human settlements but no tour operators are based here (M. Salgado-Lynn, pers. obs). Thus, human activity in this area is lower than in Area 1, but slightly higher than in Area 2 (Figure 5.1).

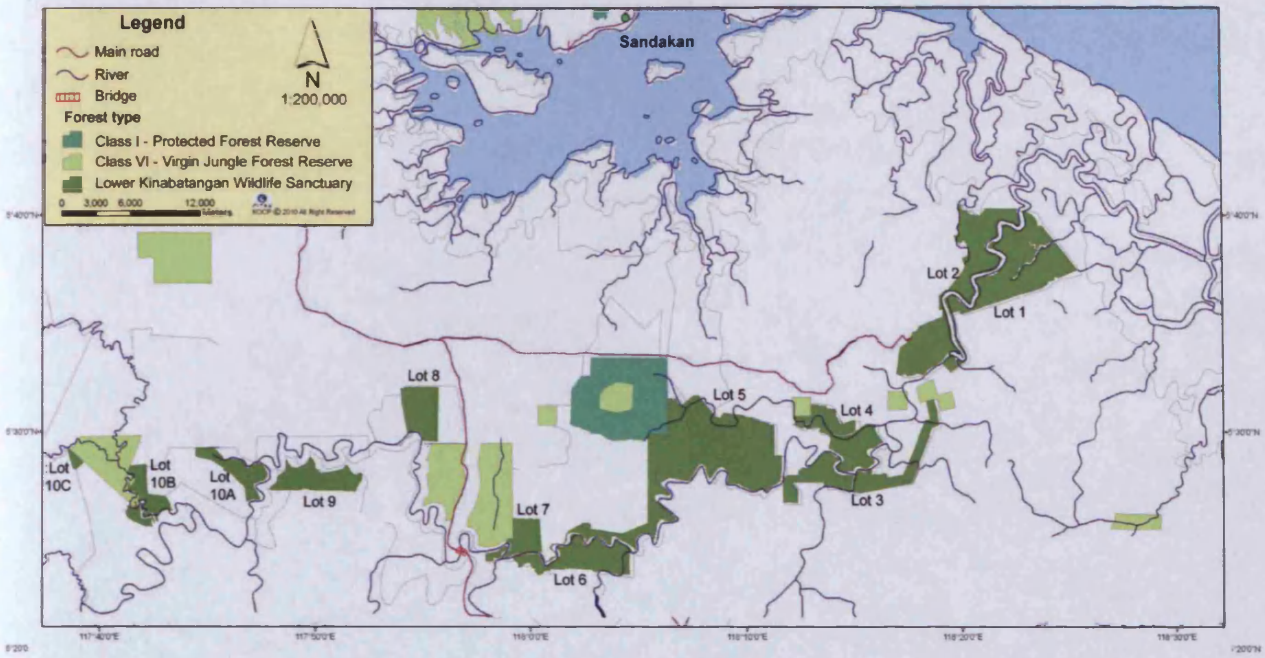


Figure 5.1. Map of the Lower Kinabatangan Wildlife Sanctuary. Courtesy of HUTAN/KOCP.

5.2.2. Sampling

Sampling in LKWS was divided in two seasons; the first covered Lots 1-4 between October 2007 and March 2008, and the second covered Lots 5-10 between June and November 2008. As long-tailed macaques and proboscis monkeys use riverine trees for sleeping sites, the Kinabatangan River was used as a transect to census the primates during their inactive periods (particularly at sunset and early night). Expeditions lasted from 5-10 days, with teams of 2-5 people covering ~330 km, equalling 660 km of riverbank including smaller tributaries. Faeces were collected at dawn to ensure freshness, and occasionally at dusk. GPS (Garmin eTrex Vista HCx) coordinates of collected samples were noted for subsequent localisation of groups and populations. Within the constraints of time and logistics, we attempted to sample as widely as possible within each population seen in the riparian forest; because individual recognition was not always possible, and being unhabituated primates, some individuals may have been sampled more than once.

A total of 1153 faecal samples were collected for molecular and parasitological analysis with those screened for parasites being selected on their size, as many samples were large enough for DNA extraction but not sufficient to be subjected to a flotation technique (see below). Therefore, a total of 290 faecal samples were screened for parasites, 144 for macaques and 146 for proboscis monkeys. These represent 22.1% of the proboscis monkey samples and 29.3% of the

macaque samples collected. Screened samples were grouped into Areas (as described above), based on the characteristics of human activities in the Lots of the Sanctuary (Table 5.1).

Table 5.17. Summary of primate stool samples screened for parasites.

		Area 1				Area 2			Area 3				
Primate		Lot1	Lot2	Lot3	Lot4	Lot5	Lot6	Lot7	Lot8	Lot9	Lot10	OS	Total
PM	C	65	87	199	89	72	31	85	0	0	33	0	661
	S	17	23	26	16	17	11	27	0	0	9	0	146
LTM	C	27	59	80	33	54	83	39	32	46	33	6	492
	S	4	19	17	6	11	27	14	12	17	14	3	144

PM-proboscis monkey, LTM- long-tailed macaque, C-number of collected samples, S-number of screened samples, OS- samples from outside the sanctuary on the south of the Kinabatangan river.

5.2.3. Parasite isolation

Faecal samples were weighed and stored in 15 ml falcon tubes with 4 ml of 0.15% sodium azide (storing solution) according to Harmon *et al.* (2007a). All samples were processed within 15 days, with the majority (70%) processed the same day of collection. Helminth eggs were isolated using a modified McMaster-Wisconsin Flotation technique (Dryden *et al.*, 2005; Egwang & Slocombe, 1982; FAO, 2007). Faecal material was stirred with 15 ml of storing solution until broken down, and subsequently filtered through a small strainer (mesh size 3 mm). The faecal suspension was homogenized and centrifuged at 400 G for 2 min. The supernatant was discarded and the pellet was resuspended in 15 ml of distilled water to be washed again under the same conditions but then filtered through a finer mesh (0.5 mm). The pellet was gently resuspended in 2 ml of flotation fluid (saturated sucrose solution, specific gravity 1.26), and 13 ml of flotation fluid were gradually added for a total of 15 ml. The tube was inverted slowly to thoroughly mix the sample, avoiding the formation of bubbles. One ml of the mix was withdrawn and used to quantify eggs using a McMaster chamber. Parasites categorised according to morphology were counted and the results were multiplied by conversion factors (depending on the original weight of each sample, Appendix one) to give the number of eggs per gram (EPG) of faeces. Flotation fluid was added to the remaining sample to form a positive meniscus in the tube, which was afterwards covered with a 22x22 mm coverslip and left to stand for 1 h. The coverslip was carefully removed and placed on a slide, which was screened using a compound microscope (Olympus BX51 or PRIOR PX033) at 100, 200 and 400x to identify particular morphological features of parasite eggs/larvae. Each slide was systematically analysed by screening the entire area under the

coverslip in progressive vertical motions, with all objects being recorded as encountered. All morphotypes were photographed for subsequent identification using an Eye-Piece Microscope EM-300M BigCatch™ camera and the software ScopePhoto 3.0 (ScopeTek). Parasite images were compared with electronic keys (DPDx; Fox; Sullivan, 1999) and/or images from Ash & Orihel (2007). The size of parasite eggs and larvae were measured using the softwares analySISGetIT 5.0 and measureIT 5.1 (Olympus Soft Imaging Solutions GmbH). No coproculture was used to match parasite eggs to larvae for positive identification, hence morphological classification was only possible to genus level. Parasite identification was independently confirmed by Prof. Michael Muehlenbein (Dept. Anthropology, Indiana University).

5.2.4. Data analysis

For this study, parasite richness was defined as the number of unique helminth species (or morphotypes) documented from individual hosts' faecal samples (Muehlenbein, 2005). Prevalence was measured as the percent of host individuals infected with a particular parasite (Margolis *et al.*, 1982; Bush *et al.*, 1997). Exact ("Clopper-Pearson") binomial confidence intervals (95%) were computed using an online web-based calculator (Pezzullo's "Interactive Statistical Calculations"). Parasite egg production is often highly variable and it may not be indicative of actual infection intensity, yet it is useful as an indication of environmental contamination. Environmental contamination (EC) is a correlate of infection risk or potential for "spill-over" to individuals of the same and different species (Chapman *et al.*, 2005b; Gillespie, 2006; Greiner & McIntosh, 2009). This parameter (EC) was estimated as the average number of EPG (eggs per gram) of faeces in infected individuals.

For the whole LKWS, prevalence of each parasite type was compared between the two species of primates by a Pearson's χ^2 test with Yates continuity correction (R package version 2.10.1). The same test was used to compare the proportion of mixed infections between primate species. Parasite species richness was compared between the two types of monkeys by a Mann-Whitney U test (Minitab® 15.1.30.0). Environmental contamination was compared between primates only for those parasite species whose prevalence was $\geq 25\%$ (lower prevalences imply reduced sample size, rendering statistical comparisons invalid), this comparison was also performed by a Mann-Whitney U test.

To assess whether parasite richness, prevalence and environmental contamination area were affected by host or locality, these associations were examined using generalised linear models (GLMs) (R package version 2.10.1). All models were initially run as a mixed model, using the

package ASReml v3 via the R statistical package interface with primate social group and faecal sample size (in grams) fitted as a spline in the random term. The random model was reduced after assessment of the log-likelihood, followed by minimization of the fixed model using the Wald statistic. Random terms were found to be of no significance, so all models were then reduced to a GLM and a minimal model was reached using stepwise deletions. Details of all terms and their interactions, incorporated in the initial models are described in Table 5.2. In particular, two localities were tested, A) river side (north/south of the Kinabatangan River) and B) area in the sanctuary (Areas 1-3, as described above); a GLM was used to assess each locality (Model A and Model B respectively). Again, prevalence and environmental contamination were only examined for parasites whose prevalence was above 25%.

Table 5.18. Terms initially included in the diversity and distribution models.

Dependent variable	Independent continuous (c) / categorical (v) variables
Richness	Primate species (c) Locality[a] (c) Size[b] of faecal sample (v) Primate species X locality Primate species X sample size Locality X sample size
Mixed infections	Primate species (c) Locality (c) Size of faecal sample (v) Primate species X locality Primate species X sample size Locality X sample size
Prevalence of *Ascaris* sp./ *Trichuris* sp./ strongylid	Primate species (c) Locality (c) Richness (c) Size of faecal sample (v) Primate species X locality Primate species X sample size Primate species X richness Locality X sample size Locality X richness Richness X sample size
Environmental contamination of *Ascaris* sp./*Trichuris* sp./ strongylid	Primate species (c) Locality (c) Richness (c) Size of faecal sample (v) Primate species X locality Primate species X sample size Primate species X richness Locality X sample size Locality X richness Richness X sample size

[a] – two different models for each locality were run, Model A) north and south of Kinabatangan River, Model B) 3 areas in the LKWS.
[b] – in grams.

Depending on the data, GLMs were run with different error distributions and link functions. Initial richness models, with a Gamma error distribution and Log link, revealed normal residuals (Appendix Fifteen). Prevalence and mixed infection models used a binomial error distribution and a logit link function. Histograms of environmental contamination showed highly skewed distributions, so transformations (log, log10, square root, exponential) were applied directly to the data without success. GLMs were then tested with several error structures (Negative binomial, Poisson, Quasi Poisson) and link functions (log, $\log_{10}$, square root), but models still revealed non-normal residuals (some examples in Appendix Fifteen). Hence, the associations between environmental contamination and the variables described in Table 5.2 were not explored further.

5.3 Results

5.3.1. Gastro-intestinal helminth parasite fauna in long-tailed macaques and proboscis monkeys of the Lower Kinabatangan Wildlife Sanctuary

Fourteen helminth parasite morphs were recovered from 144 long-tailed macaque samples and 146 proboscis monkey samples. Morphs detected from the samples include: three trematodes (dicrocoeliid, *Clonorchis* sp. and *Fasciola* sp.), two cestodes (*Taenia* sp. and *Dipylidium*-like morph), one unidentified acanthocephalan, seven nematodes (*Strongyloides* sp., *Trichuris* sp., *Anatrichosoma* sp., *Ascaris* sp., strongylids, oxyurids, and an oxyurid-like morph), and a thus-far unidentified parasite (Unknown Morph). Descriptions, images and faecal egg counts (FEC) for each parasitic type are presented in Appendix Sixteen. For the whole LKWS mean parasite richness was higher in proboscis monkeys than in long-tailed macaques (Mann-Whitney, W= 19706, d.f.= 288, P= 0.03) (Figure 5.2). Ten parasite morphotypes were common to both primate species (dicrocoeliid, *Dipylidium*-like, *Taenia* sp., acanthocephalan, oxyurid, strongylid, *Strongyloides* sp., and *Trichuris* sp.), the remainder were found only in the long-tailed macaque samples.

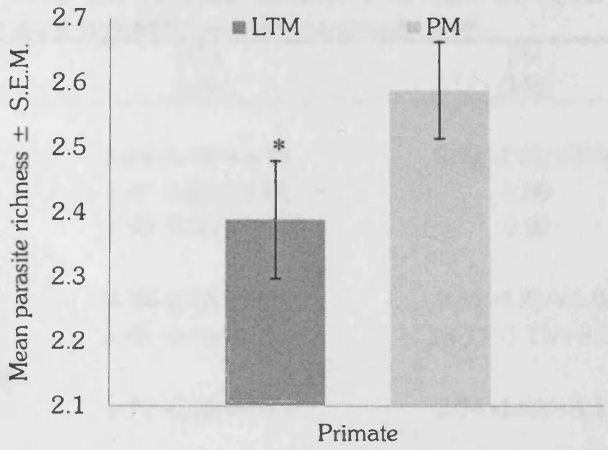


Figure 5.8. Parasite richness of long-tailed macaques (LTM) and proboscis monkeys (PM) of the LKWS.
* indicates a difference at $P< 0.05$ (Mann-Whitney U Test)

There was no significant difference in overall parasite prevalence between long-tailed macaques (95.1 -4.9/+2.9%) and proboscis monkeys (96.6 -4.39/+2.3%). The most prevalent parasites (>10%) were *Taenia* sp., *Trichuris* sp., *Ascaris* sp., strongyles and *Strongyloides* sp.; the first two were significantly more common in proboscis monkeys (χ^2= 34.8 and 67.7 respectively, d.f.= 1, $P< 0.001$) and the last two occurred more frequently in long-tailed macaques (χ^2= 56 and 8.1 respectively, d.f= 1, $P< 0.01$) (Table 5.3). For the whole LKWS, 86.1% (-6.94/+5.3%) of long-tailed macaques and 89.7% (-6.11/+4.4%) of proboscis monkeys were infected with multiple parasite species, but this was not significantly different (χ^2= 3.0037, df= 1, $P> 0.05$). On average, 2.4 different parasites were found co-infecting each primate species (range of infection 1 to 5 different parasite types). Environmental contamination (mean EPG) differed between the two primate species in a fashion similar to that of prevalence (Fig. 5.3). The mean EPG of *Ascaris* sp. (LTM= 46, PM= 100), *Taenia* sp. (LTM= 5, PM= 174) and *Trichuris* sp. (LTM= 79, PM= 690) were significantly higher in the proboscis monkey than in the long-tailed macaque (W= 17211, 16593 and 20678 respectively, d.f.= 247, $P <0.001$), while strongylids (LTM= 290, PM= 135) had a higher mean EPG in long-tailed macaques (W= 11078, d.f.= 247, $P <0.001$).

Table 5.3. Prevalence ± 0.95 confidence limits (%) of gastrointestinal parasite infections in long-tailed macaques (LTM) and proboscis monkeys (PM) from the Lower Kinabatangan Wildlife Sanctuary (sample size is in parentheses following species name).

Parasite type	LTM (144)	PM (146)	Significance
Digneans			
Dicrocoeliid	2.80 -2.04/+4.10	2.05 -1.62/+3.84	NS
Clonorchis sp.	1.40 -1.23/+3.53	0.00	-
Fasciola sp.	4.90 -2.92/+4.86	0.00	-
Cestodes			
Dipylidium-like	4.90 -2.92/+4.86	9.59 -4.25/+5.98	NS
Taenia sp.	2.80 -2.04/+4.10	28.77 -7.19/+8.06	***
Acanthocephalan			
Unknown sp.	0.70 -0.68/+3.11	2.74 -1.99/+4.13	NS
Nematodes			
Anatrichosoma sp.	2.10 -1.67/+3.87	2.05 -1.62/+3.84	NS
Ascaris sp.	62.94 -8.89/+7.48	67.12 -8.25/+7.54	NS
Oxyurid	2.10 -1.67/+3.87	4.11 -2.59/+4.62	NS
Oxyurid-like	2.80 -2.04/+4.10	0.00	-
Strongylid	88.11 -7.14/+4.31	45.21 -8.25/+8.44	***
Strongyloides sp.	16.78 -5.80/+7.00	5.48 -3.09/+5.03	**
Trichuris sp.	53.85 -15.67/+1.17	91.78 -5.70/+3.90	***
Unknown Morph	1.40 -1.23/+3.53	0.00	-
Overall	95.14 – 4.90/+2.88	96.58 -4.39/+2.30	NS

*P< 0·05, **P< 0·01, ***P< 0·001, NS> 0.05; – no chi-square test performed as one primate species had zero prevalence.

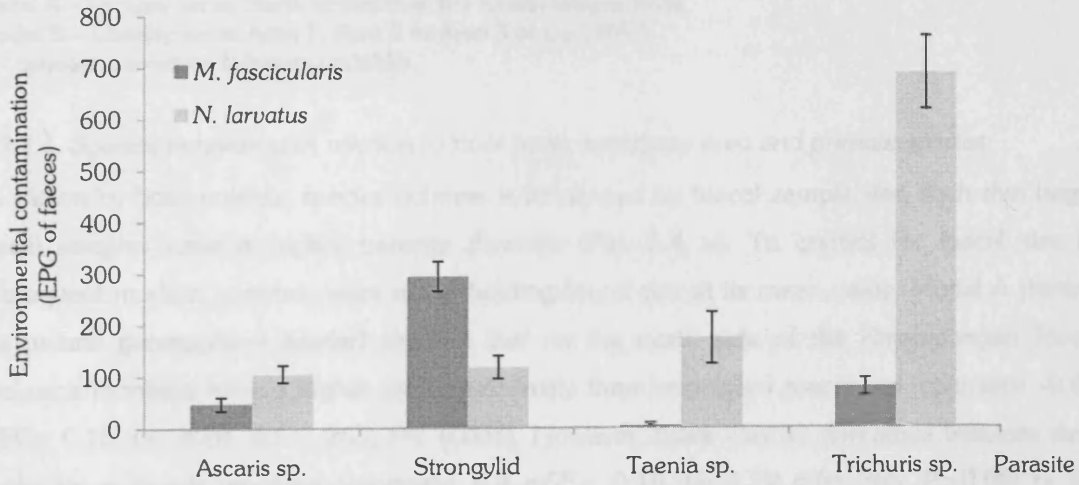


Figure 5.9. Environmental contamination with parasites found in long-tailed macaques and proboscis monkeys from the LKWS. All parasite intensities were significantly different (Mann-Whitney U Test, P< 0.001) between primate species (see text for details).

5.3.2. Parasite diversity and distribution models

The minimal model for each dependent variable is presented in Table 5.4.

Table 5.4. Minimal GLMs revealing categorical and continuous variables which are predicted to be significantly ($P<0.05$) associated with parasite richness, presence of mixed infections and parasite prevalence.

1) Parasite richness Model A			
Model term	F-statistic	d.f.	*P*
Dung size	23.7	1 , 265	<0.001
Primate species X locality	5.54	1 , 262	0.019

2) Parasite richness Model B			
Model term	F-statistic	d.f.	*P*
Dung size	8.49	1 , 263	<0.01
Primate species X locality	4.44	2 , 263	0.012

3) Mixed infections Model A			
Model term	χ^2	d.f.	*P*
Primate species X dung size	4.18	1	<0.05

4) Mixed infections Model B			
Model term	χ^2	d.f.	*P*
Dung size	11.82	1	<0.001
Primate species	11.12	1	<0.001
Locality	9.7	2	<0.001

5) *Ascaris* sp. prevalence Model A			
Model term	χ^2		
Primate species X richness	16.188	1	<0.001

6) *Ascaris* sp. prevalence Model B			
Model term	χ^2	d.f	*P*
Locality	6.19	2	<0.05
Primate species X richness	11.25	1	<0.001

7) Strongylid prevalence Model A			
Model term	χ^2	d.f	*P*
Richness	43.61	1	<0.001
Primate species X dung size	9.5	1	<0.01

8) Strongylid prevalence Model B			
Model term	χ^2	d.f	*P*
Primate species X locality	25.95	2	<0.001
Richness	23.32	1	<0.001
Dung size	6.9	1	<0.01

9) *Trichuris* sp. prevalence Model A			
Model term	χ^2	d.f	*P*
Primate species X richness	6.43	1	0.011

10) *Trichuris* sp. Prevalence Model B			
Model term	χ^2	d.f	*P*
Primate species X locality	14.2	2	<0.001
Richness	88.6	1	<0.001

Model A – Locality set as North or South of the Kinabatangan River.
Model B – Locality set as Area 1, Area 2 or Area 3 of the LKWS.
X - denotes interaction between variables.

5.3.2.1. Species richness with relation to river bank, sanctuary area and primate species

As shown by both models, species richness is influenced by faecal sample size, such that larger faecal samples have a higher parasite diversity (Fig. 5.4 a). To control for faecal size in subsequent models, contrasts were made holding faecal size at its mean value. Model A (testing the natural geographical barrier) showed that on the north side of the Kinabatangan River, proboscis monkeys have a higher parasite diversity than long-tailed macaques (contrast= -0.63 ±SE= 0.16, t= -4.01, d.f.= 265, $P< 0.001$). However, there was no difference between river banks for proboscis monkeys (contrast= 0.3 ±SE= 0.18, t= 1.72 d.f= 265, $P>0.05$) or for macaques (contrast=-0.26 ±SE= 0.16, t= -1.60, d.f.= 265, $P> 0.5$; Fig. 5.4 b). Model B (testing anthropomorphically created barriers) showed different partitioning regarding parasite diversity. Proboscis monkeys had a higher parasite diversity than the macaques in Areas 2 (contrast= -0.66 ±SE= 0.18, t= -3.64, d.f.= 263, $P< 0.001$) and 3 (contrast= -0.92 ±SE= 0.34 ,t= -2.67, d.f.= 263, $P< 0.01$). Additionally, proboscis monkeys in Area 1 had significantly

lower parasite diversity than those in Areas 2 (contrast= -0.72 ±SE= 0.18, t= -3.91, d.f.= 263, $P< 0.001$) and 3 (contrast= -0.85 ±SE= 0.37, t= -2.31, d.f.= 263, $P< 0.05$) (Fig. 5.4 c).

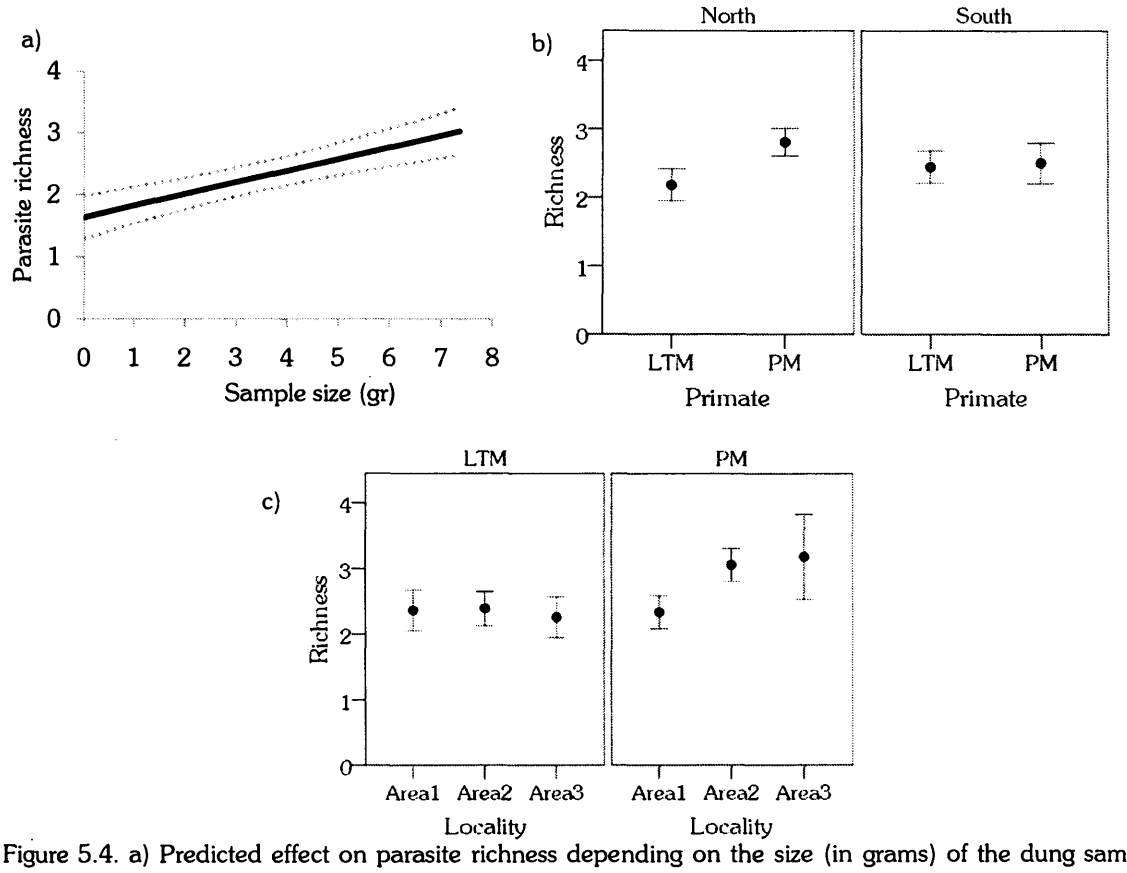


Figure 5.4. a) Predicted effect on parasite richness depending on the size (in grams) of the dung sample analysed. Dotted lines represent 95% confidence intervals. b) Predicted effect on parasite richness depending on the side of the river bank and primate species; 95% confidence intervals are shown. c) Predicted effect on parasite richness depending on sanctuary area and primate species where found; 95% confidence intervals are shown as solid bars.

5.3.2.2. Presence of mixed infections with relation to river bank, sanctuary area and primate species

As in the previous models, the size of the faecal sample effected the detection of mixed infections. From Model A, in particular, an interaction between the size of the sample and the primate species is positively associated with the presence of mixed infections; the reason for the interaction is that the trend of that relationship is subtly different between the two primate species (Fig. 5.5 a). However, the presence of mixed infections was not influenced by the Kinabatangan River. When the data was compared in terms of areas (Model B), mixed infections in each area were more common in proboscis monkeys than in long-tailed macaques. They were also more common in Area 2 for each primate (Fig. 5.5 b).

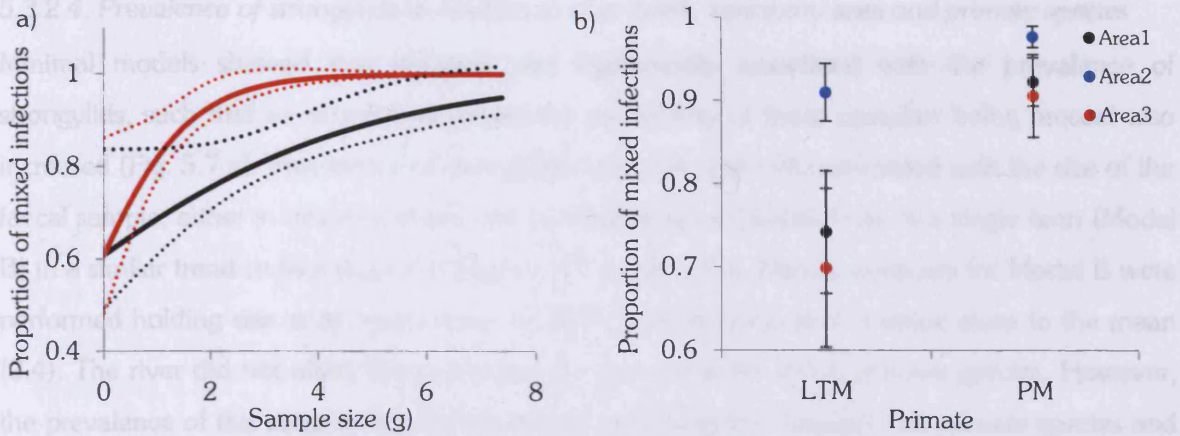


Figure 5.5. a) Predicted effect of size of the stool sample over the proportion of mixed infections detected. Proboscis monkeys are represented by red lines and long-tailed macaques by black lines; dotted lines represent 95% confidence intervals. b) Predicted effect of primate species and area of the sanctuary on the proportion of mixed infections detected. 95% confidence intervals are represented by solid bars.

5.3.2.3. Prevalence of Ascaris *sp. in relation to river bank, sanctuary area and primate species*

An interaction between primate species and parasite richness was found in both models ((A) comparing river sides or (B) comparing areas in the sanctuary) to have a significant effect on the prevalence of *Ascaris* sp. Similar to the effect of dung size on the mixed infection models, the interaction detected in the *Ascaris* sp. prevalence models reflects a subtly different trend in the association between prevalence and richness between the two primate species (to exemplify, Area 1 is depicted in Fig. 5.6 a). Model A did not show an effect of the Kinabatangan River on the prevalence of *Ascaris* sp. On the other hand, Model B indicates a higher prevalence of this parasite in Area 2 than in the other areas of the sanctuary for long-tailed macaques. The prevalence is also higher in Area 2 than in Area 1 for proboscis monkeys (Fig. 5.6 b).

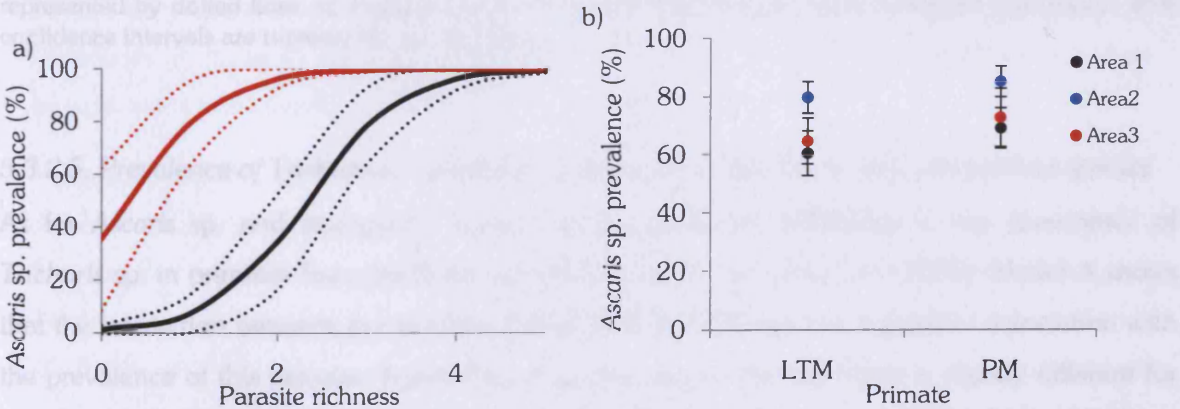


Figure 5.6. a) Predicted effect of parasite richness on *Ascaris* sp. prevalence. The red lines represent proboscis monkeys and the black lines represent long-tailed macaques. 95% confidence intervals are represented by dotted lines. b) Predicted effect of area of the sanctuary over *Ascaris* sp. prevalence. 95% confidence intervals are represented by solid bars.

5.3.2.4. Prevalence of strongylids in relation to river bank, sanctuary area and primate species

Minimal models showed that richness was significantly associated with the prevalence of strongylids, such that as richness increased the probability of these parasites being present also increased (Fig. 5.7 a). Prevalence of strongylids was also positively associated with the size of the faecal sample, either in interaction with the primate species (Model A) or as a single term (Model B) in a similar trend to that shown in Figures 5.4 a and 5.5 a. Hence, contrasts for Model B were performed holding size at its mean value (2.887), and Richness at 2, a value close to the mean (2.4). The river did not affect the prevalence of strongylids for either primate species. However, the prevalence of this parasite was influenced by an interaction between the primate species and the area of the sanctuary (Model B). In Area 1, the prevalence of hookworms was higher in long-tailed macaques than in proboscis monkeys (contrast= 5.18 ±SE= 0.87, t= 5.94, d.f.= 262, $P< 0.001$). At the same time, for proboscis monkeys, the prevalence of these parasites was higher in Area 2 than in Area 1 (contrast= -4.43 ±SE= 0.78, t= -5.65, d.f.= 262 $P< 0.001$).

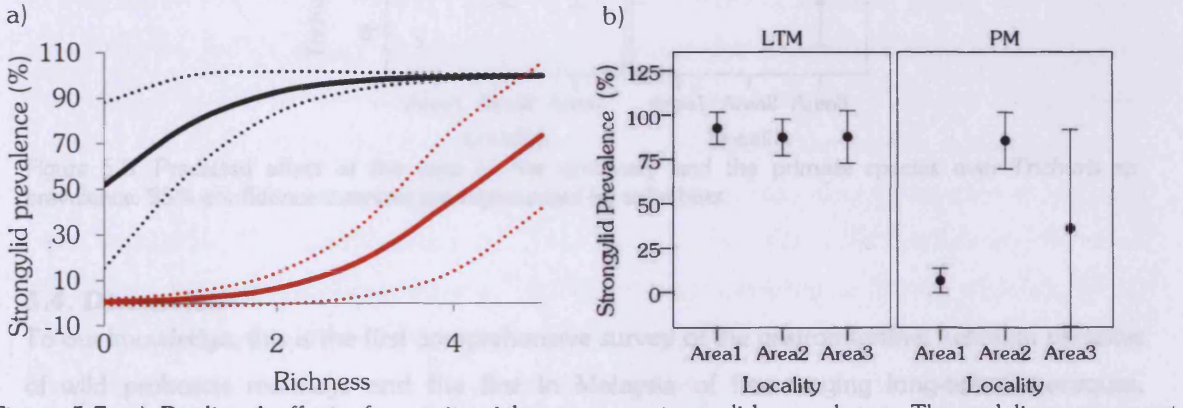


Figure 5.7. a) Predicted effect of parasite richness over strongylid prevalence. The red lines represent proboscis monkeys and the black lines represent long-tailed macaques. 95% confidence intervals are represented by dotted lines. b) Predicted effect of area of the sanctuary over strongylid prevalence. 95% confidence intervals are represented by solid bars.

5.3.2.5. Prevalence of Trichuris *sp. in relation to river bank, sanctuary area and primate species*

As for *Ascaris* sp. and strongylids, there was no significant difference in the prevalence of *Trichuris* sp. in primates from the North and South sides of the river ($P> 0.05$). Model A shows that the interaction between the primate species and the richness has a positive association with the prevalence of this parasite. Again the interaction shows that the trend is slightly different for the two monkey species in a manner similar to that of the *Ascaris* sp. and strongylid prevalence (Figs 5.6 and 5.7 a). This was also shown by Model B where as richness increased then the probability of *Trichuris* sp. being present in the sample also increased. This model also indicated that the prevalence of the whipworms was influenced by the primate species and area of the sanctuary interaction. Contrasts between these two variables (holding richness at 2 and size at its

mean) indicated that the prevalence of *Trichuris* sp. was higher in proboscis monkeys than in long-tailed macaques in Area 1 (contrast= -3.9 ±SE= 0.7, t= -5.57, d.f.= 283, $P< 0.001$), and in Area 2 (contrast= -2.49 ±SE= 0.71, t= -3.48, d.f.= 283, $P< 0.001$). At the same time, a gradient in the prevalence of this parasite was detected in proboscis monkeys, where prevalence was higher in Area 1 than in Area 2 (contrast= 1.98 ±SE= 0.87, t= 2.27, d.f.= 283, $P< 0.05$) or Area 3 (contrast= 5.19 ±SE= 1.13, t= 4.60, d.f.= 283, $P< 0.001$), and also was significantly higher in Area 2 compared with Area 3 (contrast= 3.21 ±SE= 1.06, t= 3.02, d.f.= 283, $P= 0.002$) (Fig. 5.8).

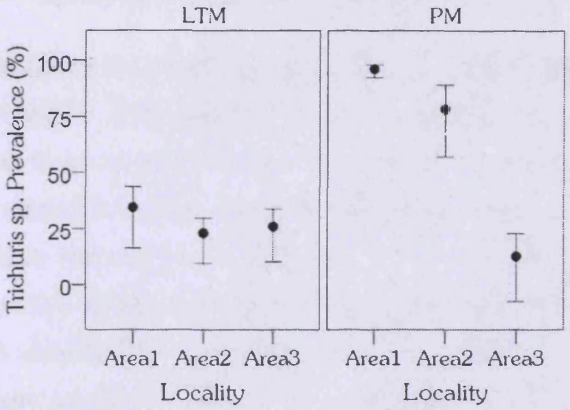


Figure 5.8. Predicted effect of the area of the sanctuary and the primate species over *Trichuris* sp. prevalence. 95% confidence intervals are represented by solid bars.

5.4. Discussion

To our knowledge, this is the first comprehensive survey of the gastrointestinal helminth parasites of wild proboscis monkeys and the first in Malaysia of free-ranging long-tailed macaques. Compared to other studies, parasite diversity was high for both primate species, and more than 50% of the parasite taxa were common to both hosts. Additionally, the results reveal significant differences (in terms of diversity, prevalence and incidence of co-infections) between the two primate species and between different areas of the LKWS.

Fourteen helminths were detected in the long-tailed macaques of the LKWS (Table 5.3, Appendix Sixteen). In a study on semi-wild Vietnamese cynomolgus macaques, Son (2002) reports five helminths which included *Trichuris trichura*, *Strongyloides fulleborni*, *Trychostrongylus* sp., *Ancylostoma duodenale* and *Oesophagostomum* sp.. Between three to four of these taxa (*Trichuris* sp., *Stronglyloides* sp., strongylidae - *Ancylostoma* sp. and/or *Oesophagostomum* sp.) were also found in the current study, but whilst the most prevalent parasite in Son's (2002) study was *Trichuris trichura*, in the Kinabatangan macaques hookworms were the most prevalent. Another study in natural populations of long-tailed macaques in Mauritius recovered only *Strongyloides* sp. and *Trichuris* sp. (see Matsubayashi *et al.* 1992).

Studies on other wild macaque species also reported only a few parasite species. For instance, Ekanayake *et al.* (2006), Gotho (2000), Gotho *et al.* (2001), Horii *et al.* (1982) and Hernandez *et al.* (2009), report 4–6 helminth parasite species in toque (*M. sinica*), Japanese (*M. fuscata*) and Sulawesi (*M. hecki/tonkeana*) macaques. Together, these studies report ten parasite taxa. Of these, *Anatrichosoma* sp., *Enterobious* sp. (Oxyuridae), *Oesophagostomum* sp. (Strongylidae), *Strongyloides* sp., and *Trichuris* sp. were are also found in LKWS *M. fascicularis*. Although the number of parasite species found in the macaques of the LKWS seems higher than those reported in other wild macaque populations, these comparisons should be treated with caution because of differences in the parasite isolation technique and the number of samples analyzed.

Ten helminths were detected in the proboscis monkeys of the LKWS, all of them shared with the macaques of the same region. This parasite diversity is similar that reported for wild African colobines, but lower than those reported for free-ranging Asian colobines. Gillespie *et al.* (2005a) detected between 3-9 helminths in three colobines (*Piliocolobus tephrosceles*, *Colobus guereza*, *C. angolensis*) from Kibale National Park, in Uganda. Mbora & Munene (2006) also reported nine helminths in *Procolobus rufomitratus* from the forest around the Tana River in Kenya, while Teichroeb *et al.* (2009) identified six helminths in *Colobus vellerosus* in Ghana. Seven of the parasites reported in these previous studies were also detected in the current study (*Ascaris* sp., *Colobenterobius/Enterobius* sp. (Oxyuridae), *Dicrocoeliidae* sp., *Oesophagostomum* sp. and other strongylids, *Strongyloides* sp., and *Trichuris* sp.). In Asian colobines, Rajendran *et al.* (2004) reported four helminth species in Nilgiri langurs (*Semnopithecus johnii*), while Do (2009) reported only one for wild Delacour's langurs *Trachypithecus delacouri*. Of all the parasites reports in Asian colobines, *Trichuris* sp., *Strongyloides* sp. and strongylids were also present in the current study.

Despite sharing 10 parasite taxa, there were still significant differences in parasite richness, prevalence of some parasites taxa, and levels of environmental contamination between long-tailed macaques and proboscis monkeys across the LKWS. The total parasite species richness was higher in the long-tailed macaques than in proboscis monkeys (14 compared to 10 taxa, respectively). This supports previous observations that total parasite species richness is lower in threatened compared to non-threatened primates (Altizer *et al.*, 2007). However, individual mean species richness was higher in proboscis monkeys than in long-tailed macaques (Fig. 5.2). In a comparative study of parasite species richness in primates, Nunn *et al.* (2003) report body mass as one of the strongest predictors of helminth species richness. With proboscis monkey males weighing on average 20 kg (Bennet & Gombek, 1993) and macaque adult males averaging 5.4 kg (Thierry, 2007) this host trait might explain the current difference in parasite richness between host species.

The most prevalent gastro-intestinal parasites (>25%) in proboscis monkeys and long-tailed macaques were *Ascaris* sp., strongyles and *Trichuris* sp. (Table 5.3). The latter two parasite groups have also been the most prevalent in previous studies (Ekanayake *et al.*, 2006; Gotho, 2000; Gotho *et al.*, 2001; Horii *et al.*, 1982; Hernandez *et al.*, 2009; Matsubayashi *et al.*, 1992; Son, 2002). These three parasites have a direct life cycle, and are usually referred to as soil-transmitted helminths (Ash & Orihel, 2007; Bethony *et al.*, 2006). Although proboscis monkeys have been reported to seldom travel on the ground (Bennet & Gombek, 1993), during the sampling for the current study it was common to find these primates walking (perhaps foraging), or even resting, on the shores of the Kinabatangan River, especially around Lots 6 and 7. Such behaviour may explain the presence of soil-transmitted helminths in a primate described as typically arboreal. Yet, the difference in prevalence between the two primate species regarding *Trichuris* sp. and strongyles is interesting. Given their transmission mode, their status as generalist parasites and the relatively high levels of environmental contamination found for each (Fig 5.3), no difference would be expected in the prevalence of these parasites between both primate species, as is the case for *Ascaris* sp. (where a difference in infection risk between the two types of primates was observed). Future work could test the parasite avoidance hypothesis (that primates should avoid recently used sleeping sites to reduce the possibility of parasitic infection from contact with accumulated faeces below sleeping sites, Hausfater & Meade, 1982), since a possible explanation for this difference in prevalence might rely on the structure of the parasite eggs, and the tree usage of the monkeys. *Trichuris* eggs can survive for extended periods in the environment compared to hookworm eggs and larvae (Speare *et al.*, 2006), so if the proboscis monkeys do not change their sleeping sites frequently this combination could explain the high prevalence of *Trichuris* sp. in this primate. Another explanation could be 'self-medication' in which primates rely on the selective use of dietary items with antiparasitic properties. For example, in African apes and monkeys condensed tannins from the diet negatively affect the presence of particular parasites (Huffman, 1997; Rothman *et al.*, 2009).

The prevalence of *Taenia* sp., a parasite group typical of pigs and cows, also differed between proboscis monkeys and long-tailed macaques. *Taenia* spp. also have a direct mode of transmission, and hosts become infected by the ingestion of contaminated vegetation (Ash & Orihel, 2007; DPDx). Again, the presence of this parasite implies host movement on the ground, thus it is interesting that proboscis monkeys harbour a higher prevalence (> 25%) than the macaques (< 3%, Table 5.3). Domestic pigs are not common in the LKWS (located in a Muslim country) however the bearded pig, or wild boar, (*Sus barbatus*) is abundant in this area, and it could be the main host for this parasite and the main reservoir for infection in primates.

In terms of human gastro-intestinal infections, the three most prevalent parasites in the current study (*Ascaris* sp., strongylids and *Trichuris* sp.) also occur at high frequency in humans in Malaysia (Singh & Cox-Singh, 2001), and are present in local orang-utans (R. Sakong personal communication). Consequently, zoonotic, epizootic, and/or anthropozoonotic transmission may occur and be promoted by various forms of anthropogenic disturbance. In the LKWS, there is greatest opportunity for contact between monkeys and humans in Area 1 (Lots 1-4), thus it was expected to find more parasite diversity, mixed infections, and higher prevalences in this area compared to the other two, especially for the macaques. This was not the case, as parasite diversity in long-tailed macaques was similar throughout the LKWS, while lower parasite diversity was found in Lots 1-4 than in Lots 5-10 for proboscis monkeys (Fig. 5.4 c). Additionally, a higher proportion of individuals carrying a mixed infection (more than one parasite species) was found in proboscis monkeys than in long-tailed macaques (Fig. 5.5 b), and most of them were found in Area 2 (Fig. 5.5 c), the area with lowest human activity. Area 2 also presented the highest prevalence of *Ascaris* sp. for both primate species (Fig. 5.6 b), and a higher prevalence of stongylids when compared to Area 1 for the proboscis monkeys (Fig. 5.7 b). However, between long-tailed macaques and proboscis monkeys strongylid prevalence was higher in Area 1 for the macaques (Fig. 5.7 b). With regards to *Trichuris* sp., proboscis monkeys presented a higher prevalence than long-tailed macaques in Areas 1 and 2, and for the proboscis monkey alone a clear gradient was apparent in Trichuris sp. prevalence (highest in Area1 and lowest in Area 3; Fig. 5.8).

A single explanation for these results is unlikely as several factors could be involved. Certainly, host population density affects parasite species richness in many systems, with diversity increasing with greater host density (Nunn *et al.*, 2003, but see Chapman *et al*, 2005b and Gillespie & Chapman, 2006). To our knowledge, the densities of proboscis monkeys and long-tailed macaques in the different areas of the LKWS have not been reported to date but, for both hosts, densities in the riparian forest seemed highest in Area 2 (personal observation). Changes in the nutritional status of the host could also influence their ability to resist parasite infections. Studies on baboons (*Papio anubis*) in Kenya and Nigeria (Eley *et al.*, 1989; Weyher, 2009; Weyher *et al.*, 2006) suggested that primates that raid agricultural crops and rubbish (as long-tailed macaques do) may reduce their parasite loads as they gain a higher nutritional status that might facilitate a more effective immune response. Another factor to consider is the forest quality (i.e. tree density) as logging may result in higher foliage density which in turn may result in a greater surface area exposed to falling faeces (Gillespie *et al.*, 2005b; Gillespie & Chapman, 2006). Being a folivorous primate, it is likely that the proboscis monkey consumes a higher volume of resources than do macaques as frugivores and may therefore ingest more parasites if

more leaves are contaminated. Finally, Nunn *et al.* (2003) suggests that animals with a larger home range that travel a longer distance per day should encounter more parasite species. The home range of the proboscis monkeys is 900 ha (Sha *et al.*, 2008) compared with 200 ha reported for macaques (Goossens *et al.*, 2003a), hence this difference might also account for the differences in parasite fauna between the primate species encountered in this study.

In addition, there were no overall differences detected in parasite fauna from primates located on either side of the Kinabatangan River. This result was expected given that both primate species are good swimmers and have been reported to cross rivers, including the Kinabantangan (Bennet & Gombek, 1993; de Ruiter & Geffen, 1998; Jalil, 2007; personal observation). A difference was detected, though, between the long-tailed macaques and the proboscis monkeys in the north side of the river, where the latter had higher parasite richness than the former (Fig. 5.4 b), but there is no obvious explanation for this.

Gastrointestinal parasite classification by faecal analyses alone is limited; however, it is the only ethical method for surveying threatened species (Gillespie, 2006). Some studies have stressed the possibility of increased parasite transmission between primates and humans via ecotourism, primate research and the human encroachment in primate habitats (Chapman *et al.*, 2005a; Goldberg *et al.*, 2007). Although there was no link between human activity and non-human primate parasite diversity in the current study, this association may be apparent for other types of parasites that were not surveyed (i.e. Amoebas, Cryptosporidium, malaria). Unfortunately, sampling in the different areas of the sanctuary could not be carried out during the same season and the differences observed may also reflect temporal variation in parasite abundance and distribution. For instance, the prevalence of *Trichuris* sp. in the proboscis monkeys appears to be a gradient, where Area 1 has the greatest prevalence of the three areas examined (Area1>Area2>Area 3, Fig. 5.8), and this could be an effect biased by the sampling scheme (covering each area in different seasons), and the sample sizes (82 in Area1, 55 in Area2 and 9 in Area3). In future studies samples should be collected at least over the course of an annual cycle to assess the effects of climate, host activity, and life and physiological stages of both parasite and primate on infections. Molecular analysis (such as sequencing and Q-PCR) could also be used to improve the classification and quantification of the gastrointestinal parasites (Chapman *et al.*, 2009; Prichard & Tait, 2001; Zarlenga & Higgins, 2001). Molecular methods may enhance the sensitivity and specificity of the detection process and also reduce much of the subjectivity inherent in interpreting morphological data. For example, multiplex PCRs can substantially simplify analyses of mixed parasite populations, as described by Zarlenga *et al.* (2001) to differentiate gastrointestinal parasites in cattle. In addition, the final PCR product in a "real-time" assay can be used to deduce the starting number of target molecules in a given sample. If the

target molecule is parasite-derived, Q-PCR assays can be developed to ascertain parasite levels in infected hosts (Harmon *et al.*, 2007b; Zarlenga & Higgins, 2001). This could be done directly using the DNA extracted from the faeces and will be advantageous compared to time-consuming and relatively imprecise morphological identification and egg counting, but species-specific primers would need to be developed. Lacking this, coprocultures should be performed to identify parasite species, and plant/fruit parasite identification keys should also be considered to identify cryptic taxa (i.e. the 'oxyurid-like', '*Dipylidium*-like' and the 'Unknown Morph', found in the current study). In a study on howler monkey parasites (*Allouata pigra*) in Mexico, some unidentified morphs were subsequently identified as plant or fruit parasites (Diana Ramírez, personal communication).

In conclusion, parasite richness, prevalence and mixed infections, are likely to be affected by complex interactions among environmental, demographic, behavioural and genetic factors (Gillespie *et al.*, 2005), therefore an explanation for the variations found in this study is difficult. Identifying differences in parasites harboured by threatened and non-threatened host taxa represents one crucial step towards understanding the role of infectious agents in animal conservation (Altizer *et al.*, 2007). Hence, the results of this study represent a baseline for future work regarding parasite-host ecology in primates of the LKWS, and could be used in conservation and management plans.

CHAPTER SIX

General Discussion

Two important topics addressed in conservation biology are genetic variation and the role of infectious diseases in the management of fragmented wild populations of threatened species (Frankham, 2009; Smith *et al.*, 2009). Genetic impacts (particularly genetic variation, inbreeding and extinction risk) of population fragmentation depend critically upon connectivity and gene-flow among fragments (Frankham, 2006). Most of these parameters have been evaluated on the basis of neutral markers which are not expected to accurately capture the adaptive variation necessary for populations to thrive both in the short and long term, given changing environmental conditions. Therefore, current discussions in evolutionary ecology and conservation genetics have re-emphasized the importance of including markers of coding genes to identify adaptive and evolutionary relevant processes (Bruford, 2009; Hauffe & Sbordoni, 2009; Kohn *et al.*, 2006; Wayne & Morin; 2004). In addition, the reduced genetic diversity caused by habitat fragmentation may potentially increase susceptibility to infectious disease (Frankham, 2006). In situations where host populations aggregate into remaining patches of a fragmented landscape, the transmission success of pathogens may increase and this may lead to increases in the prevalence of some of them (McCallum & Dobson, 2006). In the case of primates, habitat fragmentation poses a severe threat to populations around the world (Cowlishaw & Dunbar, 2000; Mittermeier *et al.*, 2009), with clear evidence that effects of disease interact with habitat loss and other stress factors, with often catastrophic consequences (Nunn & Altizer, 2006; some examples in Chapter One).

The current research project attempted to address the questions of genetic variation, gene flow and parasite ecology for two primate species, the proboscis monkey (*Nasalis larvatus*) and the long tailed macaque (*Macaca fascicularis*). The Lower Kinabatangan Wildlife Sanctuary (LKWS) in Sabah, offers an ideal environment with which to test the effects of natural and anthropomorphic mediated habitat fragmentation on neutral and adaptive genetic diversity and on host-parasite ecology. In particular, the high primate diversity in the LKWS allows comparisons to be made between primates with different social systems and dispersal patterns. In this final Chapter, the outcomes of the proposed hypotheses (Chapters Three, Four and Five) are discussed in the context of the LKWS landscape, the previous primate population genetic studies in the area, the biology of the primates studied and in relation to the conservation and management of the area with recommendations for future studies.

It was hypothesised that long-tailed macaques, which exhibit male- biased dispersal, would have a less pronounced population structure than that obtained by previous mtDNA analysis (Jalil, 2007). In contrast, the female-transfer behaviour of the proboscis monkey was expected to reveal a pattern of genetic structure similar to that reported by Jalil (ibid) from mtDNA. Both species present high levels of genetic diversity (mean H_E= 0.8 and 0.68 for long-tailed macaques and proboscis monkeys respectively) and gene flow (F_{ST}= 0.005 and 0.012, respectively) despite the presence of a potential natural barrier: the Kinabatangan River. No population structuring for either primate species was detected by Genetix 4.05 (Belkhir *et al.*, 1996-2004) and GENELAND 3.1.5 (Guillot *et al.*, 2005a, b; Guillot, 2008). However, STRUCTURE (version 2.3.1; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009; Pritchard *et al.*, 2000) weakly assigned two different populations (for both primate species), perhaps a reflection of recent human-mediated habitat fragmentation (Appendices Six and Ten). These results are somewhat comparable to those of Jalil (2007) where analysis of molecular variance from mtDNA data indicated restricted gene-flow between populations of *M. fascicularis* on the north and south side of the Kinabatangan River. Genetic variation will be shared for some period of time between populations which have recently separated, even in the absence of gene exchange (Hey, 2006; Waples & Gaggiotti, 2006). Therefore, genetic differences due to a migration-drift balance will render a F_{ST} value which, by itself, cannot be distinguished from that yielded by an accumulation of genetic changes over time in completely isolated populations (Waples & Gaggiotti, 2006). These two scenarios could be distinguished by a non-equilibrium method such as the Isolation and Migration (IM) model developed by Hey and Nielsen (2004). The IM model allows variation in populations sizes over time and they are not assumed to be in migration-drift equilibrium (Hey & Nielsen, 2004; 2007). The current study did not evaluate this method due to time constraints, nevertheless it would be interesting to explore it further.

Rates of allelic dropout above 20% were detected within the datasets of both primate species. Such genotyping errors can bias population genetic studies (Dewoody *et al*, 2006, Hoffman & Amos, 2005; Roon *et al.*, 2005) and, with an apparent lack of population structuring, it was unexpected to have positive significant values of F_{IS} (indicating homozygotes excess) across every loci of long-tailed macaques (Table 3.9). These excesses can be explained by genotyping errors and presence of null alleles common problems when working with faecal samples and heterologous primers. DNA extracted from faeces is usually of poor quality (possibly degraded or accompanied by PCR inhibitors, especially in faeces that contains plant, bacterial and protozoan remnants, characteristic of folivorous primates such as the proboscis monkey) and quantity (often in the picogram range) (Broquet *et al.*, 2007; Launhardt, 1998; Monteiro *et al.*, 1997, Morin *et al.*, 2001; Vallet *et al*, 2008, Taberlet *et al.*, 1999). Reliable amplification of nuclear DNA (at least

3 positive PCRs) was possible for 71% of the long-tailed macaque samples, and for only 47% of the proboscis monkey samples. Samples were stored in 70% ethanol, a protocol previously used for primates in this area (Goossens *et al.*, 2003b; Jalil, 2007), however, DNA degradation might have occurred. A double step of ethanol-silica gel has proved useful (95.2% extraction rate success) for other primate species (Nsubuga *et al.*, 2004) and could be tested in future studies. Additionally, an extraction protocol based on the use of CTAB, an extensively used compound for plant DNA extractions (Bhattacharjee *et al.*, 2009; Murray & Thompson, 1980; Porebsky *et al.*, 1997; Vallet *et al.* 2007), should be used in the future to increase extraction success rates (i.e. 94%, Quéméré *et al.*, 2009), especially when working with folivorous primates such as the proboscis monkey. Additionally, rates of overall genotyping error rates and allelic drop were 16-33% and 15-28% for macaques and proboscis respectively. Various methods have been proposed to limit genotyping errors in non-invasive studies and the most accepted is the "multi-tubes" technique (Taberlet et al. 1996). However, this approach is very expensive, time-consuming and requires higher template volume, thus its application was not feasible for the current study. The simulation method (GEMINI) of Valière *et al.* (2002) was used for the macaque dataset to determine the number of PCR repetitions needed to obtain an accurate genotype. Although 94% of genotyping accuracy was predicted by GEMINI, error rates were still high (see above). GEMINI requires error rates for the input files prior to the determination of PCR repeats to obtain an accurate genotype, these "pre-error rates" were calculated by PEDANT (Johnson & Haydon, 2007) instead of through a pilot study using the "multi-tube" approach. A pilot study using the original Taberlet *et al.* (1996) approach to determine the "pre-error rates" might have given a different result regarding the error rates and hence the number of PCR repetitions needed to obtain an accurate genotype might have been different as well. Comparatively, neither the Taberlet *et al.* (1996) nor the Valière *et al.* (2002) methods were used for the proboscis monkey dataset, due to constrains of samples volume, and still the error rates were similar to those estimated for the macaques, whereas the estimation of null alleles differed between those two species.

Results from MICROCHECKER (version 2.2.3, van Oosterhout *et al.*, 2004) suggest that null alleles exist in every loci screened on the long-tailed macaques. This result was unexpected as 14 out of the 15 loci screened in this study had been used either in long-tailed macaques or in other species of the same genus confidently (Tables 3.3 and 3.4). The divergence time between hominoids and cercopithecids is 25 mya (Kumar & Hedger, 1998; Page & Goodman, 2001) hence it is highly probable that modifications on the molecular structure of regions surrounding the microsatellites occurred (although they were not sequenced as part of this study). Ten percent of the studies revised by Dakin and Avise (2004) report the redesign of primers in order to solve

the problem of null alleles. This is congruent with our proboscis monkey results since after testing almost 40 human derived microsatellite primers (Table 3.5) and obtaining spurious results the problem was solved when the markers were redesigned (Chapter Two). Only three of the eight loci screened in the proboscis monkey showed evidence for null alleles, but this could be an effect of allelic dropout rather than true null alleles and this explanation cannot be overruled for the macaques either. A new set of 499 microsatellite markers is now available for *Macaca fascicularis* (see Higashino *et al.*, 2009) and should be used in subsequent studies.

To date, previous population genetic analyses in primates of the LKWS have focused on neutral (demographically mediated) genetic variation (Goossens *et al.*, 2005 & 2006; Jalil, 2007, *et al.*, 2008). If the primate populations in the Kinabatangan have undergone recent habitat-mediated contraction, the separation of populations might be too short to leave a signal at neutral loci so that differences between populations might only be detectable at genes under selection. Therefore, this study presented the hypothesis that a more pronounced population structuring between river sides (or forest fragments) would be rendered through the analyses of a non-neutral molecular marker, such as the MHC. The genetic signatures of natural selection may be superimposed on the signatures of drift and/or gene flow (features that affect neutral variation). This hypothesis was not tested due to the technical difficulties encounter while performing the microsatellite analyses (low amplification success and high genotyping errors), and to the finite DNA availability. According to Knapp (2005c) identification of MHC alleles in any species is a complicated undertaking, no matter what type of sample is collected, and determination of haplotypes and homozygosity can be extremely problematic when using non-invasive samples since more errors can be introduced by the presence of few, damaged templates. In order to have a more accurate MHC allele identification, only samples from the microsatellite analysis that produced three positive PCR reactions with consistent genotypes were used (Table 4.1). Although over 300 samples were available for each primate species after DNA extraction, only 36 long-tailed macaque and 15 proboscis monkey samples were useful to characterize DRB alleles. The samples number was then inadequate to sustain a population genetic study and test the second hypothesis outlined in the introduction of this thesis. However, to take advantage of the available samples, the study was then made a preliminary one, where the genetic variability of the functionally important MHC gene *-DRB* (exon 2) was investigated on proboscis monkeys and long tailed macaques.

Future studies should investigate the role played by human barriers such as oil palm plantations, riparian villages, or roads in the development of genetic differentiation between remaining forest patches of the LKWS. Habitat fragmentation/connectivity is a fundamental concern in conservation biology as it affects extinction risk. To maintain connectivity, the creation of

corridors to increase the movement of both individuals and their genes is a popular approach in conservation management (Haddad & Tewsbury, 2006). Since the mid 1950s (but especially during the past 25 years), most of the Kinabatangan forest has been subjected to large-scale commercial timber exploitation or has been cleared for oil-palm development, resulting in a highly fragmented forest structure (McMorrow & Talip, 2001). Thus the aim of the LKWS is to create a corridor linking forest fragments isolated from one another by a wasteland of scrubland, agriculture and human settlements. To date, some of the fragments are partially connected by riparian corridors of variable width (some being as small as 20 m wide). In addition to the patchy landscape, the sanctuary is bisected along its length by the Kinabatangan River, and two major roads create extra partitioning, especially between Lots 8 - 10 and the remainder (see Fig. 5.1 in Chapter Five).

The effects of the LKWS geographical mosaic over the genetic diversity and structure of the orang-utan (Goossens *et al.*, 2005, 2006; Jalil *et al.*, 2008), the proboscis monkey and long tailed macaque (Jalil, 2007; current study) have been investigated using two classes of neutral molecular marker. High levels of genetic diversity are suggested by the studies based on mtDNA (Jalil, 2007; *et al.*, 2008), and microsatellites (Goossens *et al.*, 2005; Chapter Three, current study). The long-tailed macaques show the highest level of nuclear genetic diversity (H_E= 0.8), followed by the orang-utan (H_E= 0.74, Goossens *et al.*, 2005) and the proboscis monkey (H_E= 0.68). The preliminary *Mhc-DRB* analysis from this study also suggests higher levels of nuclear genetic diversity in the macaques than in the proboscis monkey, where surprisingly only five –*DRB* sequences were identified. However only a few proboscis monkey individuals were screened with just one set of generalist –*DRB* primers (Chapter Four). In contrast, the control region mtDNA analysis revealed a higher genetic diversity in the proboscis monkeys than in the other two species (Jalil, 2007). These results however probably reflect the different dispersal mechanisms of these primates, the ancestral populations that inhabited Borneo and/or the recent human-mediated habitat contraction. For instance, Jalil (2007) explains that proboscis monkey's female biased dispersal and the size of the harem could have contributed to the establishment of large numbers of mtDNA lines during the Pliocene. At the end of that era, several cold phases led to contraction of suitable habitat that was likely to reduce the population sizes of this species. As mtDNA is inherited by females only, Jalil (2007) suggests that the high level of mtDNA diversity in proboscis monkeys is proportioned to the large numbers of breeding females in the ancestral population before the onset of the glaciations. In the case of orang-utans, Goossens *et al.* (2005, 2006) suggest that recent anthropogenic environmental changes are the main cause of the population decline, and that this recent habitat loss and degradation may have led to the concentration of the surviving individuals in the remaining forest patches along the Kinabatangan

River, hence accounting for the high levels of genetic variation. In addition, Goossens *et al.* (2005) suggests that the very large numbers of orang-utans thought to have existed in the last centuries and millennia throughout the Kinabatangan area, together with the long generation time and lifespan of this species, might have allowed populations to retain diversity for long periods after habitat loss. Thus, the recent habitat contraction and the ancient population numbers could also explain the high levels of neutral and adaptive genetic diversity estimated for the long-tailed macaques in the current study.

Genetic methods such as F_{ST}, assignment tests, and coalescence can define the extent of connectivity and infer rates of gene flow between populations (Frankham, 2006). This allows rational scientific management of threatened species in the wild. In a highly fragmented ecosystem, such as the LKWS, it is important to determine whether there are barriers to gene flow, particularly for endangered species such as the orang-utan and the proboscis monkey. Goossens *et al.* (2005) and Jalil *et al.* (2008) demonstrated the importance of the river in shaping the genetic structure of the declining orang-utan populations in the LKWS. Gene flow is the rule rather than the exception between forest fragments from the same side of the river, despite the fragmentation, but gene flow was limited or not existent across the river. In contrast, the river did not pose a barrier to gene flow for proboscis monkey when assessed by mtDNA (Jalil, 2007) or microsatellite data (Chapter Three), in accordance with the swimming behaviour of this species. Perhaps surprisingly, Jalil (2007) found slight evidence of the river acting as a dispersal barrier for the long tailed macaques and with many of the closely related mtDNA haplotypes distributed on the same side of the river. In the current study, no evidence was found regarding a partitioning by the river of the long tailed macaque populations. Results by STRUCTURE suggest the presence of two populations, for both the proboscis monkey and the macaque, but not necessarily delimited by the river. This pattern could be a reflection of the recent human-mediated habitat fragmentation (Appendices Six and Ten) and should be explored further with an Isolation-Migration model approach (see above). Nevertheless, the results of both Jalil (2007) and the current study suggest that the LKWS supports admixed populations of both, proboscis monkeys and long tailed macaques. In species which do not follow a classical riverine-restricted pattern of genetic structure, the question therefore arises as to what determines their genetic structure. In particular, while neutral patterns of genetic structure may show limited differentiation, non-neutral markers may show a different pattern, sometimes of greater conservation significance (e.g. Bonin & Bernatchez, 2009; Bonin *et al.*, 2007; Aguilar *et al.*, 2004). Hence, information regarding adaptive genetic variation is still needed.

Bruford *et al.* (2010) incorporated the genetic data from Goossens *et al.* (2005, 2006) and Jalil *et al.* (2008) into a population viability analysis (PVA) for the LKWS orang-utan populations. A

combination of modest translocation rates (one female every 20 years) and corridor establishment is suggested to enable even the most isolated subpopulations to retain demographic stability and constrain localised inbreeding levels. As the natural demographic isolation between the populations north and south of the river has already been demonstrated (Goossens *et al* 2005; Jalil *et al.*, 2008), Bruford *et al.* (2010) recommended to maintain riverside populations as separate Management Units, as long as this is practically feasible, given the significantly greater F_{ST} values among riversides than within. An assessment such as this is needed for the proboscis monkey population in the LKWS. If translocations were required, the results from Jalil (2007) and the current study suggest that movement of individuals could potentially be performed between populations from both sides of the river or from distant forest patches without disrupting the gene-pools. However, better characterization of the fragmented subpopulations, for instance the number of individuals and densities in each Lot, or forest quality assessments are needed. Goossens *et al.* (2003a) determined a density of 6.9 and 6.3 proboscis monkey and long-tailed macaque individuals/km^2 respectively for the LKWS. However, those estimations are based on a survey that covered only a portion of Lots 2, 3 and 4, and observations during the current study (5 years later) suggest that densities are much higher in the region of Lots 5 – 7 and much lower in Lots 8 – 10. In particular, proboscis monkeys were absent from Batu Putih bridge (the division between Lots 7 and 8) up to Lok Kan River, and not many groups were observed in that locality either, when only four years ago, Jalil (2007) recorded the presence of several groups in Lots 8-10, including Lok Kan. The causes of this decline are unknown, but could be attributed to inadequate forest quality/food supplies, additional habitat loss, hunting, or disease, and should be investigated further.

Diseases are important when considering the conservation of endangered species (Cunningham, 1996; McCallum & Dobson, 2006, Smith *et al.*, 2009). Reserves created to protect animals from habitat loss and poaching could carry costs of increased disease risk, in part because these reserves may maintain populations at higher densities, especially in the absence of top carnivores or hunting by humans. Corridors that facilitate movement between populations could represent poorer-quality habitats and might also expose animals to parasites from domesticated animals, humans and other wildlife species that also use the corridors (Nunn & Altizer, 2006). Furthermore, as human population density continues to increase, speeding the reduction and fragmentation of primate habitats, greater human-primate contact is inevitable and even higher rates of parasite transmission are likely between humans and monkeys and between monkey species (Gillespie, 2006; Goldberg *et al.*, 2008). When moving between forest patches and because of the proximity to human settlements and to domestic animals, primates may be exposed to a wider range of parasitic vectors and/or intermediate hosts (Trejo-Macías *et al.*,

2007). These features are common characteristics of the LKWS that need to be considered in the management plans of the area. In addition, when animals are translocated they may "import" new parasites or immunologically naive animals may be released into an area where potentially-pathogenic parasites are endemic (Cunningham, 1996). Since translocation has already been advised for the orang-utan, it is important to identify host-parasite ecological patterns in the region. It was hypothesised that long-tailed macaques would have higher parasite diversity and to harbour more mixed infections than proboscis monkeys, since they are known to come closer to human habitation and to domestic fauna. Another hypothesis was that given the swimming ability of both primate species, no effect was expected from the natural barrier presented by the Kinabatangan River. Contrarily, it was predicted that areas of the sanctuary with higher rates of human activities, the parasite richness, the frequency of co-infections and the prevalence of particular parasites would be higher in both primate species, compared to areas with less human activity.

The parasite survey in the current study (Chapter Five) suggests that most of the gastro-intestinal parasites found are ubiquitous in the sanctuary, perhaps in agreement with the patterns of dispersal and gene flow detected in the hosts. The Kinabatangan River did not affect parasite diversity, or prevalence of particular parasites, or the proportion of individuals carrying a mixed infection. However, mapping of the parasite's distributions revealed that a few taxa were confined to particular areas of the LKWS (*Taenia* sp. in Lots 1-4, oxyurids in Lots 5-7, and oxyurid-like in Lots 8-10, data not shown), so future landscape epidemiology studies should be encouraged. A clearer pattern was distinguished for *Trichuris* sp. in proboscis monkeys, where the prevalence of this parasite was predicted to decrease gradually from downriver (Area 1 – Lots 1-4, see methods in Chapter Five) to upriver (Area 3), but caution must be applied since the number of samples differed between areas (82 in Area1, 55 in Area2 and 9 in Area3). In a hypothetical situation, if translocations of proboscis monkeys were needed given the low numbers of individuals present in Lots 8-10, translocating animals from Lots 5-7 would not pose an issue according to population genetics analyses, but they could cause a problem regarding disease transmission by increasing levels of environmental contamination. Yet, it must be noted that this prevalence gradient and the distributions above mentioned could be biased by the sampling scheme. Future studies on host-parasite ecology in the LKWS should cover (at least) the course of an annual cycle to assess the effects of climate, host activity, and life and physiological stages of both parasite and primate on infections.

It was expected that areas with high rates of human – non-human primate contact would have a positive correlation with parasite diversity, or prevalence of particular parasites, or the proportion of individuals carrying a mixed infection, but this was not always the case as host-parasite

dynamics are likely to be affected by complex interactions between environmental, and host demographic, behavioural and genetic factors. Future studies could evaluate in a more direct way the effects of humans (including tourists and researchers) and livestock on the parasite communities of LKWS primates. Moreover, surveys including more parasite and host taxa are desirable. For instance, the unidentified morph present in the macaque samples was also observed in a sample that presumably belonged to an otter (not shown) that was opportunistically collected in the current study. Another example comes from malaria, where *Plasmodium knowlesi* (the macaque malaria parasite) has been found in humans in Sabah (Singh *et al.*, 2004, Symposium on zoonosis & emerging diseases, 2008), but no research has been done to identify whether any of the human *Plasmodium* spp. infect non-human primates in this region (an interesting fact on the light of recent evidence on *P. falciparum* originating from gorillas, Liu *et al.*, 2010). In Uganda, Goldberg *et al.* (2008) found evidence that the forest fragmentation is a cause of bacterial transmission among humans, non-human primates and livestock, and the genetic similarity of the parasite between the three taxa increased as the anthropogenic disturbance within forest fragments increased. In addition, Goldberg *et al.* (2007) found that chimpanzees harboured bacteria genetically more similar to those of humans employed in chimpanzee-directed research and tourism than to those of humans from a local village. Hence, and also because different strains of parasites may vary in their virulence, parasite genetics should also be evaluated to determine levels of zoononsis/anthropozonoosis in the area. Additionally, some guidelines for primate ecotourism are presented by Muehlenbein & Ancrenaz (2009) and should be implemented or adapted to the conditions in the LKWS.

Reserve design requires improved understanding of how edge effects, logging, and fragmentation impact patterns of parasitism (Nunn & Altizer, 2006). Therefore, future research on parasites in the LKWS could also address host densities and edge effects. For example, in a study on recently immigrated African colobines into the same forest fragment Chapman *et al.* (2005b) noticed that the richness of parasite communities did not increase with increased host density, but the population numbers of red colobus increased when the prevalence and intensity of *Trichuris* sp. decreased while the numbers of the black-and-white colobus decreased as the prevalence and intensity of this parasite remained high. Chapman *et al.* (2006b) also noticed that the proportion of individuals with multiple infections was greater in edge than forest interior groups and that primates that raided agricultural corps had less severe infections compared to those individuals that did not. Finally, Gillespie and Chapman (2006) and Gillespie *et al.* (2005b) demonstrated that various forms of anthropogenic disturbance, specifically selective logging and forest fragmentation, alter the dynamics of gastrointestinal parasite infection in the human and nonhuman primate populations in the region of Kibale National Park, Uganda. Therefore, the

identification of physical (i.e., size, location, and topography) and biological (i.e., tree diversity, tree density, stump density, and primate and other reservoir hosts density) attributes than can be used to predict parasite infection dynamics in the LKWS primate populations should also be attempted in future studies.

Maintaining the ability of wild populations to respond evolutionarily to parasite-mediated selection could be one of the best long-term strategies for mitigating the risks of infectious diseases (Crandall *et al.*, 2000). Two main types of pathogen-driven modes of selection ('heterozygote advantage hypothesis' and 'frequency-dependence selection') have been suggested as important in retaining high levels of genetic diversity at the MHC in humans and vertebrates (reviewed in Bernatchez & Landry, 2003; Sommer, 2005; Spurgin & Richardson, 2010; but see van Oosterhout, 2009). The heterozygote advantage hypothesis (Doherty & Zinkernagel, 1975) proposes that individuals heterozygous at MHC loci are able to respond to a greater range of pathogen peptides than homozygotes and, consequently, benefit from increased resistance to pathogens. Heterozygotes are, therefore, more likely to have higher relative fitness and, as a result, on average more MHC alleles will persist in the population. The rare allele advantage hypothesis assumes that MHC diversity is maintained through frequency-dependent coevolutionary processes between hosts and parasites (Takahata and Nei, 1990). The most resistant allele will be favoured and spread through the population. However, it will not go into fixation because when the resistant allele becomes common, this increases selection on parasites to evade the recognition by this common allele. In the end this leads to an increased variability within a population. In a population subject to heterozygote advantage, one may therefore expect to observe associations between MHC heterozygosity and between both pathogen load and diversity (Spurgin & Richardson, 2010). The current study suggests that parasite richness and proportion of individuals carrying a mixed infection is higher in the proboscis monkey than in the long-tailed macaque (Chapter Five) and that MHC diversity/heterozygosity is lower in the proboscis monkeys compared to the macaques (Chapter Four). Could this be an indication of selection through the "heterozygote advantage" hypothesis? The answer to this question is beyond the scope of the current study, but it clearly poses a research line that should be explored in the future. Other studies regarding MHC diversity and pathogens should look for immunogenetic variants that result disadvantage/advantageous with respect to parasite burden and whether they are influenced by habitat fragmentation. For example, Schwensow *et al.* (2007) found evidence for a specific MHC supertype that was linked to infected individuals, a higher number of different nematode infections and high intensity of infection per individual, at the same time one rare supertype was revealed to be advantageous with respect to parasite

burden. This type of information could be used in the LKWS to identify vulnerable populations likely to be at risk from new pathogens.

The future of the primates in Kinabatangan depends on the adequate management of the sanctuary. The current study was challenging in many aspects, thus no full answers can be given to the posed hypothesis yet. However, even with its limitations, this study provides useful baseline information regarding neutral and adaptive genetic diversity, and host-parasite ecology of proboscis monkeys and long-tailed macaques of the LKWS. Moreover, it identifies lines for future research to answer questions on genetic structure, fitness and disease risk on these primates. The inclusion of these data and that from future studies is needed in order to have a complete management plan that will ensure the long-term survival of the proboscis monkey and the long-tailed macaque in the Kinabatangan floodplain.

REFERENCES

Abegg C, Thierry B. 2002. Macaque evolution and dispersal in insular south-east Asia. Biological Journal of the Linnean Society, 75: 555-76.

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. Proceedings of the National Academy of Sciences of the United States of America 101: 3490-4.

Aldrich-Blake FPG. 1980. Long-tailed macaques. In: Malayan Forest Primates: ten years' study in tropical rain forest. Chivers DJ (Ed.). Plenum Press, New York. New York, USA. pp. 364.

Allendorf FW, Luikart G. 2007. Conservation and the genetics of populations. Blackwell Publishing. pp 642.

Altizer SM, Harvell D, Friedle E. 2003. Rapid evolutionary dynamics and disease threats to biodiversity. Trends in Ecology and Evolution, 18: 589-96.

Altizer SM, Nunn CL, Lindenfors P. 2007. Do threatened hosts have fewer parasites? A comparative study in primates. The Journal of Animal Ecology, 76(2): 304-14.

Ancrenaz M, Ambu L, Sunjoto I, Ahmad E, Manokaran K, Meijaard E, Lackman I. 2010. Recent surveys in the forests of Ulu Segama Malua, Sabah, Malaysia, show that orang-utans (*P. p. morio*) can be maintained in slightly logged forests. PLoS One, 5(7): e11510.

Ancrenaz M, Goossens B, Gimenez O, Sawang A, & Lackman-Ancrenaz I. 2004. Determination of ape distribution and population size using ground and aerial surveys: a case study with orang-utans in lower Kinabatangan, Sabah, Malaysia. Animal Conservation, 7(4): 375-385.

Ancrenaz M, Setchell JM, James SS, Synior J, Dakog KH, Seventri AD, Goossens B. 2003. Preliminary results concerning the status of the Kinabatangan gibbon population. In: Lower Kinabatangan Scientific Expedition 2002. Mohamed M, Takano A, Goossens B, Indran R. (Eds.)Universiti Malaysia Sabah. Malaysia. pp. 37-45.

Anthony N, Johnson-Bawe M, Jeffery K, Clifford S, Abernethy K, Tutin C, *et al.*. 2007. The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. Proceedings of the National Academy of Sciences, 104(51): 20432.

Arandjelovic M, Guschanski K, Schubert G, Harris TR, Thalmann O, Siedel H, *et al.*. 2009. Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. Molecular Ecology Resources, 9(1): 28-36.

Ash LR, Orihel TC. 2007. Atlas of Human Parasitology, 5th Edition. American Society for Clinical Pathology Press, Chicago, Illinois, USA. pp. 525.

Atkinson CT, Dusek RJ, Woods KL, Iko WM. 2000. Pathogenicity of avian malaria in experimentally-infected Hawaii Amakihi. Journal of Wildlife Diseases, 36: 197-204.

Atkinson CT, Woods KL, Dusek RJ, Sileo RS, Iko WM. 1995. Wildlife disease and conservation in Hawaii: pathogenicity of avian malaria (*Plasmodium relictum*) in experimentally infected iiwi (*Vestiaria coccinea*). Parasitology, 111: S59-69.

Avise JC, 2000. Phylogeography: the history and formation of species. Harvard University Press. pp. 447.

Avise JC, 2008. Phylogeography: retrospect and prospect. J. Biogeogr. 36: 3–15.

Ayres JM, Clutton-Brock TH (1992) River boundaries and species range size in Amazonian primates. American Naturalist, 140, 531–537.

Azmi R, 1998. Natural vegetation of the Kinabatangan floodplain. Part 1: Background and preliminary checklist. Report WWF- Malaysia, Kota Kinabalu, Sabah.

Babik W. 2010. Methods for MHC genotyping in non-model vertebrates. Molecular Ecology Resources, 10(2): 237-251.

Balloux F, Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology, 11: 155-65.

Bayes MK, Smith KL, Alberts SC, Altmann J, Bruford MW. 2000. Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. Conservation Genetics, 1: 173-76.

Beebe T, Rowe G. 2004. An introduction to molecular ecology. Oxford University Press, Oxford. pp. 346.

Belkhir K, Borsa P, Chikhi L, Raufaste N & Bonhomme F. 1996-2004 GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier France.

Bennett EL, Davis G. 1994. The ecology of Asian colobines. In: Colobine Monkeys: their ecology, behaviour and evolution. Davies AG & Oates JF (Eds.) Cambridge University Press, Cambridge UK. pp.129-171.

Bennett EL, Gombek F. 1993. Proboscis Monkeys of Borneo. Natural History Publications. Borneo Snd. Bhd, Kota Kinabalu.

Bennett EL, Sebastian AC. 1988. Social organization and ecology of proboscis monkeys (*Nasalis larvatus*) in mixed coastal forest in Sarawak. International Journal of Primatology 9:233-253.

Bergl RA, Vigilant L. 2007. Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). Molecular Ecology, 16(3): 501-16.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology, 16: 363-77.

Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, *et al.*. 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet, 367(9521): 1521-32.

Bhagavatula J, Singh L. 2006. Genotyping faecal samples of Bengal tiger (*Panthera tigris tigris*) for population estimation: a pilot study. BMC Genetics: 7: 48.

Bhargava A, Fuentes FF. 2010. Mutational dynamics of microsatellites. Molecular Biotechnology, 44(3): 250-66

Bhattacharjee R, Ferguson M, Gedil M, Dumet D, Ingelbrecht I. 2009. Field collection, preservation and large scale DNA extraction procedures for cassava (*Manihot esculenta Crantz*). African Journal of Biotechnology, 8(15): 3424–3430.

Blancher A, Tisseyre P, Dutaur M, Apoil P, Maurer C, Quesniaux V, *et al.*. 2006. Study of Cynomolgus monkey (*Macaca fascicularis*) MhcDRB (*Mafa-DRB*) polymorphism in two populations. Immunogenetics, 58(4): 269-282.

Bonhomme M, Blancher A, Crouau-Roy B. 2005. Multiplexed microsatellites for rapid identification and characterization of individuals and populations of Cercopithecidae. American Journal of Primatology, 67: 385-391.

Bonin A, Bernatchez L. 2009. Challenges in assessing adaptive genetic diversity: overview of methods and empirical illustrations. In: Population Genetics for Animal Conservation. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C. (Eds.). Cambridge University Press, Cambridge. p. 123-147.

Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P. 2007. Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. Conservation biology, 21(3): 697-708.

Bonneaud C, Pérez-Tris J, Federici P, Chastel O, Sorci G. 2006. Major Histocompatibility alleles associated with local resistance to malaria in a passerine. Evolution, 60: 383-89.

Bontrop RE, Otting N, de Groot N, Doxiadis GGM. 1999. Major histocompatibility complex class II polymorphisms in primates. Immunological Reviews, 167: 339-50.

Bontrop RE. 2006. Comparative genetics of MHC polymorphisms in different primate species: duplications and deletions. Human immunology, 67(6): 388-97.

Boonratana R. 2000. Ranging behavior of proboscis monkeys (*Nasalis larvatus*) in the Lower Kinabatangan, Northern Borneo. International Journal of Primatology, 21: 497-518.

Boonratana R. 2002. Social organisation of proboscis monkey (*Nasalis larvatus*) in the Lower Kinabatangan, Sabah, Malaysia. Malayan Nature Journal, 56: 57-75.

Bos DH, Turner SM, DeWoody JA. 2007. Haplotype inference from diploid sequence data: evaluating performance using non-neutral MHC sequences. Hereditas, 144(6): 228-234.

Bradley BJ, Boesch C, Vigilant L. 2000. Identification and redesign of human microsatellite markers for genotyping wild chimpanzee (*Pan troglodytes verus*) and gorilla (*Gorilla gorilla gorilla*) DNA from faeces. Conservation Genetics, 1: 289-92.

Broquet T, Ménard N, Petit E. 2007. Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conservation Genetics, 8: 249-260.

Brown JH, Jardetzky T, Saper MA, Samraoui B, Bjorkman PJ, Wiley DC, *et al.*. 1988. A hypothetical model of the foreign antigen binding site of Class II histocompatibility molecules. Nature, 332: 845-850.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, *et al.*. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature, 364: 33–39.

Bruford MW, Ancrenaz M, Chikhi L, Lackman-Ancrenaz I, Andau M, Ambu L, Goossens B. 2010. Projecting genetic diversity and population viability for the fragmented orang-utan population in the Kinabatangan floodplain, Sabah, Malaysia. Endangered Species Research, 12: 249-262.

Bruford MW, Ciofi C, Funk SM. 1998. Characteristics of microsatellites. In: Molecular tools for screening biodiversity. Karp A, Isaac PG, Ingram DS (Eds.). Chapman & Hall. London, UK. pp. 202-205.

Bruford MW, Hanotte O, Brookfield JFY, Burke T. 1992. Single-locus and multilocus DNA fingerprinting. In: Molecular Genetic Analysis at Populations: A Practical Approach. Hoezel AR (Ed.) IRL Press, Oxford. pp. 225-265

Bruford MW, Wayne RK. 1993. Microsatellites and their application to population genetic studies. Current Opinion in Genetics & Development, 3(6): 939-43.

Bruford MW. 2009. Future-proofing genetic units for conservation: time's up for subspecies as the debate gets out of neutral! In: Population Genetics for Animal Conservation. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C. (Eds.). Cambridge University Press, Cambridge. p. 227-240.

Buschiazzo E, Gemmell NJ. 2006. The rise, fall and renaissance of microsatellites in eukaryotic genomes. BioEssays: news and reviews in molecular, cellular and developmental biology, 28(10): 1040-50.

Bush AO, Lafferty KD, Lotz JM, Shostak AW. 1997. Parasitology meets ecology on its own terms: Margolis *et al.* revisited. Journal of Parasitology, 83(4): 575-583.

CAIMS (Conservation Areas Information and Monitoring Systems). 2005. Class 6 Forest Reserves. Retrieved 12 February, 2007 from: http://www.sabah.gov.my/htan_caims/Level%201%20 frame%20pgs/map_search_fr.htm

Carlsson J. 2008. Effects of microsatellite null alleles on assignment testing. The Journal of Heredity, 99(6): 616-23.

Chambers GK, MacAvoy ES. 2000. Microsatellites: consensus and controversy. Comparative biochemistry and physiology. Part B, Biochemistry & Molecular Biology, 126(4): 455-476.

Chapman CA, Gillespie TR, Speirs ML. 2005b. Parasite prevalence and richness in sympatric colobines: effects of host density. American Journal of Primatology, 67(2): 259-266.

Chapman CA, Gilliespie TR, Goldberg TL. 2005a. Primates and the ecology of their infectious diseases: how will anthropogenic change affect host-parasite interactions? Evolutionary Anthropology, 14: 134–144.

Chapman CA, Huffman MA, Ryan SJ, Sengupta R, Goldberg TL. 2009. Ways forward in the study of primate parasite ecology. In: Primate Parasite Ecology. The dynamics and study of host-parasite relationships. Huffman MA, Chapman CA. (Eds.). Cambridge University Press, Cambridge. p. 487-505.

Chapman CA, Peres CA. 2001. Primate conservation in the new millennium: the role of scientists. Evolutionary Anthropology, 10(1): 16–33.

Chapman CA, Speirs ML, Gillespie TR, Holland T, Austad KM. 2006b. Life on the edge: gastrointestinal parasites from the forest edge and interior primate groups. American Journal of Primatology, 68(4): 397.

Chapman CA, Wasserman MD, Gillespie TR, Speirs ML, Lawes MJ, Saj TL, *et al.*. 2006a. Do food availability, parasitism, and stress have synergistic effects on red colobus populations living in forest fragments? American Journal of Physical Anthropology, 131(4): 525-534.

Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution, 24(3): 621-631.

Ciofi C, Funk SM, Coote T, Cheesman DJ, Hammond RL, Sacceri IJ, Bruford MW. 1998. Genotyping with microsatellite markers. In: Molecular tools for screening biodiversity. Karp A, Isaac PG, Ingram DS. (Eds.). Chapman & Hall. London, UK. p. 195-201.

Cleaveland S, Hess GR, Dobson AP, Laurenson MK, McCallum HI, Roberts MG, Woodroffe R. 2002. The role of pathogens in biological conservation. In: The Ecology of Wildlife Diseases. Hudson PJ, Rizzoli A, Grenfell BT, Heesterbeek H, Dobson AP. (Eds.). Oxford University Press, Oxford. p. 139-150.

Cleaveland S, Laurenson MK, Taylor LH. 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philosophical Transcripts of the Royal Society of London B, 356: 991–999.

Clifford SL, Anthony NM, Bawe-Johnson M, Abernethy KA, Tutin CE, White LJ, Bermejo M, Goldsmith ML, Mcfarland K, Jeffery KJ, Bruford MW, Wickings EJ. 2004. Mitochondrial DNA phylogeography of western lowland gorillas (*Gorilla gorilla gorilla*). Molecular Ecology, 13: 1551–1561.

Clifford SL, Jeffrey K, Bruford MW, Wickings EJ. 1999. Identification of polymorphic microsatellite loci in the gorilla (*Gorilla gorilla gorilla*) using human primers: application to noninvasively collected hair samples. Molecular Ecology, 8: 1551-1556.

Clisson I, Lathuilliere M, Crouau-Roy B. 2000. Conservation and evolution of microsatellite loci in primate taxa. American Journal of Primatology, 50(3): 205-14.

Cohen S. 2002. Strong positive selection and habitat-specific amino acid substitution patterns in MHC from an estuarine fish under intense pollution stress. Molecular Biology and Evolution, 19(11): 1870-1880.

Collins AC, Dubach JM 2000. Biogeographic and ecological forces responsible for speciation in Ateles. International Journal of Primatology, 21: 421–444.

Coote T, Bruford MW. 1996. Human microsatellites applicable for analysis of genetic variation in apes and old world monkeys. The Journal of Heredity, 87: 406-410.

Couette S. 2007. Morphologic and genetic differentiation of night monkey (*Aotus infulatus*, Primates, Platyrrhinians, Cebidae) populations from the right and left banks of Rio Tocantins (Brazil). Comptes Rendus Biologies, 330(2): 148-158.

Coulon A, Fitzpatrick JW, Bowman R, Stith BM, Makarewich CA, Stenzler LM, *et al.* 2008. Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). Molecular Ecology, 17(7): 1685-1701.

Cowlishaw G, Dunbar RIM. 2000. Primate Conservation Biology. University of Chicago Press, Chicago Il. pp. 498.

Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution, 15(7): 290–295.

Craul M, Radespiel U, Rasolofoson DW, Rakotondratsimba G, Rakotonirainy O, Rasoloharijaona S, Randrianambinina B, Ratsimbazafy J, Ratelolahy F, Randrianamboavaonjy T, Rakotozafy L. 2008. Large rivers do not always act as species barriers for *Lepilemur* sp.. Primates, 49(3): 211-218.

Cunningham AA. 1996. Disease Risks of Wildlife Translocations. Conservation Biology, 10: 349-353.

Dakin EE, Avise JC. 2004. Microsatellite null alleles in parentage analysis. Heredity, 93(5): 504-509.

Daszak P, Berger L, Cunningham AA, Hyatt AD, Green DE, Speare R. 1999. Emerging infectious diseases and amphibian population declines. Emerging Infectious Diseases, 5: 735-748.

Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious diseases of wildlife - Threats to biodiversity and human health. Science, 287: 443-449.

Daszak P, Cunningham AA, Hyatt, AD. 2001. Anthropogenic environmental change and the emergence of infectious disease in wildlife. Acta Tropica 78: 103–16.

Daszak P, Epstein JH, Kilpatrick AM, Aguirre AA, Karesh WB, Cunningham AA. 2007. Collaborative research approaches to the role of wildlife in zoonotic disease emergence. Current Topics in Microbiology and Immunology, 315: 463–475.

Daszak P, Tabor GM, Kilpatrick AM, Epstein J, Plowright R. 2004. Conservation medicine and a new agenda for emerging diseases. Annals of the New York Academy of Science, 1026: 1-11.

Davies AG, Oates JF. 1994. Colobine monkeys: their ecology, behaviour and evolution. Cambridge University Press, Cambridge. UK.

Davison GWH. 2002. Rehabilitation and restoration of habitat neat the Kinabatangan Wildlife Sanctuary, Sabah, Malaysia. Retrieved 6 February 2007, from: http://www.iucn.org/themes/fcp/publications/files/flr_costarica/flr_davison_ext.doc

De Barba M, Waits LP. 2009. Multiplex pre-amplification for noninvasive genetic sampling: is the extra effort worth it? Molecular Ecology Resources, 10(4): 659-665.

de Ruiter JR, Geffen E. 1998. Relatedness of matrilines, dispersing males and social groups in long-tailed macaques (*Macaca fascicularis*). Proceedings of the Royal Society B: Biological Sciences, 265 (1391): 79.

de Ruiter JR, Van Hooff JARAM, Scheffrahn W. 1994. Social and genetic aspects of paternity in wild long-tailed macaques (*Macaca fascicularis*). Behaviour, 129: 203–224.

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc B.: 39: 1–38.

DeSalle R, Amato G. 2004. The expansion of conservation genetics. Nature Reviews Genetics, 5(9): 702-12.

Deucher A, Chiang T, Schrijver I. 2010. Rare sequence variation in the genome flanking a short tandem repeat locus can lead to a question of "nonmaternity". The Journal of Molecular Diagnostics, 12(3): 384-389.

Dewoody J, Nason JD, Hipkins VD. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes, 6(4): 951-957.

Do H. 2009. Gastrointestinal parasites of the Delacour's langur (*Trachypithecus delacouri*): Comparison between caged, semi-wild, and free-ranging individuals. Vietnamese Journal of Primatology, 3: 83-88.

Dobson AP, Grenfell B (Eds.) 1995. Ecology of infectious disease in natural populations. Cambridge University Press, Cambridge.

Doherty PC, Zinkernagel RM. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. Nature, 256: 50-52.

Domínguez-Domínguez O, Vázquez-Domínguez E. 2009. Filogeografía: aplicaciones en taxonomía y conservación. Animal Biodiversity and Conservation, 32(1): 59-70.

Doxiadis GG, de Groot N, de Groot NG, Rotmans G, de Vos-Rouweler AJ, Bontrop RE, *et al.*. 2010. Extensive DRB region diversity in cynomolgus macaques: recombination as a driving force. Immunogenetics, 62(3): 137-147.

Doxiadis GG, Rouweler AJ, de Groot NG, Louwerse A, Otting N, Verschoor EJ, *et al.*. 2006. Extensive sharing of MHC class II alleles between rhesus and cynomolgus macaques. Immunogenetics, 58(4): 259-268.

Doxiadis GGM, Otting N, de Groot NG, Noort R, Bontrop RE. 2000. Unprecedented polymorphism of *Mhc-DRB* region configurations in rhesus macaques. The Journal of Immunology, 164: 3193-3199.

Doxiadis GGM, Otting N, de Groot NG, Rouweler AJM, Noort R, Verschoor EJ, *et al.*. 2003. Evolutionary stability of MHC class II haplotypes in diverse rhesus macaque populations. Immunogenetics, 55(8): 540-551.

DPDx -Division of Parasitic Diseases. CD-ROM for Laboratory identification of Parasitic Diseases and website. Centres for Disease Control and Prevention, Georgia. Last accessed: 28 July 2010, from http://www.dpd.cdc.gov/dpdx/Default.htm

Dryden MW, Payne PA, Ridley R, Smith V. 2005. Comparison of common fecal flotation techniques for the recovery of parasite eggs and oocysts. Veterinary Therapeutics, 6: 15-28.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol., 11: 2571–2581.

Egwang TG, Slocombe JOD. 1982. Evaluation of the Cornell-Wisconsin centrifugal flotation technique for recovering Trichostrongylid eggs from bovine feces. Canadian Journal of Comparative Medicine, 46: 133-137.

Ekanayake DK, Arulkanthan A, Horadagoda NU, Sanjeevani GK, Kieft R, Gunatilake S, *et al.*. 2006. Prevalence of cryptosporidium and other enteric parasites among wild non-human primates in Polonnaruwa, Sri Lanka. The American Journal of Tropical Medicine and Hygiene, 74(2): 322-329.

Eley RM, Strum SC, Muchemi G, Reid GDF. 1989. Nutrition, body condition, activity patterns, and parasitism of free-ranging troops of olive baboons (*Papio anubis*) in Kenya. American Journal of Primatology, 18: 209-219.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics, 5(6): 435-45.

Engelhardt A, Heistermann M, Hodges JK, Nürnberg P, Niemitz C. 2006. Determinants of male reproductive success in wild long-tailed macaques (*Macaca fascicularis*) male monopolisation, female mate choice or post-copulatory mechanisms? Behavioral Ecology and Sociobiology, 59(6): 740-752.

Environmental and Conservation Working Group. 2001. Sabah structure plan 2020. Retrieved 15 February, 2007 from: http://www.sabah.gov.my/jpas/programs/ecd-cab/technical/StructPlan.pdf.

Eriksson J, Hohmann G, Boesch C, Vigilant L. 2004. Rivers influence the population genetic structure of bonobos (*Pan paniscus*). Molecular Ecology, 13: 3425–3435.

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.

Evans BJ, Supriatna J, Melnick DJ. 2001. Hybridization and population genetics of two macaque species in Sulawesi, Indonesia. Evolution, 55(8); 1686–1702.

Excoffier L, Lischer HE. 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

Faircloth BC. 2008. MSATCOMMANDER: Detection of Microsatellite Repeat Arrays and Automated, Locus-Specific Primer Design. Molecular Ecology Resources, 8: 92-94

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics, 164: 1567–1587.

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes, 7: 574–578.

Fleagle J. 1998. Primate adaptation and evolution. 2nd Edition. Academic Press.

Fogelqvist J, Niittyvuopio A, Ågren J, Savolainen O, Lascoux M. 2010. Cryptic population genetic structure: the number of inferred clusters depends on sample size. Molecular Ecology Resources, 10(2): 314-323.

Food and Agriculture Organization (FAO) of the United Nations. 2007. The RVC/FAO guide to veterinary diagnostic parasitology: Faecal examination of farm animals for helminth parasites. Compiled by: Gibbons LM, Jacobs DE, Fox MT and Hansen J. Last accessed 28 July 2010, from: http://www.rvc.ac.uk/review/Parasitology/Index/Index.htm

Fox JC. Veterinary Clinical Parasitology Images. Oklahoma State University, College of Veterinary Medicine. Parasitology teaching resources. Last accessed 28 July 2010, from: http://instruction.cvhs.okstate.edu/jcfox/HTDOCS/CLINPARA/clinpara.htm

Frankham R, Ballou JD, Briscoe DA. 2002. Introduction to conservation genetics. Cambridge University Press, Cambridge, UK. pp. 617.

Frankham R. 2006. Genetics and landscape connectivity. In: Connectivity conservation. Crooks KR, Sanjayan M. (Eds.) Cambridge University Press, Cambridge. p. 72-96.

Frankham R. 2009. Where are we in conservation genetics and where do we need to go? Conservation Genetics, 11(2): 661-663.

Frantz AC, Pope LC, Carpenter PJ, Roper TJ, Wilson GJ, Delahay RJ, *et al.*. 2003. Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology, 12(6): 1649-1661.

Garamszegi LZ, de Groot NG, Bontrop RE. 2009. Correlated evolution of nucleotide substitution rates and allelic variation in *Mhc-DRB* lineages of primates. BMC Evolutionary Biology, 9: 73.

Garnier JN, Bruford MW, Goossens B. 2001. Mating system and reproductive skew in the black rhinoceros. Molecular Ecology, 10(8): 2031-2041.

Garrick RC, Sunnucks P, Dyer RJ. 2010. Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. BMC Evolutionary Biology, 10: 118-135.

Gascon C, Malcolm J, Patton J, da Silva M, Bogart JP, Lougheed SC, *et al.* 2000. Riverine barriers and the geographic distribution of Amazonian species. Proceedings of the National Academy of Sciences of the United States of America, 97(25): 13672.

Gillespie TR, Chapman CA, Greiner EC. 2005b. Effects of logging on gastrointestinal parasite infections and infection risk in African primates. Journal of Applied Ecology, 42(4): 699-707.

Gillespie TR, Chapman CA. 2006. Prediction of Parasite Infection Dynamics in Primate Metapopulations Based on Attributes of Forest Fragmentation. Conservation Biology, 20(2): 441-448.

Gillespie TR, Greiner EC, Chapman CA. 2005a. Gastrointestinal parasites of the colobus monkeys of Uganda. The Journal of Parasitology, 91(3): 569-573.

Gillespie TR. 2006. Noninvasive Assessment of Gastrointestinal Parasite Infections in Free-Ranging Primates. International Journal of Primatology, 27(4): 1129-1143.

Glenn TC, Schable NA. 2005. Isolating microsatellite DNA loci. Methods in Enzymology, 395: 202-222.

Goldberg TL, Gillespie TR, Rwego IB, Estoff EL, Chapman CA. 2008. Forest Fragmentation as Cause of Bacterial Transmission among Nonhuman Primates, Humans, and Livestock, Uganda. Emerging Infectious Diseases, 14(9): 1375-1382.

Goldberg TL, Gillespie TR, Rwego IB, Wheeler E, Estoff EL, Chapman CA. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. Biological Conservation, 135: 511–517.

Gonder MK, Disotell TR, Oates JF. 2006. New genetic evidence on the evolution of chimpanzee populations and implications for taxonomy. International Journal of Primatology, 27: 1103-1127.

Gonder MK, Oates JF, Disotell TR, Forstner MRJ, Morales JC, Melnick DJ. 1997. A new west African chimpanzee subspecies? Nature, 388: 337.

Goossens B, Anthony N, Jeffery K, Johnson-Bawe M, Bruford MW. 2003b. Collection, storage and analysis of non-invasive genetic material in primate biology. In: Field and laboratory methods in primatology: A practical Guide. Setchell JM, Curtis DJ. (Eds.). Cambridge University Press. Cambridge, UK. pp. 295-308.

Goossens B, Bruford MW. 2009. Non-invasive genetic analysis in conservation. In: Population Genetics for Animal Conservation. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C. (Eds.). Cambridge University Press, Cambridge. p 167 -201.

Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau P, Bruford MW. 2006a. Genetic signature of anthropogenic population collapse in orang-utans. Public Library of Science Biology, 4: 285-91.

Goossens B, Chikhi L, Jalil MF, Ancrenaz M, Lackman-Ancrenaz I, Mohamed M, Andau P, Bruford W. 2005. Patterns of genetic diversity and migration in increasingly fragmented and

declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia. Molecular Ecology, 14: 441-56.

Goossens B, Funk SM, Vidal C, Latour S, Jamart A, Ancrenaz, Wickings EJ, Tutin CEG, Bruford MW. 2002. Measuring genetic diversity in translocation programmes: principles and application to a chimpanzee release project. Animal Conservation, 5: 225-36.

Goossens B, Setchell JM, James SS, Funk SM, Chikhi L, Abulani A, Ancrenaz M, Lackman-Ancrenaz I, Bruford MW. 2006b. Philopatry and reproductive success in Bornean orang-utans (*Pongo pygmaeus*). Molecular Ecology, 15: 2577-88.

Goossens B, Stechell JM, Abulani DMA, Jalil MF, James SS, Aris SH, Lakim MH, Seventri AD, Sariningsih SS, Ancrenaz M. 2003a. A boat survey of primates in the Lower Kinabatangan Wildlife Sanctuary. In: Lower Kinabatangan Scientific Expedition 2002. Mohamed M, Takano A, Goossens B, Indran R. (Eds.): Universiti Malaysia Sabah. Malaysia. p. 37-45.

Goossens, B, Chikhi L, Utami S, de Ruiter J, Bruford MW. 2000. A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. Conservation genetics, 1: 157-62.

Gotoh S, Takenaka O, Watanabe K, Hamada Y, Kawamoto Y, Watanabe T, *et al.*. 2001. Hematological values and parasite fauna in free-ranging *Macaca hecki* and the *M. hecki/M. tonkeana* hybrid group of Sulawesi Island, Indonesia. Primates, 42(1): 27-34.

Gotoh S. 2000. Regional differences in the infection of wild Japanese macaques by gastrointestinal helminth parasites. Primates, 41(3): 291-298.

Greiner EC, McIntosh A. 2009. Collection methods and diagnostic procedures for primate parasitology. In: Primate parasite ecology. The dynamics and study of host-parasite relationships. Huffman MA, Chapman CA. (Eds.). Cambridge University Press, Cambridge. p. 3-27.

Guillot G, Estoup A, Mortier F, Cosson J-F. 2005a. A spatial statistical model for landscape genetics. Genetics. 170:1261–1280.

Guillot G, Mortier F, Estoup A. 2005b. Geneland: a computer package for landscape genetics. Molecular Ecology Notes. 5:708–711.

Guillot G. 2008. Inference of structure in subdivided populations at low levels of genetic differentiation. The correlated allele frequencies model revisited. Bionformatics, 24: 2222–2228.

Haddad NM, Tewksbury JJ. 2006. Impacts of corridors on populations and communities. In: Connectivity conservation. Crooks KR, Sanjayan M. (Eds.) Cambridge University Press, Cambridge. p. 416-443.

Hadfield RM, Pullen JG, Davies KF, Wolfensohn SE, Kemnitz JW, Weeks DE, Bennett ST, Kennedy SH. 2001. Toward developing a genome-wide microsatellite marker set for linkage analysis in the rhesus macaque (*Macaca mulatta*): Identification of 76 polymorphic markers. American Journal of Primatology, 54: 223-31.

Hahn NE, Proulx D, Muruthi PM, Alberts S, Altmann J. 2003. Gastrointestinal parasites in Free-ranging Kenyan Baboons (*Papio cynocephalus* and *P. anubis*). International Journal of Primatology, 24(2): 271-279.

Hancock JM. 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Microsatellites: Evolution and Applications. Goldstein D, Schlötterer C. (Eds.). Oxford University Press, New York. p. 1-9.

Hansen H, Ben-David M, Mcdonald DB. 2008. Effects of genotyping protocols on success and errors in identifying individual river otters (*Lontra canadensis*) from their faeces. Molecular Ecology Resources, 8(2); 282-289.

Harcourt AH & Schwartz MW. 2001. Primate evolution: a biology of Holocene extinction and survival on the southeast Asian Sunda Shelf islands. American Journal of Physical Anthropology 114:4-17.

Harmon AF, Williams ZB, Holler LD, Hildreth MB. 2007a. Comparison of three different preservatives for morphological and real-time PCR analyses of *Haemonchus contortus* eggs. Veterinary Parasitology, 145: 361-365.

Harmon AF, Williams, ZB, Zarlenga DS, Hildreth MB, 2007b. Real-time PCR for quantifying *Haemonchus contortus* eggs and potential limiting factors. Parastiol Res, 101(1): 71-76.

Hasegawa H, Matsuo K, Onuma M. 2003. *Enterobius* (*Colobenterobius*) *serratus* sp. nov. (Nematoda: Oxyuridae) from the Proboscis Monkey (*Nasalis larvatus*) (Wurmb, 1787) (Primates: Cercopithecidae: Colobinae), in Sarawak, Borneo, Malaysia. Comparative Parasitology, 70(2): 128-131.

Hauffe C, Sbordoni V. 2009. Introduction. In: Population Genetics for Animal Conservation. Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C. (Eds.). Cambridge University Press, Cambridge. p. 1-21.

Hausfater G, Meade BJ. 1982. Alternation of sleeping groves by yellow baboons (*Papio cynocephalus*) as a strategy for parasite avoidance. Primates, 23(2): 287-297.

He W, Lin L, Shen F, Zhang W, Zhang Z, King E, Yue B. 2008. Genetic diversities of the giant panda (*Ailuropoda melanoleuca*) in wanglang and baoxing nature reserves. Conservation Genetics 9: 1541–1546.

Hedrick P. 2001. Conservation genetics: where are we now? Trends in Ecology & Evolution, 16(11): 629-636.

Hedrick PW. 1985. Genetics of populations. Jones and Bartlett publishers, Inc, University of Kansas, Boston.

Hernandez AD, Macintosh AJ, Huffman MA. 2009. Primate Parasite Ecology: patterns and predictions from an on-going study of Japanese macaques. In: Primate Parasite Ecology. The Dynamics and Study of Host-Parasite Relationships Huffman MA, Chapman CA. (Eds.). Cambridge University Press. p. 387-401.

Hess G. 1996. Disease in metapopulation models: implications for conservation. Ecology, 77: 1617–1632.

Hey J. 2006. Recent advances in assessing gene flow between diverging populations and species. Current Opinion in Genetics & Development, 16(6): 592-596.

Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics, 167(2): 747-760.

Hey J, Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proceedings of the National Academy of Sciences of the United States of America, 104(8): 2785-2790.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, *et al.*. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution, 54(1): 291-301.

Higashino A, Osada N, Suto Y, Hirata M, Kameoka Y, Takahashi I, *et al.* 2009. Development of an integrative database with 499 novel microsatellite markers for *Macaca fascicularis*. BMC Genetics, 10: 24.

Hill AVS, Allsopp CEM, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, Brewster D, McMichael AJ, Greenwood BM. 1991. Common West African HLA antigens are associated with protection from severe malaria. Nature, 352: 595-600.

Hoelzer GA, Dittus WPJ, Ashley MV, Melnick DJ. 1994. The local distribution of highly divergent mitochondrial DNA haplotypes in toque macaques *Macaca sinica* at Polonnaruwa, Sri Lanka. Molecular Ecology 3: 451-458.

Hoffman JI, Amos W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Molecular ecology, 14(2): 599-612.

Holderegger R, Kamm U, Gugerli F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology, 21(6), 797-807.

Horii Y, Imada I, Yanagida T, Usui M, Mori A. 1982. Parasite changes and their influence on the body weight of Japanese monkeys (*Macaca fuscata fuscata*) of the Koshima troop. Primates, 23(3): 416-431.

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

Hudson PJ, Rizzoli A, Grenfell BT, Heesterbeek H, Dobson AP. 2006. The ecology of wildlife diseases. Oxford University Press, Oxford. pp. 197.

Huffman MA. 1997. Current evidence for self-medication in primates: A multidisciplinary perspective. American Journal of Physical Anthropology, 104(S25): 171-200.

Hughes A. 2008. Major Histocompatibility Complex (MHC) Genes: Evolution. Encyclopedia of Life Sciences. John Wiley & Sons, Ltd: Chichester. doi: 10.1002/9780470015902.a0005133.pub2.

Hughes AL, Hughes MK, Howell CY, Nei M. 1997. Natural selection at the class II major histocompatibility complex loci of mammals. In: Infection, polymorphism and evolution. Hamilton WD, Howard JC. (Eds.). Chapman & Hall in association with The Royal Society. Suffolk, Great Britain. p. 89-98.

Hughes AL, Nei M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature, 335(6186): 167–170.

Hughes AL, Nei M. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proceedings of the National Academy of Sciences of the United States of America, 86(3): 958-62.

Hughes AL, Yeager M. 1998. Natural selection at major histocompatibility complex loci of vertebrates. Annual Review of Genetics, 32: 415-435.

IUCN (2010) IUCN Red List of Threatened Species. Version 2010.1. http://www.iucnredlist.org. Last accessed : 19 June 2010

Jalil MF, Cable J, Sinyor J, Lackman-Ancrenaz I, Ancrenaz M, Bruford MW, *et al.*. 2008. Riverine effects on mitochondrial structure of Bornean orangutans (*Pongo pygmaeus*) at two spatial scales. Molecular Ecology, 17(12): 2898-2909.

Jalil MFB. 2007. Comparative phylogeography of three primate species in the Lower Kinabatangan Wildlife Sanctuary, Sabah, Malaysia. Unpublished Thesis submitted for the degree of Doctor of Philosophy. School of Biosciences, Cardiff University. pp. 201.

Jarne P, Lagoda PJ. 1996. Microsatellites, from molecules to populations and back. Trends in Ecology & Evolution, 11(10): 424-429.

Johnson MB, Clifford SL, Goossens B, Nyakaana S, Curran B, White LJ, Wickings EJ, Bruford MW. 2007. Complex phylogeographic history of central African forest elephants and its implications for taxonomy. BMC Evolutionary Biology, 7: 244.

Johnson PC, Haydon DT. 2007. Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. Genetics, 175(2): 827-842.

Johnson PC, Haydon DT. 2009. Software for quantifying and simulating microsatellite genotyping error. Bioinformatics and Biology Insights, 1: 71-75.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Mammalian protein metabolism. Munro HN. (Ed.). Academic, New York, p. 21–132.

Kawabe M, Mano T. 1972. Ecology and behavior of the wild proboscis monkey, *Nasalis larvatus* (Wurmb) in Sabah, Malaysia. Primates, 13: 213-228.

Kawamoto Y, Kawamoto S, Matsubayashi K, Nozawa K, Watanabe T, Stanley M, Perwitasari-Farajallah D. 2007. Genetic diversity of longtail macaques (*Macaca fascicularis*) on the island of Mauritius: an assessment of nuclear and mitochondrial DNA polymorphisms. Journal of Medical Primatology, 37(1): 45-54.

Keet DF, Kriek NPJ, Bengis RG, Grobler DG, Michel A. 2000. The rise and fall of tuberculosis in a free–ranging chacma baboon troop in the Krueger National Park. Onderstepoort Journal of Veterinary Research, 67: 115-122.

Kikuchi T, Hara M, Terao K. 2007. Development of a microsatellite marker set applicable to genome-wide screening of cynomolgus monkeys (*Macaca fascicularis*). Primates, 48: 140-46.

Kirkpatrick RC. 2007. The Asian colobines: diversity among leaf-eating monkeys. In: Primates in perspective. Capmpell CJ, Fuentes A, MacKinnon KC, Panger M, Bearder SK. (Eds.). Oxford University Press, New York. p. 186-200

Klein J. 1986. Natural History of the Major Histocompatibility Complex, 1st edn. John Wiley & Sons: New York, Chichester, Brisbane, Toronto, Singapore

Knapp LA, Cadavid LF, Watkins DI. 1998. The MHC-E locus is the most well conserved of all known primate class I histocompatibility genes. The Journal of Immunology, 160: 189-196.

Knapp LA. 2005a. The ABCs of MHC. Evolutionary Anthropology, 14(1): 28-37.

Knapp LA. 2005b. Denaturing gradient gel electrophoresis and its use in the detection of major histocompatibility complex polymorphism. Tissue Antigens, 65(3): 211-219.

Knapp LA. 2005c. Facts, faeces and setting standards for the study of MHC genes using noninvasive samples. Molecular Ecology, 14(5): 1597-1599.

Kohn MH, Murphy WJ, Ostrander EA, Wayne RK. 2006. Genomics and conservation genetics. Trends in Ecology and Evolution, 21(11): 629-637.

Kohn MH, Wayne RK. 1997. Facts from feces revisited. Trends in Ecology and Evolution, 12: 223-227.

Kriener K, O'hUigin C, Tichy H, Klein J. 2000. Convergent evolution of major histocompatibility complex molecules in humans and New World monkeys. Immunogenetics, 51(3): 169-178.

Kumar S, Hedges SB. 1998. A molecular timescale for vertebrate evolution. Nature, 392: 917-920

Lafferty KD, Gerber LR. 2002. Good medicine for conservation biology: the intersection of epidemiology and conservation theory. Conservation Biology, 16: 593-604.

Lafont BA, Buckler-White A, Plishka R, Buckler C, Martin MA. 2003. Characterization of pig-tailed macaque classical MHC class I genes: implications for MHC evolution and antigen presentation in macaques. Journal of Immunology, 171(2): 875-885.

Langefors A, Lohm J, Grahn M, Andersen O, von Schantz. 2001. Association between major histocompatibility complex class IIB alleles and resistance to *Aeromonas salmonicida* in Atlantic salmon. Proceedings of the Royal Society of London B, 268: 479-485.

Latch EK, Dharmarajan D, Glaubitz JC, Rhodes OE. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics, 7: 295–302.

Lathuillière M, Ménard N, Gautier-Hion A, Crouau-Roy B. 2001. Testing the reliability of noninvasive genetic sampling by comparing analysis of blood and fecal samples in Barbary macaques (*Macaca sylvanus*). American Journal of Primatology, 55: 151-158.

Launhardt K, Epplen C, Epplen JT, Winkler P. 1998. Amplification of microsatellites adapted from human systems in faecal DNA of wild Hanuman langurs (*Presbytis entellus*). Electrophoresis, 19: 1356-1361.

Lee H, Li W, Lee A, Rodine P, Graham RR, Ortmann WA, *et al.*. 2006. Microsatellite typing for DRB1 alleles: application to the analysis of HLA associations with rheumatoid arthritis. Genes and Immunity, 7(7): 533-543.

Leendertz F, Pauli G, Maetz-Rensing K, Boardman W, Nunn C, Eellerbrok H, Jensen SA, Junglen S, Boesch C. 2006. Pathogens as drivers of population declines: The importance of systematic monitoring in great apes and other threatened mammals. Biological Conservation, 131(2); 325-337.

Leendertz FH, Elerbrok H, Boesch C, Couacy-Hymann E, Mätz-Rensig K, Hakenbeck R, Gergmann C, Abaza P, Junglen S, Moebious Y, Vigilant L, Formenty P, Pauli G. 2004. Anthrax kills wild chimpanzees in a tropical rainforest. Nature, 430: 451-452.

Lenz TL, Wells K, Pfeiffer M, Sommer S. 2009. Diverse MHC IIB allele repertoire increases parasite resistance and body condition in the Long-tailed giant rat (*Leopoldamys sabanus*). BMC Evolutionary Biology, 9: 269

Lenz TL, Wells K, Pfeiffer M, Sommer S. 2009. Diverse MHC IIB allele repertoire increases parasite resistance and body condition in the Long-tailed giant rat (*Leopoldamys sabanus*). BMC evolutionary biology, 9: 269.

Leroy EM, Rouquet P, Formentry P, Souquière S, Kilbourne AM, Froment JM, Bermejo M, Smit S, Karesh W, Swanepoel R, Zaki SR, Rollin PE. 2004. Multiple Ebola Virus Transmission Events and Rapid Decline of Central African Wildlife. Science, 303: 387-390.

Leuchte N, Berry N, Köhler B, Almond N, LeGrand R, Thorstensson R, *et al.*. 2004. *MhcDRB*-sequences from cynomolgus macaques (*Macaca fascicularis*) of different origin. Tissue Antigens, 63(6): 529-37.

Li Y, Korol AB, Fahima T, Beiles A, Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology, 11(12): 2453-2465.

Little KA. 2003. Genetic relatedness and socioecology of wild Hanuman langur monkeys. PhD Thesis, University College London, University of London.

Liu W, Li Y, Learn GH, Rudicell RS, Robertson JD, Keele BF, *et al.*. 2010. Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. Nature, 467: 420-425.

Liu Z, Ren B, Wu R, Zhao L, Hao Y, Wang B, *et al.*. 2009. The effect of landscape features on population genetic structure in Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) implies an anthropogenic genetic discontinuity. Molecular Ecology, 18(18); 3831-3846.

Long JL. 2003. Introduced mammals of the world: their history, distribution and influence. Wallingford, Oxon: CABI Publishing. pp. 589.

Lowe S, Browne M, Boudjelas S, De Poorter M. 2000. 100 of the World's Worst Invasive Alien Species A selection from the Global Invasive Species Database. Published by The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN), p. 12.

Luikart G, England PR. 1999. Statistical analysis of microsatellite DNA data. Trends in Ecology and Evolution, 7: 253-256.

Lukas D, Bradley BJ, Nsubuga AM, Doran-Sheehy D, Robbins MM, Vigilant L. 2004. Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. Molecular Ecology, 13: 3389-3402.

MacDonald DW. 1992. Cause of wild dog deaths. Nature, 360: 633-34.

Majail J, Webber DA. 2006. Human dimension in conservation works in the Lower Kinabatangan: Sharing PFW's experience. Working paper presented in the Fourth Sabah and Sarawak Environmental Convention, 5 – 7 September 2006. Retrieved 6 February 2007, from: http://www.sabah.gov.my/jpas/news/Conv06/conv06.html

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18: 189–197.

Margolis L, Esch GW, Holmes JC, Kuris AM & Schad GA. 1982. The use of ecological terms in parasitology (report of an ad hoc committee of the American Society of Parasitologists). The Journal of Parasitology, 68(1): 131–133.

Marsh SG, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, *et al.*. 2005 Nomenclature for Factors of the HLA System, 2004. Human Immunology, 66: 571–636.

Mathew D. 2004. WWF-Malaysia Programmes and Initiatives on IWRM Working paper presented at Malaysian water week 2004. Retrieved 12 February, 2007 from: http://www.water.gov.my/web/water%20forum//Malaysian%20Water%20Forum/Status%20Report%202/WWF%20%20Malaysia%20Programmes%20and%20Initiatives%20on%20IWRM.pdf

Matsubayashi K, Gotoh S, Kawamoto Y, Watanabe T, Nozawa K, Takasaka M, *et al.*. 1992. Clinical examinations on crab-eating macaques in Mauritius. Primates, 33(2): 281–288.

Matsuda I, Tuuga A, Higashi S. 2009. Ranging Behavior of Proboscis Monkeys in a Riverine Forest with Special Reference to Ranging in Inland Forest. International Journal of Primatology, 30(2): 313-325.

Matsuda I. 2008. Feeding and Ranging Behaviors of Proboscis Monkey (*Nasalis larvatus*) in Sabah, Malaysia. Thesis submitted for the degree of Doctor of Philosophy. School of Environmental Earth Science, Hokkaido University. pp. 112.

May RM. 1988. Conservation and disease. Conservation Biology 2: 28-30.

Mbora DN, Munene E. 2006. Gastrointestinal parasites of critically endangered primates endemic to Tana River, Kenya: Tana River red colobus (*Procolobus rufomitratus*) and crested mangabey (*Cercocebus galeritus*). The Journal of Parasitology, 92(5): 928-932.

McCallum H, Dobson A. 2006. Disease and connectivity. In: Connectivity conservation. Crooks KR, Sanjayan M. (Eds.) Cambridge University Press, Cambridge. p. 479-501.

McKelvey KS, Schwartz MK. 2004. Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. Journal of Wildlife Management, 68: 439-48.

McMorrow J, Talip MA. 2001. Decline of forest area in Sabah, Malaysia: relationship to state policies, land code and land capability. Global Environmental Change, 11: 217–230.

Meijaard E, Nijman V. 2000. Distribution and conservation of the proboscis monkey (*Nasalis larvatus*) in Kalimantan, Indonesia. Biological Conservation, 92: 15-24.

Meyer-Lucht Y, Otten C, Püttker T, Pardini R, Metzger JP, Sommer S. 2010. Variety matters: adaptive genetic diversity and parasite load in two mouse oppossums from the Brazilian Atlantic Forest. Conservation Genetics, 11 (5): 2001 – 2013.

Meyers LA, Bull JJ. 2002. Fighting change with change: adaptive variation in an uncertain world. Trends in Ecology and Evolution, 17: 551-57.

Mills LS, Citta JJ, Lair KP, Schwartz MK, Tallmon DA. 2000. Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. Ecological Applications, 10: 283-294.

Mills LS. 2007. Conservation of Wildlife Populations: Demography, genetics and management. Blackwell Publishing. pp 407

Mittermeier RA, Wallis J, Rylands AB, Ganzhorn JU, Oates JF, Williamson EA, Palacios E, Heymann EW, Kierulff CM, Yongcheng L, Supriatna J, Roos C, Walker S, Cortes-Ortiz L, Schwitzer C. (Eds.). 2009. Primates in peril: the World's 25 most endangered primates 2008-2010. IUCN/SSC Primate Specialist Group (PSG), International Primatological Society (IPS), and Conservation International (CI), Arlington, VA. pp. 84.

Modolo L, Martin RD, Van Schaik CP, Van Noordwijk MA, Krützen M. 2008. When dispersal fails: unexpected genetic separation in Gibraltar macaques (*Macaca sylvanus*). Molecular Ecology, 17: 4027-4038.

Mona S, Crestanello B, Bankhead-Dronnet S, Pecchioli E, Ingrosso S, D'Amelio S, *et al.*. 2008. Disentangling the effects of recombination, selection, and demography on the genetic variation at a major histocompatibility complex class II gene in the alpine chamois. Molecular Ecology, 17: 4053-4067.

Mondol S, Navya R, Athreya V, Sunagar K, Selvaraj VM, Ramakrishnan U, *et al.*. 2009. A panel of microsatellites to individually identify leopards and its application to leopard monitoring in human dominated landscapes. BMC Genetics, 10: 79.

Monteiro L, Bonnemaison D, Vekris A, Petry KG, Bonnet J, Vidal R, *et al.*. 1997. Complex polysaccharides as PCR inhibitors in feces: *Helicobacter pylori* model. Journal of Clinical Microbiology, 35(4): 995–998.

Morin PA, Chambers KE, Boesch C, Vigilant L. 2001. Quantitative polymerase chain reaction analysis of DNA from non-invasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Molecular Ecology, 10: 1835-1844.

Muehlenbein MP, Ancrenaz M. 2009. Minimizing pathogen transmission at primate ecotourism destinations: the need for input from travel medicine. Journal of Travel Medicine, 16: 229-32.

Muehlenbein MP. 2005. Parasitological analyses of the male chimpanzees (*Pan troglodytes schweinfurthii*) at Ngogo, Kibale National Park, Uganda. American Journal of Primatology, 65(2): 167-179.

Murai T, Mohamed M, Bernard H, Mahedi PA, Saburi R, Higashi S. 2007. Female transfer between one-male groups of proboscis monkeys (*Nasalis larvatus*). Primates, 48: 117-121.

Murai T. 2004. Social behaviors of all-male proboscis monkeys when joined by females. Ecological Research, 19(4): 451-454.

Murai T. 2006. Mating behaviors of the proboscis monkey (*Nasalis larvatus*). American Journal of Primatology, 68:832-837.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic acids research, 8(19): 4321-4325.

Natural History Museum Host-Parasite Database. Retrieved 26 July 2010, from: http://www.nhm.ac.uk/research-curation/research/projects/host-parasites/database/index.jsp

Nei M, Gojobori T. 1986 Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution, 3: 418–426.

Nsubuga AM, Robbins MM, Roeder AD, Morin PA, Boesch C, Vigilant L. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular ecology, 13(7): 2089-2094.

Nunn CL, Altizer S, Jones KE, Sechrest W. 2003. Comparative tests of parasite species richness in primates. American Naturalist, 162(5): 597–614.

Nunn CL, Altizer S. 2005. The Global Mammal Parasite Database: an online resource for infectious disease records in wild primates. Evolutionary Anthropology, 14: 1-2.

Nunn CL, Altizer S. 2006. Infectious diseases in primates: Behavior, ecology and evolution. Oxford Series in Ecology and Evolution. Oxford University Press, UK. pp. 384.

O'Connor SL, Blasky AJ, Pendley CJ, Becker EA, Wiseman RW, Karl JA, *et al.*. 2007. Comprehensive characterization of MHC class II haplotypes in Mauritian cynomolgus macaques. Immunogenetics, 59(6): 449–462.

Oppelt C, Starkloff A, Rausch P, Von Holst D, Rödel HG. 2010. Major histocompatibility complex variation and age-specific endoparasite load in subadult European rabbits. Molecular Ecology: 4155-4167.

Oppelt C, Starkloff A, Rausch P, von Holst D, Rödel HG. 2010. Major histocompatibility complex variation and age-specific endoparasite load in subadult European rabbits. Molecular ecology, 4155-4167.

Packer C, Altizer S, Appel M, Brown E, Martenson J, O'Brien SJ, Roelke-Parker M, Hofmann-Lehmann R and Lutz H. 1999. Viruses of the Serengeti: patterns of infection and mortality in African lions. Journal of Animal Ecology, 68:1161–1178.

Page SL, Goodman M. 2001. Catarrhine phylogeny: noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. Molecular Phylogenetics and Evolution, 18(1): 14-25.

Pan D, Li Y, Hu H, Meng S, Men Z, Fu Y, *et al.*. 2005. Microsatellite polymorphisms of Sichuan golden monkeys. Chinese Science Bulletin, 50(24): 2850–2855.

Pang C, Prudente C, Giong D, Moey G. 2002. Kinabatangan Wildlife Sanctuary Gazetted. World Wildlife Fund (WWF) – Malaysia. Retrieved 12 February, 2007 from: http://www.wwfmalaysia.org/PFW/news/news1.htm

Pang C, Prudente C, Giong D, Moey G. 2005. Kinabatangan Gazettement: A win for Sabahans and Wildlife. World Wildlife Fund (WWF) – Malaysia. Retrieved 12 February, 2007 from: http://www.wwfmalaysia.org/PFW/news/news050912.htm

Pang C, Prudente C. 2003. A vision for the future: "Kinabatangan, a corridor of life". World Wildlife Fund (WWF) – Malaysia. Retrieved 12 February, 2007 from: http://www.wwfmalaysia.org/PFW/Vision%20for%20the%20future.htm

Paterson S, Wilson K, Pemberton JM. 1998. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (*Ovis aries L.*). Proceedings of the National Academy of Sciences USA, 95: 3714-3719.

Payne J. 1989. A tourism feasibility study for the proposed Kinabatangan Wildlife Sanctuary. A report submitted to the Ministry of Tourism and Environmental Development, Sabah. Final Project Report, WWF Project No. 158/89. WWF Malaysia. Kuala Lumpur, Malaysia. pp. 51

Pearse DE, Crandall KA. 2004. Beyond FST: Analysis of population genetic data for conservation. Conservation Genetics, 5: 585–602.

Pedersen AB, Jones KE, Nunn CL, Altizer S. 2007. Infectious diseases and extinction risk in wild mammals. Conservation Biology, 21(5): 1269-1279.

Penedo CMT, Bontrop RE, Heijmans CMC, Otting N, Noort R, Rouweler AJM, de Groot N, de Groot NJ, Ward T, Doxiadis GGM. 2005. Microsatellite typing of the rhesus macaque MHC region. Immunogenetics, 57: 198-209.

Penn DJ, Ilmonen P. 2005. Major Histocompatibility Complex (MHC). Encyclopedia of Life Sciences. John Wiley & Sons, Ltd: Chichester. doi: 10.1038/npg.els.0003986.

Penn DJ, Potts WK. 1999. The evolution of mating preferences and major histocompatiblity complex genes. The American Naturalist, 153: 145-164.

Penn DJ. 2002. The scent of genetic compatibility: sexual selection and the major histocompatibility complex. Ethology, 108: 1-21.

Pérez-Ponce De León G, Garcia-Prieto L. 2001. Los parásitos en el contexto de la biodiversidad y la conservación. Biodiversitas, 34: 11–15.

Perwitasari-Farajallah D, Kyes RC, Iskandar E. 2010. Microsatellite DNA polymorphisms for colony management of long-tailed macaques (*Macaca fascicularis*) population on the Tinjil Island. Biodiversitas, 11(2): 55–58.

Pezzulo. Website "Interactive Statistical Calculations". Last accessed 28 July 2010, from: http://StatPages.org/confint.html.

Pfaff CL, Parra EJ, Bonilla C, *et al.*. 2001. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. American Journal of Human Genetics, 68: 198-207.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. Heredity, 96: 7-21.

Piggott MP, Banks SC, Stone N, Banffy C, Taylor AC. 2006. Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. Molecular Ecology, 15(1): 81-91.

Piggott MP, Taylor AC. 2003. Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildlife Research, 30(1): 1.

Piggott MP. 2004. Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildlife Research, 31(5): 485.

Pompanon F, Bonin A, Bellemain E, Taberlet P. 2005. Genotyping errors: causes, consequences and solutions. Nature Reviews, 6: 847-859.

Porebski S, Bailey LG & Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Molecular Biology Reporter, 15(1): 8-15.

Prichard R, Tait A. 2001. The role of molecular biology in veterinary parasitology. Veterinary Parasitology, 98(1-3): 169-194.

Primmer CR, Moller AP, Ellegren H. 1996. A wide-range survey of cross-species microsatiellite amplification in birds. Molecular Ecology, 5(3): 365-378.

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

Pritchard JK, Wen X, Falush D. 2010. Documentation for STRUCTURE software: Version 2.3. pp. 39.

Quéméré E, Louis EE, Ribéron A, Chikhi L, Crouau-Roy B. 2009. Non-invasive conservation genetics of the critically endangered golden-crowned sifaka (*Propithecus tattersalli*): high diversity and significant genetic differentiation over a small range. Conservation Genetics, 11(3): 675-687.

R Development Core Team. 2010. R: A Language and Environment for StatisticalComputing. R Foundation for Statistical Computing,Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org

Rajendran S, Saseendran PC, Subramanian H, Chitra R, Yuvaraj N. 2004. Gastro-intestinal parasitic infection in Nilgiri Langur (*Semnopithecus johnii*) at Kalakkad-Mundanthurai Tiger Reserve, Tamil Nadu. Zoos' Print Journal, 19(4): 1454.

Ralls K, Pilgrim KL, White PJ, Paxinos EE, Schwartz MK, Fleischer RC. 2001. Kinship, social relationships, and den sharing in kit foxes. Journal of Mammalogy, 82(3): 858-866.

Randi E, Tabarroni C, Rimondi S, Lucchini V, Sfougaris A. 2003. Phylogeography of the rock partridge (*Alectoris graeca*). Molecular Ecology, 12(8): 2201-2214.

Rautner M, Hardiono M, Alfred RJ. 2005. Borneo: Treasure Island at Risk. Status of forest, wildlife and related threats on the island of Borneo. World Wildlife Fund Germany. Frankfurt am Main, Germany. pp. 78.

Raveendran M, Harris RA, Milosavljevic A, Johnson Z, Shelledy, Cameron J, Rogers J. 2006. Designing new microsatellite markers for linkage and population genetic analyses in rhesus macaques and other nonhuman primates. Genomics, 88: 706-710.

Raymond M, Rousset F. 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Eucumenicism. J Hered, 86: 248-249

Reed DH. 2004. Extinction risk in fragmented habitats. Animal conservation, 7: 181–91.

Regip J, Alfred RJ, Indran R. 2004. Integrated river basin management planning for the Kinabatangan catchment, Sabah: Approach and strategy. Working paper presented at the Malaysian Water Week, 7 – 12 June 2004. Retrieved 7 February 2007, from: http://www.water.gov.my/web/water%20forum//Malaysian%20Water%20Forum/Review%20Report%20Sector%202/Integrated%20River%20Basin%20Management%20Planning%20for%20Kanabatangan%20Catchment%20Sabah,%20%20Approach%20and%20Strate.pdf.

Richard AF. 1985. Sympatry, competition, and the niche. In: Primates in nature. Richard AF. (Ed.) W.H. Freeman and Co.

Robinson J, Mistry K, McWilliam H, Lopez R, Marsh SG. 2010. IPD--the Immuno Polymorphism Database. Nucleic Acids Research, 38(Database issue): D863-9.

Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SGE. 2003. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Research, 31: 311-14.

Robinson J, Waller MJ, Stoehr P, Marsh SG. 2005. IPD--the Immuno Polymorphism Database. Nucleic Acids Research, 33(Database issue): D523-6.

Roeder AD, Bonhomme M, Heijmans C, Bruford MW, Crouau-Roy B, Doxiadis G, Otting N, INPRIMAT Consortium. 2009. A large panel of microsatellite markers for genetic studies in the infra-order catarrhini. Folia Primatologica, 80 (2):63-9.

Roon DA, Waits LP, Kendall KC. 2005. A simulation test of the effectiveness of several methods for error-checking non-invasive genetic data. Animal Conservation, 8(2): 203-215.

Rothman JM, Pell AN, Bowman DD. 2009. How does diet quality affect the parasite ecology of mountain gorillas? In: Primate parasite ecology. The dynamics and study of host-parasite relationships. Huffman MA, Chapman CA. (Eds.). Cambridge University Press, Cambridge. p. 441-462.

Rousset F. 2008. GENEPOP'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources, 8, 103-106.

Salgado-Lynn M, Stanton DW, Sakong R, Cable J, Goossens B, Bruford MW. 2010. Microsatellite markers for the proboscis monkey (*Nasalis larvatus*). Conservation Genetics Resources, 2: 159-163.

Salo J, Kalliola R, Häkkinen I, Mäkinen Y, Niemelä P, Puhakka M, *et al.*. 1986. River dynamics and the diversity of Amazon lowland forest. Nature, 322(6076): 254–258.

Schad J, Ganzhorn JU, Sommer S. 2005. MHC constitution and parasite burden in the Malagasy mouse lemur, *Microcebus murinus*. Evolution, 59: 439-50.

Schad J, Sommer S, Ganzhorn JU. 2004. MHC variability of a small lemur in the littoral forest fragments of southeastern Madagascar. Conservation genetics, 5: 299-309.

Schmid-Hempel P, Koella JC. 1994. Variability and its implications for host-parasite interactions. Parasitology Today, 10: 98-102.

Schwartz MK, McKelvey KS. 2008. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conservation Genetics, 10(2), 441-452.

Schwensow N, Fietz J, Dausmann KH, Sommer S. 2007. Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. Heredity, 99: 265-277.

Scott ME. 1988. The impact of infection and disease on animal populations: implications for conservation biology. Conservation Biology 2: 40-56.

Seabra SG, Quartau JA, Bruford MW. 2009. Spatio-temporal genetic variation in sympatric and allopatric Mediterranean Cicada species (*Hemiptera, Cicadidae*). Biological Journal of the Linnean Society, 96: 249-265.

Sha JC, Bernard H, Nathan S. 2008. Status and conservation of proboscis monkeys (*Nasalis larvatus*) in Sabah, East Malaysia. Primate Conservation, 23: 107–120.

Silva MC, Edwards SV. 2009. Structure and evolution of a new avian MHC class II B gene in a sub-Antarctic seabird, the thin-billed prion (Procellariiformes: *Pachyptila belcheri*). Journal of Molecular Evolution, 68(3): 279-91.

Singh B, Cox-Singh J. 2001. Parasites that cause problems in Malaysia: soil-transmitted helminths and malaria parasites. Trends in parasitology, 17(12): 597-600.

Singh B, Kim Sung L, Matusop A, Radhakrishnan A, Shamsul SS, Cox-Singh J, Thomas A, Conway DJ. 2004. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. Lancet 363: 1017-24.

Slatkin M. 2008. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics, 9(6): 477–485.

Slierendregt BL, van Noort JT, Bakas RM, Otting N, Jonker M, Bontrop RE. 1992. Evolutionary stability of trans-species major histocompatibility complex class II DRB lineages in humans and rhesus monkeys. Human Immunology, 35: 29–39.

Smith KF, Acevedo-Whitehouse K, Pedersen AB. 2009. The role of infectious diseases in biological conservation. Animal Conservation, 12(1): 1-12.

Smith KF, Sax DF, Lafferty KD. 2006. Evidence for the role of infectious disease in species extinction and endangerment. Conservation Biology, 20(5): 1349-1357.

Sommer S. 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology, 2: 16-33.

Son VD. 2002. Intestinal Parasites of *Macaca fascicularis* in a Mangrove Forest, Ho Chi Minh City, Vietnam. Laboratory Primate Newsletter, 41(2): 4–5.

Speare R, Latasi FF, Nelesone T, Harmen S, Melrose W, Durrheim D, Heukelbach J. 2006. Prevalence of soil transmitted nematodes on Nukufetau, a remote Pacific island in Tuvalu. BMC Infectious Disease, 6: 110

Spurgin LG, Richardson DS. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. Proceedings of the Royal Society B: Biological Sciences, 277: 979-988.

Steiper ME, Wolfe ND, Karesh WB, Kilbourn AM, Bosi EJ, Ruvolo M. 2006. The phylogenetic and evolutionary history of a novel alpha-globin-type gene in orangutans (*Pongo pygmaeus*). Infection, Genetics and Evolution, 6(4): 277-286.

Stephens M, Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. American Journal of Human Genetics, 73(5): 1162-1169.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics, 68: 978-989.

Stephens M, Smith NJ, Donnelly P. 2004. Documentation for PHASE, version 2.1. Seattle: Department of Statistics, University of Washington. pp. 36.

Stockwell CA, Hendry AP, Kinnison MT. 2003. Contemporary evolution meets conservation biology. Trends in Ecology and Evolution, 18(2): 94-101.

Stoner KE. 1996. Prevalence and Intensity of Intestinal Parasites in Mantled Howling Monkeys (*Alouatta palliata*) in Northeastern Costa Rica: Implications for Conservation Biology. Conservation Biology, 10(2): 539-546.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, *et al.*. 2007. Putting the "landscape" in landscape genetics. Heredity, 98(3): 128-142.

Storfer A. 1999. Gene flow and endangered species translocations: a topic revisited. Biological Conservation, 87(2): 173–180.

Stuart MD, Strier KB. 1995. Primates and parasites: a case for a multidisciplinary approach. International Journal of Primatology, 16(4): 577–593.

Sullivan JT. 1999. Electronic Atlas of Parasitology CD-ROM v 2.02. University of San Francisco, McGraw-Hill.

Supriatna JA, Yanuar M, Wibisono HT, Asinaga R, Sidik I, Iskandar S. 1996. A preliminary survey of Long-tailed and Pig-tailed Macaques (*Macaca fascicularis* and *Macaca nemestrina*) in Lampung, Bengkulu and Jambi Provinces, Southern Sumatra, Indonesia. Tropical Biodiversity 3: 131-139.

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research, 24: 3189-3194.

Taberlet P, Waits LP, Luikart G. 1999. Noninvasive genetic sampling: look before you leap. Trends in Ecology and Evolution, 8: 323-327.

Takahata N, Nei M. 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics, 124: 967-978.

Tamura K, Dudley J, Nei M,'Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution, 24; 1596-1599.

Teichroeb JA, Kutz SJ, Parkar U, Thompson RC, Sicotte P. 2009. Ecology of the gastrointestinal parasites of *Colobus vellerosus* at Boabeng-Fiema, Ghana: possible anthropozoonotic transmission. American Journal of Physical Anthropology, 140(3): 498-507.

Telfer PT, Souquière S, Clifford SL, *et al.*. 2003. Molecular evidence for deep phylogenetic divergence in Mandrillus sphinx. Molecular Ecology, 12: 2019–2024.

Thierry B. 2007. The Macaques: a double-layered social organization. In: Primates in perspective. Capmpell CJ, Fuentes A, MacKinnon KC, Panger M, Bearder SK. (Eds.). Oxford University Press, New York. p. 186-200.

Thompson RC, Kutz SJ, Smith A. 2009. Parasite zoonoses and wildlife: emerging issues. International Journal of Environmental Research and Public Health, 6(2): 678-693.

Thompson RC, Lymbery AJ, Smith A. 2010. Parasites, emerging disease and wildlife conservation. International Journal for Parasitology, 40(10): 1163-1170.

Tollenaere C, Bryja J, Galan M, Cadet P, Deter J, Chaval Y, *et al.*. 2008. Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. Journal of evolutionary biology, 21(5): 1307-1320.

Tollenaere C, Bryja J, Galan M, Cadet P, Deter J, *et al.*. 2008 Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. Journal of Evolutionary Biology, 21: 1307-1320.

Trachtenberg E, Korber B, Sollars C, Kepler TB, Hraber PT, Hayes E, Funkhouser R, Fugate M, Theiler J, Hsu YS, Kunstman K, Wu S, Phair J, Erlich H, Wolkinsky S. 2003. Advantage of rare HLA supertype in HIV disease progression. Nature Medicine, 9: 928-35.

Trejo-Macías G, Estrada A & Mosqueda Cabrera MÁ. 2007. Survey of Helminth Parasites in Populations of *Alouatta palliata mexicana* and *A. pigra* in Continuous and in Fragmented Habitat in Southern Mexico. International Journal of Primatology, 28(4): 931-945.

Turner SM, Chaves-Campos J, DeWoody JA. 2009. Parental relatedness and major histocompatibility effects on early embryo survivorship in Atlantic salmon. Genetica, 137(1): 99-109.

UNEP-WCMC. UNEP-WCMC Species Database: CITES-Listed Species. United Nations Environment Programme (UNEP) – World Conservation Monitoring Centre (WCMC). http://sea.unep-wcmc.org. Accessed 19 June 2010

Valière N, Berthier P, Mouchiroud D, Pontier D. 2002. Gemini: Software for Testing the Effects of Genotyping Errors and Multitubes Approach for Individual Identification. Molecular Ecology Notes, 2(1): 83-86.

Valière N, Bonenfant C, Toïgo C, Luikart G, Gaillard JM, Klein F. 2007. Importance of a pilot study for non-invasive genetic sampling: genotyping errors and population size estimation in red deer. Conservation Genetics, 8: 69-78.

Vallet D, Petit EJ, Gatti S, Levréro F, Ménard N. 2007. A new 2CTAB/PCI method improves DNA amplification success from faeces of Mediterranean (*Barbary macaques*) and tropical (lowland gorillas) primates. Conservation Genetics, 9(3): 677-680.

Vallinoto M, Araripe J, Rego PS, Tagliaro CH, Sampaio I, Schneider H, *et al.*. 2006. Tocantins river as an effective barrier to gene flow in *Saguinus niger* populations. Genetics and Molecular Biology, 29(2): 215-219.

van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4(3): 535-538.

van Oosterhout C, Joyce DA, Cummings SM, Blais J, Barson NJ, Ramnarine IW, *et al.*. 2006. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). Evolution, 60(12): 2562-2574.

van Oosterhout C. 2009. A new theory of MHC evolution: beyond selection on the immune genes. Proceedings of the Royal Society B: Biological Sciences, 276: 657-665.

van Schaik CP, van Amerongen AJJ, van Noordwijk MA. 1996. Riverine refuging in Sumatran long tailed macaques (*Macaca fascicularis*). In: Evolution and ecology of macaque societies. Fa J & Lindburg D. (Eds.). Cambridge University Press. Cambridge. UK. p. 160-181.

Vaz J. 1997. The Kinabatangan Floodplain: an introduction. WWF Malaysia, Ministry of tourism and environment, Sabah. Sabah, Malaysia. pp. 66.

Vernesi C, Bruford MW, Bertorelle G, Pecchioli E, Rizzoli A, Hauffe HC, *et al.*. 2008. Where's the conservation in conservation genetics? Conservation Biology, 22(3): 802-804.

Vogel TU, Evans DT, Urvater JA, O'Connor DH, Hughes AL, Watkins DI. 1999. Major histocompatibility complex class I genes in primates: co-evolution with pathogens. Immunological Reviews, 167: 327.37.

von Segesser F, Menard N, Gaci B, Martin RD. 1999. Genetic differentiation within and between isolated Algerian subpopulations of Barbary macaques (*Macaca sylvanus*): evidence from microsatellites. Molecular Ecology, 8(3): 433-442.

Waits LP, Paetkau D. 2005. Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. Journal of Wildlife Management, 69: 1419-1433.

Wallis J, Lee DR. 1999. Primate conservation: The prevention of disease transmission. International Journal of Primatology, 20: 803-826.

Wallis J. 2000. Prevention of disease transmission in primate conservation. Annals of the New York Academy of Sciences, 916: 691-694.

Walsh PD, Abernethy KA, Bermejo M, Beyers R, De Wascher P, Akou ME, Huijbregts B, Mambounga DI, Toham AK, Kilbourn AM, Lahm SA, Latour S, Maisels F, Mbina C, Mihindou Y, Oblang SN, Effa EN Strakey MP, Teifer P, Thibault M, Tutin CEG, White LJT, Wilkie DS. 2003. Catastrophic ape decline in Western Equatorial Africa. Nature, 422: 611-614.

Waples RS, Gaggiotti O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular ecology, 15(6): 1419-1439.

Warren KS, Verschoor EJ, Langenhuijzen S, Heriyanto, Swan RA, Vigilant L, *et al.* 2001. Speciation and intrasubspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus*. Molecular Biology and Evolution, 18(4): 472-480.

Wayne RK, Morin PA. 2004. Conservation genetics in the new molecular age. Frontiers in Ecology and the Environment, 2(2): 89–97.

Wedekind C, Walker M, Little TJ. 2005. The course of malaria in mice: Major Histocompatibility Complex (MHC) effects, but no general MHC heterozygote advantage in single-strain infections. Genetics, 170: 1427-30.

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

Weisser WW. 2001. The effects of predation on dispersal. In: Dispersal. Clobert J, Danchin E, Dhondt AA, Nichols JD (Eds.). Oxford University Press, New York. pp. 180-188.

Westerdahl H, Waldenström J, Hansson B, Hasselquist D, von Schantz T, Bensch S. 2005. Associations between malaria and MHC genes in a migratory songbird. Proceedings of the Royal Society B, 272: 1511-1518.

Weyher AH, Ross C, Semple S. 2006. Gastrointestinal Parasites in Crop Raiding and Wild Foraging *Papio anubis* in Nigeria. International Journal of Primatology, 27(6): 1519-1534.

Weyher AH. 2009. Crop raiding: the influence of behavioral and nutritional changes on primate-parasite relationships. In: Primate parasite ecology. The dynamics and study of host-parasite relationships. Huffman MA, Chapman CA. (Eds). Cambridge University Press, Cambridge. p. 409-422.

Williams ES, Thorne ET, Appel MJ, Belitsky DW. 1988. Canine distemper in black-footed ferrets (*Mustela nigripes*) from Wyoming. Journal of Wildlife Diseases, 24: 385-398.

Williams ES, Yuill T, Artois M, Fischer J, Haigh SA. 2002. Emerging infectious diseases in wildlife. Revue Scientifique et Technique, Office International des Épizooties, 21: 139-157.

Wolfe ND, Escalante AA, Karesh WB, Kilbourn A, Spielman A, Lal AA. 1998. Wild primate populations in emerging infectious disease research: the missing link? Emerging Infectious Diseases, 4: 149–158.

Wolfheim JH. 1983. Primates of the world: distribution, abundance and conservation. Seattle. University of Washington Press.

Woodford MH, Butynski TM, Karesh WB. 2002. Habituating the great apes: the disease risks. Oryx, 36: 153-160.

Woodroffe R. 1999. Managing disease threats to wild mammals. Animal conservation, 2: 185-193.

World Wildlife Fund (WWF) - Asia/Pacific. 2005. Endangered species found in the Kinabatangan. Retrieved 15 October, 2006, from: http://www.panda.org/about_wwf/where_we_work/asia_pacific/where/malaysia/kinabatangan/species/index.cfm

World Wildlife Fund (WWF) – Malaysia. 2003. Managing rivers wisely: Kinabatangan case study. Retrieved 2 February, 2007 from: http://www.panda.org/about_wwf/what_we_do/freshwater/our_solutions/rivers/irbm/cases/kinabatangan_river_case_study/index.cfm

Xia X, Xie Z. 2001. DAMBE: Data Analysis in Molecular Biology. Journal of Heredity, 92: 371-373.

Xia X. 2000. Data analysis in molecular biology and evolution. Kluwer Academic publishers, Boston. pp. 276.

Yeager CP. 1995. Does intraspecific variation in social systems explain reported differences in the social structure of the proboscis monkey (*Nasalis larvatus*)? Primates 36: 575-582.

Young AG, Clarke GM (Eds.). 2000. Genetics, demography and viability of fragmented populations. Cambridge University Press, Cambridge, UK. pp. 438.

Zarlenga DS, Chute MB, Gasbarre LC, Boyd PC. 2001. A multiplex PCR assay for differentiating economically important gastrointestinal nematodes of cattle. Veterinary Parasitology, 97: 199-209.

Zarlenga DS, Higgins J. 2001. PCR as a diagnostic and quantitative technique in veterinary parasitology. Veterinary Parasitology, 101(3-4): 215-230.

Zhan XJ, Zhang ZJ, Wu H, Goossens, B, Li M, Jiang SW, *et al.* 2007. Molecular analysis of dispersal in giant pandas. Molecular Ecology, 16(18): 3792-3800.

Zhan XJ, Zheng X, Bruford MW, Wei F, Tao Y. 2009. A new method for quantifying genotyping errors for noninvasive genetic studies. Conservation Genetics, 11(4): 1567-1571.

Zhang B, Fang SG, Xi YM. 2006. Major Histocompatibility Complex variation in the endangered crested ibis *Nipponia nippon* and implications for reintroduction. Biochemical Genetics, 44: 113-23.

Zhang DX, Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. Molecular Ecology, 12(3): 563–584.

APPENDIX ONE

Description of samples collected for proboscis monkeys (PM – *Nasalis larvatus*) and long-tailed macaques (Mafa – *Macaca fascicularis*) of the Lower Kinabatangan Wildlife Sanctuary. Additional samples are included from a few orangutans (Popy- *Pongo pygmeaus*), pig-tailed macaques (Mane – *Macaca nemestrina*), and elephants (Elma – *Elephas maximus borneensis*). Each sample has a unique identifier 'name' with GPS coordinates of the collection site on the day of collection. The weight of the faecal sample (g) is cited along with the conversion factor (equivalent to the total volume of flotation fluid divided by the volume counted in the McMaster chamber multiplied times the weight of the faeces, i.e. 15 ml / 0.3 ml x weight) to calculate the eggs per gram (epg) of faeces in each sample (multiply the number of eggs counted times the conversion factor). DNA was extracted from most samples twice ('duplicate') and date of extraction is also given. The storage details of these samples are presented and whether mtDNA was amplified from one or both replicates. Finally, it is indicated whether or not a sucrose floatation was performed for detection of parasite eggs in the faecal samples.

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-1-1	N5°31.467', E118°17.774'	231007	2.18	23.58	Duplicate	05/11/07	Box 2 Nala	Y (Both)	Y
PM 3-1-2	N5°31.467', E118°17.774'	231007	1.54	32.47	Duplicate	05/11/07	Box 2 Nala	Y (Few B)	Y
PM 3-1-3	N5°31.467', E118°17.774'	231007	0.46	108.70	Duplicate	05/11/07	Box 2 Nala	Y (Only B)	Y
PM 3-1-4	N5°31.467', E118°17.774'	231007	0.48	104.17	Duplicate	05/11/07	Box 2 Nala	N (Few B)	Y
PM 3-1-5	N5°31.467', E118°17.774'	231007	0.85	58.82	Duplicate	05/11/07	Box 2 Nala	Y (Both)	Y
PM 3-1-6	N5°31.467', E118°17.774'	231007	1.75	28.57	Duplicate	05/11/07	Box 2 Nala	Y (Both)	Y
PM 3-1-7	N5°31.467', E118°17.774'	231007	2.1	23.81	Duplicate	05/11/07	Box 2 Nala	Y (Both)	Y
PM 2-1-1	N5°33.196', E118°19.371'	231007	0.2	250.00	Duplicate	05/11/07	Box 1 Nala	Y (Both)	Y
PM 2-1-2	N5°33.196', E118°19.371'	231007	0.84	59.52	Duplicate	05/11/07	Box 1 Nala	Y (Both)	Y
PM 2-1-3	N5°33.196', E118°19.371'	231007	3.66	13.66	Duplicate	05/11/07	Box 1 Nala	Y WeirdA	Y
PM 2-1-4	N5°33.196', E118°19.371'	231007	0.82	60.98	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-1-5	N5°33.196', E118°19.371'	231007	0.02	2500.00	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-2-1	N5°33.422', E118°20.188'	231007	1.26	39.68	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-2-2	N5°33.422', E118°20.188'	231007	0.55	90.91	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-2-3	N5°33.422', E118°20.188'	231007	1.77	28.25	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-2-4	N5°33.422', E118°20.188'	231007	2.04	24.51	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
Popy 3-1-1	N5°33.498', E118°19.736'	231007	7.45	6.71	Duplicate	06/11/07	Box 7 Mafa		Y
PM 1-1-1	N5°40.025', E118°23.134'	251007	1.07	46.73	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 1-1-2	N5°40.025', E118°23.134'	251007	1.24	40.32	Duplicate	06/11/07	Box 1 Nala	Y (A?)	Y
PM 1-1-3	N5°40.025', E118°23.134'	251007	0.13	384.62	Duplicate	06/11/07	Box 1 Nala	Y (B?)	Y
PM 1-1-4	N5°40.025', E118°23.134'	251007	1.35	37.04	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 1-1-5	N5°40.025', E118°23.134'	251007	0.74	67.57	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 1-1-6	N5°40.025', E118°23.134'	251007	1.46	34.25	Duplicate	06/11/07	Box 1 Nala	Y (Few A)	Y
PM 1-1-7	N5°40.025', E118°23.134'	251007	1.45	34.48	Duplicate	06/11/07	Box 1 Nala	Y (Few B)	Y
PM 1-1-8	N5°40.025', E118°23.134'	251007	1.23	40.65	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-3-1	N5°39.528', E118°22.098'	251007	0.8	62.50	Duplicate	06/11/07	Box 1 Nala	Y (Both)	Y
PM 2-3-2	N5°39.528', E118°22.098'	251007	0.83	60.24	Duplicate	07/11/07	Box 1 Nala	Y (Both)	Y
PM 2-3-3	N5°39.528', E118°22.098'	251007	0.22	227.27	Duplicate	07/11/07	Box 1 Nala	Y (Both)	Y
PM 2-4-1	N5°39.528', E118°22.098'	251007	0.72	69.44	Duplicate	07/11/07	Box 1 Nala	Y (Both)	Y
PM 2-4-2	N5°39.528', E118°22.098'	251007	0.8	62.50	Duplicate	07/11/07	Box 1 Nala	N	Y
PM 2-4-3	N5°39.528', E118°22.098'	251007	0.59	84.75	Duplicate	07/11/07	Box 2 Nala	Y (Both)	Y
PM 2-4-4	N5°39.528', E118°22.098'	251007	1.46	34.25	Duplicate	07/11/07	Box 2 Nala	Y (Both)	Y
PM 2-4-5	N5°39.528', E118°22.098'	251007	1.54	32.47	Duplicate	07/11/07	Box 2 Nala	Y (Both)	Y
PM 4-1-1	N5°28.859', E118°15.446'	271007			Duplicate	07/11/07	Box 3 Nala	Y (Few A)	N
PM 4-1-2	N5°28.859', E118°15.446'	271007	1.62	30.86	Duplicate	07/11/07	Box 3 Nala	Y (Both)	Y
PM 4-1-3	N5°28.859', E118°15.446'	271007	1.71	29.24	Duplicate	07/11/07	Box 3 Nala	Y (Few A)	Y
Mane	Gomantong Caves	271007	-6.64	-7.53	Duplicate	07/11/07	Box 1 Mafa		Y

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-2-1	N5°33.497', E118°19.668'	121107	0.321	155.76	Duplicate	28/12/07	Box 2 Nala	Y (Only A)	Y
PM 3-3-1	N5°33.526', E118°19.821'	121107	0.291	171.82	Duplicate	26/12/07	Box 2 Nala	Y (Only B)	Y
PM 3-3-2	N5°33.526', E118°19.821'	121107	0.22	227.27	Only A	26/12/07	Box 2 Nala	Y	Y
PM 3-3-3	N5°33.526', E118°19.821'	121107	0.474	105.49	Duplicate	26/12/07	Box 2 Nala	Y (Both)	Y
PM 3-3-4	N5°33.526', E118°19.821'	121107	0.7926	63.08	Duplicate	26/12/07	Box 2 Nala	Y (Few A)	Y
Mane 3-1-1	N5°33.033', E118°20.076'	121107	1.125	44.444444					Y
Mane 3-1-2	N5°33.033', E118°20.076'	121107	0.541	92.421442					Y
Elma 2-1-1	N5°32.719', E118°18.726'	121107	0.408	122.54902					Y
Elma 2-1-2	N5°32.719', E118°18.726'	121107	0.549	91.074681					Y
PM 1-3-1	N5°33.172', E118°20.335'	131107			Duplicate	26/12/07	Box 1 Nala	Y (Both)	Y
Mafa 3-1-1	N5°33.240', E118°19.979'	131107			Duplicate	31/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-2-1	N5°31.789', E118°17.607'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-1-1	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	N (Few A)	Y
Mafa 2-1-2	N5°31.831', E118°17.464'	141107							N
Mafa 2-1-3	N5°31.831', E118°17.464'	141107							N
Mafa 2-1-4	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Few B)	Y
Mafa 2-1-5	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-1-6	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Few A)	Y
Mafa 2-1-7	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Few B)	Y
Mafa 2-1-8	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	Y (Few B)	Y
Mafa 2-1-9	N5°31.831', E118°17.464'	141107			Duplicate	26/12/07	Box 1 Mafa	N (Few A)	Y
Mafa 2-1-10	N5°31.831', E118°17.464'	141107							Y
PM 2-5-1	N5°32.999', E118°18.993'	141107			Duplicate	27/12/07	Box 2 Nala	Y (Both)	Y
PM 3-4-1	N5°33.228', E118°19.536'	141107			Duplicate	26/12/07	Box 2 Nala	Y (Both)	Y
PM 3-4-2	N5°33.228', E118°19.536'	141107			Duplicate	26/12/07	Box 2 Nala	Y (Both)	Y
PM 3-4-3	N5°33.228', E118°19.536'	141107							N
PM 1-2-1	N5°33.134', E118°20.058'	141107			Duplicate	26/12/07	Box 1 Nala	Y (Few 2)	N
PM 1-2-2	N5°33.134', E118°20.058'	141107			Duplicate	26/12/07	Box 1 Nala	Y (Few B)	N
PM 1-2-3	N5°33.134', E118°20.058'	141107							N
PM 1-2-4	N5°33.134', E118°20.058'	141107			Duplicate	26/12/07	Box 1 Nala	Y (Both)	Y
PM 1-2-5	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Few A)	Y
PM 1-2-6	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	N
PM 1-2-7	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	Y
PM 1-2-8	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	Y
PM 1-2-9	N5°33.134', E118°20.058'	141107							N
PM 1-2-10	N5°33.134', E118°20.058'	141107							N
PM 1-2-11	N5°33.134', E118°20.058'	141107							N
PM 1-2-12	N5°33.134', E118°20.058'	141107							N
PM 1-2-13	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	Y
PM 1-2-14	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	Y
PM 1-2-15	N5°33.134', E118°20.058'	141107							N
PM 1-2-16	N5°33.134', E118°20.058'	141107			Duplicate	27/12/07	Box 1 Nala	Y (Both)	N
Mafa 2-2-1	N5°34.201', E118°19.867'	141107	0.16	312.50	Duplicate	26/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 3-3-1	N5°32.119', E118°17.544'	141107	0.36	138.89	Duplicate	26/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-3-1	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-2	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-3	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-4	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-5	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-6	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-7	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-8	N5°35.226', E118°19.665'	151107	0.96	52.08	Duplicate	27/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-3-9	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-10	N5°35.226', E118°19.665'	151107			Duplicate	31/12/07	Box 1 Mafa	Y (Both)	N
Mafa 2-3-11	N5°35.226', E118°19.665'	151107							N
Mafa 2-3-12	N5°35.226', E118°19.665'	151107			Duplicate	31/12/07	Box 1 Mafa	Y (Both)	N
Mafa 2-3-13	N5°35.226', E118°19.665'	151107			Duplicate	27/12/07	Box 1 Mafa	Y (Both)	N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 2-3-14	N5°35.226', E118°19.665'	151107			Duplicate	31/12/07	Box 1 Mafa	Y (Both)	N
Elma 1-1-1	N5°35.042', E118°19.780'	241007							N
PM 2-6-1	N5°33.908', E118°20.012'	151107							N
PM 2-6-2	N5°33.908', E118°20.012'	151107							N
PM 2-6-3	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-4	N5°33.908', E118°20.012'	151107							N
PM 2-6-5	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-6	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-7	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	N	N
PM 2-6-8	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	N	N
PM 2-6-9	N5°33.908', E118°20.012'	151107	1.56	32.05	Duplicate	27/12/07	Box 2 Nala	Y (Only B)	Y
PM 2-6-10	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-11	N5°33.908', E118°20.012'	151107							N
PM 2-6-12	N5°33.908', E118°20.012'	151107							N
PM 2-6-13	N5°33.908', E118°20.012'	151107							N
PM 2-6-14	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-15	N5°33.908', E118°20.012'	151107							N
PM 2-6-16	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-17	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-18	N5°33.908', E118°20.012'	151107							N
PM 2-6-19	N5°33.908', E118°20.012'	151107							N
PM 2-6-20	N5°33.908', E118°20.012'	151107							N
PM 2-6-21	N5°33.908', E118°20.012'	151107							N
PM 2-6-22	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-23	N5°33.908', E118°20.012'	151107							N
PM 2-6-24	N5°33.908', E118°20.012'	151107							N
PM 2-6-25	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	FEW	N
PM 2-6-26	N5°33.908', E118°20.012'	151107							N
PM 2-6-27	N5°33.908', E118°20.012'	151107							N
PM 2-6-28	N5°33.908', E118°20.012'	151107							N
PM 2-6-29	N5°33.908', E118°20.012'	151107							N
PM 2-6-30	N5°33.908', E118°20.012'	151107							N
PM 2-6-31	N5°33.908', E118°20.012'	151107							N
PM 2-6-32	N5°33.908', E118°20.012'	151107							N
PM 2-6-33	N5°33.908', E118°20.012'	151107							N
PM 2-6-34	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-6-35	N5°33.908', E118°20.012'	151107							N
PM 2-6-36	N5°33.908', E118°20.012'	151107							N
PM 2-6-37	N5°33.908', E118°20.012'	151107			Single	03/01/08	Box 2 Nala	Y (Few)	N
PM 2-6-38	N5°33.908', E118°20.012'	151107							N
PM 2-6-39	N5°33.908', E118°20.012'	151107							N
PM 2-6-40	N5°33.908', E118°20.012'	151107	0.86	58.14	Duplicate	27/12/07	Box 2 Nala	Y (Both)	Y
PM 2-7-1	N5°32.999', E118°18.993'	151107							N
PM 2-7-2	N5°32.999', E118°18.993'	151107							N
PM 2-7-3	N5°32.999', E118°18.993'	151107			Single	04/01/08	Box 2 Nala	Y	N
PM 2-7-4	N5°32.999', E118°18.993'	151107							N
PM 2-7-5	N5°32.999', E118°18.993'	151107			Single	03/01/08	Box 2 Nala	Y	N
PM 2-7-6	N5°32.999', E118°18.993'	151107	0.66	75.76	Duplicate	28/12/07	Box 2 Nala	Y (Both)	Y
PM 2-7-7	N5°32.999', E118°18.993'	151107							N
PM 2-7-8	N5°32.999', E118°18.993'	151107			Single	04/01/08	Box 2 Nala	Y	N
PM 2-7-9	N5°32.999', E118°18.993'	151107							N
PM 2-7-10	N5°32.999', E118°18.993'	151107							N
PM 2-7-11	N5°32.999', E118°18.993'	151107							N
PM 2-7-12	N5°32.999', E118°18.993'	151107							N
PM 2-5-2	N5°32.999', E118°18.993'	151107			Single	04/01/08	Box 1 Nala	N	N
PM 2-5-3	N5°32.999', E118°18.993'	151107							N
PM 2-5-4	N5°32.999', E118°18.993'	151107	0.76	65.79	Duplicate	28/12/07	Box 2 Nala	Y (Few A)	Y

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 2-5-5	N5°32.999', E118°18.993'	151107							N
PM 2-5-6	N5°32.999', E118°18.993'	151107							N
PM 2-5-7	N5°32.999', E118°18.993'	151107							N
PM 2-5-8	N5°32.999', E118°18.993'	151107							N
PM 2-5-9	N5°32.999', E118°18.993'	151107							N
PM 2-5-10	N5°32.999', E118°18.993'	151107							N
PM 2-5-11	N5°32.999', E118°18.993'	151107							N
PM 2-5-12	N5°32.999', E118°18.993'	151107			Single	04/01/08	Box 2 Nala	Y	N
Mafa 2-4-1	N5°32.999', E118°18.993'	151107	1.36	36.76	Duplicate	27/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-4-2	N5°32.999', E118°18.993'	151107							N
Mafa 2-4-3	N5°32.999', E118°18.993'	151107							N
Mafa 1-1-1	N5°33.289', E118°20.417'	161107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-4-1	N5°33.124', E118°20.028'	161107	1.26	39.68	Duplicate	28/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 3-4-2	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-3	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-4	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-5	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-6	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-7	N5°33.124', E118°20.028'	161107							N
Mafa 3-4-8	N5°33.124', E118°20.028'	161107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 1-2-1	N5°33.134', E118°20.058'	161107			Single	300609	Box 7 Mafa	Y	N
Mafa 1-2-2	N5°33.134', E118°20.058'	161107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 1-2-3	N5°33.134', E118°20.058'	161107			Single	300609	Box 7 Mafa	Y	N
Mafa 1-2-4	N5°33.134', E118°20.058'	161107	1.56	32.05	Duplicate	28/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 1-2-5	N5°33.134', E118°20.058'	161107			Duplicate	31/12/07	Box 1 Mafa	Y (Only A)	N
Mafa 1-2-6	N5°33.134', E118°20.058'	161107			Single	300609	Box 7 Mafa	N	N
Mafa 1-2-7	N5°33.134', E118°20.058'	161107			Single	02/01/08	Box 1 Mafa	N	N
Mane 1-1-1	N5°33.134', E118°20.058'	161107							N
Mane 1-1-2	N5°33.134', E118°20.058'	161107	0.26	192.31					Y
Mafa 2-4-4	N5°32.999', E118°18.993'	161107	0.46	108.70	Duplicate	27/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 2-4-5	N5°32.999', E118°18.993'	161107							N
Mafa 2-4-6	N5°32.999', E118°18.993'	161107							N
Mafa 2-4-7	N5°32.999', E118°18.993'	161107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 2-4-8	N5°32.999', E118°18.993'	161107							N
Mafa 3-5-1	N5°32.479', E118°18.554'	161107							N
Mafa 3-5-2	N5°32.479', E118°18.554'	161107							N
Mafa 3-5-2	N5°32.479', E118°18.554'	161107							N
Mafa 3-5-4	N5°32.479', E118°18.554'	161107	0.96	52.08	Duplicate	27/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 3-5-5	N5°32.479', E118°18.554'	161107							N
Mafa 3-5-6	N5°32.479', E118°18.554'	161107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 3-5-7	N5°32.479', E118°18.554'	161107							N
PM 3-5-1	N5°31.473', E118°17.767'	181107							N
PM 3-5-2	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 2 Nala	Y	N
PM 3-5-3	N5°31.473', E118°17.767'	181107							N
PM 3-5-4	N5°31.473', E118°17.767'	181107							N
PM 3-5-5	N5°31.473', E118°17.767'	181107							N
PM 3-5-6	N5°31.473', E118°17.767'	181107							N
PM 3-5-7	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 2 Nala	N	N
PM 3-5-8	N5°31.473', E118°17.767'	181107							N
PM 3-5-9	N5°31.473', E118°17.767'	181107							N
PM 3-5-10	N5°31.473', E118°17.767'	181107							N
PM 3-5-11	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 2 Nala	Y	N
PM 3-5-12	N5°31.473', E118°17.767'	181107							N
PM 3-5-13	N5°31.473', E118°17.767'	181107							N
PM 3-5-14	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 2 Nala	N	N
PM 3-5-15	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 2 Nala	Y	N
PM 3-5-16	N5°31.473', E118°17.767'	181107							N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-5-17	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 2 Nala	Y	N
PM 3-5-18	N5°31.473', E118°17.767'	181107	1.46	34.25	Duplicate	28/12/07	Box 2 Nala	Y (Both)	Y
PM 3-5-19	N5°31.473', E118°17.767'	181107							N
PM 3-5-20	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 2 Nala	Y	N
PM 3-5-21	N5°31.473', E118°17.767'	181107							N
PM 3-5-22	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-5-23	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 2 Nala	Y	N
PM 3-5-24	N5°31.473', E118°17.767'	181107							N
PM 3-5-25	N5°31.473', E118°17.767'	181107							N
PM 3-5-26	N5°31.473', E118°17.767'	181107							N
PM 3-5-27	N5°31.473', E118°17.767'	181107							N
PM 3-5-28	N5°31.473', E118°17.767'	181107							N
PM 3-5-29	N5°31.473', E118°17.767'	181107							N
PM 3-5-30	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 2 Nala	Y	N
PM 3-5-31	N5°31.473', E118°17.767'	181107							N
PM 3-5-32	N5°31.473', E118°17.767'	181107							N
PM 3-5-33	N5°31.473', E118°17.767'	181107							N
PM 3-5-34	N5°31.473', E118°17.767'	181107							N
PM 3-5-35	N5°31.473', E118°17.767'	181107							N
PM 3-5-36	N5°31.473', E118°17.767'	181107							N
PM 3-5-37	N5°31.473', E118°17.767'	181107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-5-38	N5°31.473', E118°17.767'	181107							N
PM 3-5-39	N5°31.473', E118°17.767'	181107							N
PM 3-5-40	N5°31.473', E118°17.767'	181107							N
PM 3-5-41	N5°31.473', E118°17.767'	181107							N
PM 3-5-42	N5°31.473', E118°17.767'	181107							N
PM 3-5-43	N5°31.473', E118°17.767'	181107							N
PM 3-5-44	N5°31.473', E118°17.767'	181107	2.66	18.80	Duplicate	28/12/07	Box 3 Nala	Y (Both)	Y
PM 3-5-45	N5°31.473', E118°17.767'	181107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-5-46	N5°31.473', E118°17.767'	181107							N
PM 3-5-47	N5°31.473', E118°17.767'	181107							N
PM 3-2-2	N5°33.497', E118°19.668'	181107							N
PM 3-2-3	N5°33.497', E118°19.668'	181107							N
PM 3-2-4	N5°33.497', E118°19.668'	181107							N
PM 3-2-5	N5°33.497', E118°19.668'	181107							N
PM 3-2-6	N5°33.497', E118°19.668'	181107							N
PM 3-2-7	N5°33.497', E118°19.668'	181107	1.66	30.12	Duplicate	28/12/07	Box 2 Nala	Y (Both)	Y
PM 3-2-8	N5°33.497', E118°19.668'	181107							N
PM 3-2-9	N5°33.497', E118°19.668'	181107			Single	03/01/08	Box 2 Nala	Y	N
PM 3-2-10	N5°33.497', E118°19.668'	181107							N
PM 3-2-11	N5°33.497', E118°19.668'	181107							N
PM 3-2-12	N5°33.497', E118°19.668'	181107							N
PM 3-2-13	N5°33.497', E118°19.668'	181107							N
PM 3-2-14	N5°33.497', E118°19.668'	181107			Single	03/01/08	Box 2 Nala	Y	N
PM 3-2-15	N5°33.497', E118°19.668'	181107							N
PM 3-2-16	N5°33.497', E118°19.668'	181107							N
PM 3-2-17	N5°33.497', E118°19.668'	181107			Single	03/01/08	Box 2 Nala	Y	N
PM 3-2-18	N5°33.497', E118°19.668'	181107			Single	03/01/08	Box 2 Nala	Y	N
PM 3-2-19	N5°33.497', E118°19.668'	181107							N
Mafa 3-6-1	N5°28.709', E118°13.472'	181107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-6-2	N5°28.709', E118°13.472'	181107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
PM 3-6-1	N5°27.937', E118°14.821'	191107							N
PM 3-6-2	N5°27.937', E118°14.821'	191107							N
PM 3-6-3	N5°27.937', E118°14.821'	191107							N
PM 3-6-4	N5°27.937', E118°14.821'	191107							N
PM 3-6-5	N5°27.937', E118°14.821'	191107	2.442	20.48	Duplicate	02/01/08	Box 3 Nala	Y (Only A)	Y
PM 3-6-6	N5°27.937', E118°14.821'	191107							N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-6-7	N5°27.937', E118°14.821'	191107							N
PM 3-6-8	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-9	N5°27.937', E118°14.821'	191107							N
PM 3-6-10	N5°27.937', E118°14.821'	191107							N
PM 3-6-11	N5°27.937', E118°14.821'	191107							N
PM 3-6-12	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-13	N5°27.937', E118°14.821'	191107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-6-14	N5°27.937', E118°14.821'	191107							N
PM 3-6-15	N5°27.937', E118°14.821'	191107							N
PM 3-6-16	N5°27.937', E118°14.821'	191107							N
PM 3-6-17	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-18	N5°27.937', E118°14.821'	191107	2.774	18.02	Duplicate	28/12/07	Box 3 Nala	Y (Both)	Y
PM 3-6-19	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-20	N5°27.937', E118°14.821'	191107							N
PM 3-6-21	N5°27.937', E118°14.821'	191107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-6-22	N5°27.937', E118°14.821'	191107							N
PM 3-6-23	N5°27.937', E118°14.821'	191107							N
PM 3-6-24	N5°27.937', E118°14.821'	191107							N
PM 3-6-25	N5°27.937', E118°14.821'	191107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-6-26	N5°27.937', E118°14.821'	191107			Single	05/01/08	Box 3 Nala	Y	N
PM 3-6-27	N5°27.937', E118°14.821'	191107							N
PM 3-6-28	N5°27.937', E118°14.821'	191107							N
PM 3-6-29	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-30	N5°27.937', E118°14.821'	191107							N
PM 3-6-31	N5°27.937', E118°14.821'	191107							N
PM 3-6-32	N5°27.937', E118°14.821'	191107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-6-33	N5°27.937', E118°14.821'	191107							N
PM 3-6-34	N5°27.937', E118°14.821'	191107							N
Mafa 3-7-1	N5°29.285', E118°13.445'	191107			Duplicate	28/12/07	Box 1 Mafa	Y (Only A)	N
Mafa 3-7-2	N5°29.285', E118°13.445'	191107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-7-3	N5°29.285', E118°13.445'	191107			Duplicate	28/12/07	Box 1 Mafa	Y (Only A)	N
Mafa 3-8-1	N5°28.302', E118°12.884'	191107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-8-2	N5°28.302', E118°12.884'	191107			Duplicate	28/12/07	Box 1 Mafa	Y (Both)	N
Mafa 3-9-1	N5°28.428', E118°13.255'	191107							N
Mafa 3-9-2	N5°28.428', E118°13.255'	191107							N
Mafa 3-9-3	N5°28.428', E118°13.255'	191107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 3-9-4	N5°28.428', E118°13.255'	191107							N
Mafa 3-9-5	N5°28.428', E118°13.255'	191107							N
Mafa 3-9-6	N5°28.428', E118°13.255'	191107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 3-9-7	N5°28.428', E118°13.255'	191107							N
Mafa 3-9-8	N5°28.428', E118°13.255'	191107	1.106	45.21	Duplicate	31/12/07	Box 1 Mafa	Y (Both)	Y
Popy 2-1-1	N5°35.292', E118°19.668'	191107	2.66	18.80	Duplicate	02/01/08	Box 7 Mafa		Y
PM 1-4-1	N5°40.028', E118°23.146'	201107							N
PM 1-4-2	N5°40.028', E118°23.146'	201107							N
PM 1-4-3	N5°40.028', E118°23.146'	201107							N
PM 1-4-4	N5°40.028', E118°23.146'	201107							N
PM 1-4-5	N5°40.028', E118°23.146'	201107							N
PM 1-4-6	N5°40.028', E118°23.146'	201107							N
PM 1-4-7	N5°40.028', E118°23.146'	201107							N
PM 1-4-8	N5°40.028', E118°23.146'	201107	1.96	25.51	Duplicate	28/12/07	Box 1 Nala	Y (Both)	Y
PM 1-4-9	N5°40.028', E118°23.146'	201107							N
PM 1-4-10	N5°40.028', E118°23.146'	201107							N
PM 1-4-11	N5°40.028', E118°23.146'	201107							N
PM 1-4-12	N5°40.028', E118°23.146'	201107							N
PM 1-4-13	N5°40.028', E118°23.146'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-4-14	N5°40.028', E118°23.146'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-4-15	N5°40.028', E118°23.146'	201107			Single	07/01/08	Box 1 Nala	Y	N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 1-4-16	N5°40.028', E118°23.146'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-4-17	N5°40.028', E118°23.146'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-4-18	N5°40.028', E118°23.146'	201107							N
PM 1-4-19	N5°40.028', E118°23.146'	201107							N
PM 1-4-20	N5°40.028', E118°23.146'	201107							N
PM 1-4-21	N5°40.028', E118°23.146'	201107							N
Popy 1-1-1	N5°39.382', E118°22.194'	201107	2.26	22.12	Duplicate	02/01/08	Box 7 Mafa		Y
PM 1-5-1	N5°38.361', E118°21.572'	201107							N
PM 1-5-2	N5°38.361', E118°21.572'	201107							N
PM 1-5-3	N5°38.361', E118°21.572'	201107							N
PM 1-5-4	N5°38.361', E118°21.572'	201107							N
PM 1-5-5	N5°38.361', E118°21.572'	201107							N
PM 1-5-6	N5°38.361', E118°21.572'	201107							N
PM 1-5-7	N5°38.361', E118°21.572'	201107			Single	04/01/08	Box 1 Nala	N	N
PM 1-5-8	N5°38.361', E118°21.572'	201107							N
PM 1-5-9	N5°38.361', E118°21.572'	201107							N
PM 1-5-10	N5°38.361', E118°21.572'	201107	3.06	16.34	Duplicate	28/12/07	Box 1 Nala	Y (Both)	Y
PM 1-5-11	N5°38.361', E118°21.572'	201107							N
PM 1-5-12	N5°38.361', E118°21.572'	201107			Single	04/01/08	Box 1 Nala	Y	N
PM 1-5-13	N5°38.361', E118°21.572'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-5-14	N5°38.361', E118°21.572'	201107							N
PM 1-5-15	N5°38.361', E118°21.572'	201107							N
PM 1-5-16	N5°38.361', E118°21.572'	201107							N
PM 1-5-17	N5°38.361', E118°21.572'	201107							N
PM 1-5-18	N5°38.361', E118°21.572'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 1-5-19	N5°38.361', E118°21.572'	201107			Single	07/01/08	Box 1 Nala	Y	N
PM 2-9-1	N5°38.811', E118°21.927'	201107							N
PM 2-9-2	N5°38.811', E118°21.927'	201107							N
PM 2-9-3	N5°38.811', E118°21.927'	201107							N
PM 2-9-4	N5°38.811', E118°21.927'	201107			Single	04/01/08	Box 2 Nala	Y	N
PM 2-9-5	N5°38.811', E118°21.927'	201107	0.36	138.89	Duplicate	28/12/07	Box 2 Nala	Y (Few A)	Y
PM 2-9-6	N5°38.811', E118°21.927'	201107							N
PM 2-9-7	N5°38.811', E118°21.927'	201107							N
Mafa 1-3-1	N5°34.852', E118°19.864'	201107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 1-3-2	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-3-3	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-3-4	N5°34.852', E118°19.864'	201107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 1-3-5	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-3-6	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-3-7	N5°34.852', E118°19.864'	201107	1.06	47.17	Duplicate	31/12/07	Box 1 Mafa	Y (B?)	Y
Mafa 1-3-8	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-3-9	N5°34.852', E118°19.864'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-4-1	N5°38.361', E118°21.572'	201107			Single	30/06/09	Box 7 Mafa	Y	N
Mafa 1-4-2	N5°38.361', E118°21.572'	201107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 1-4-3	N5°38.361', E118°21.572'	201107	1.16	43.10	Duplicate	31/12/07	Box 1 Mafa	N	Y
PM 3-7-1	N5°27.153', E118°15.040'	221107							N
PM 3-7-2	N5°27.153', E118°15.040'	221107	2.06	24.27	Duplicate	02/01/08	Box 3 Nala	Y (Both)	Y
PM 3-7-3	N5°27.153', E118°15.040'	221107							N
PM 3-7-4	N5°27.153', E118°15.040'	221107							N
PM 3-7-5	N5°27.153', E118°15.040'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-7-6	N5°27.153', E118°15.040'	221107							N
PM 3-7-7	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-8	N5°27.153', E118°15.040'	221107							N
PM 3-7-9	N5°27.153', E118°15.040'	221107							N
PM 3-7-10	N5°27.153', E118°15.040'	221107							N
PM 3-7-11	N5°27.153', E118°15.040'	221107							N
PM 3-7-12	N5°27.153', E118°15.040'	221107							N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-7-13	N5°27.153', E118°15.040'	221107							N
PM 3-7-14	N5°27.153', E118°15.040'	221107	2.16	23.15	Duplicate	02/01/08	Box 3 Nala	Y (Both)	Y
PM 3-7-15	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-16	N5°27.153', E118°15.040'	221107							N
PM 3-7-17	N5°27.153', E118°15.040'	221107							N
PM 3-7-18	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-19	N5°27.153', E118°15.040'	221107							N
PM 3-7-20	N5°27.153', E118°15.040'	221107							N
PM 3-7-21	N5°27.153', E118°15.040'	221107							N
PM 3-7-22	N5°27.153', E118°15.040'	221107							N
PM 3-7-23	N5°27.153', E118°15.040'	221107							N
PM 3-7-24	N5°27.153', E118°15.040'	221107							N
PM 3-7-25	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-26	N5°27.153', E118°15.040'	221107							N
PM 3-7-27	N5°27.153', E118°15.040'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-7-28	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-29	N5°27.153', E118°15.040'	221107							N
PM 3-7-30	N5°27.153', E118°15.040'	221107							N
PM 3-7-31	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-7-32	N5°27.153', E118°15.040'	221107							N
PM 3-7-33	N5°27.153', E118°15.040'	221107							N
PM 3-7-34	N5°27.153', E118°15.040'	221107	1.96	25.51	Duplicate	02/01/08	Box 3 Nala	Y (Few A)	Y
PM 3-7-35	N5°27.153', E118°15.040'	221107							N
PM 3-7-36	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	N	N
PM 3-7-37	N5°27.153', E118°15.040'	221107			Single	07/01/08	Box 3 Nala	Y	N
Mafa 3-10-1	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 3-10-2	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-3	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 3-10-4	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-5	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-6	N5°27.482', E118°15.139'	221107	1.86	26.88	Duplicate	31/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 3-10-7	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 3-10-8	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 1 Mafa	N	N
Mafa 3-10-9	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 3-10-10	N5°27.482', E118°15.139'	221107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 3-10-11	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-12	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-13	N5°27.482', E118°15.139'	221107							N
Mafa 3-10-14	N5°27.482', E118°15.139'	221107							N
PM 3-8-1	N5°27.740', E118°15.180'	221107							N
PM 3-8-2	N5°27.740', E118°15.180'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-8-3	N5°27.740', E118°15.180'	221107							N
PM 3-8-4	N5°27.740', E118°15.180'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-8-5	N5°27.740', E118°15.180'	221107							N
PM 3-8-6	N5°27.740', E118°15.180'	221107							N
PM 3-8-7	N5°27.740', E118°15.180'	221107							N
PM 3-8-8	N5°27.740', E118°15.180'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-8-9	N5°27.740', E118°15.180'	221107							N
PM 3-8-10	N5°27.740', E118°15.180'	221107			Single	07/01/08	Box 3 Nala	FEW	N
PM 3-8-11	N5°27.740', E118°15.180'	221107							N
PM 3-8-12	N5°27.740', E118°15.180'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-8-13	N5°27.740', E118°15.180'	221107	1.96	25.51	Duplicate	02/01/08	Box 3 Nala	Y (Both)	Y
PM 3-8-14	N5°27.740', E118°15.180'	221107							N
PM 3-8-15	N5°27.740', E118°15.180'	221107							N
PM 3-8-16	N5°27.740', E118°15.180'	221107							N
PM 3-8-17	N5°27.740', E118°15.180'	221107							N
PM 3-8-18	N5°27.740', E118°15.180'	221107			Single	03/01/08	Box 3 Nala	Y	N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 3-8-19	N5°27.740', E118°15.180'	221107							N
PM 3-8-20	N5°27.740', E118°15.180'	221107							N
PM 3-8-21	N5°27.740', E118°15.180'	221107			Single	04/01/08	Box 3 Nala	N	N
PM 3-8-22	N5°27.740', E118°15.180'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-8-23	N5°27.740', E118°15.180'	221107							N
PM 3-8-24	N5°27.740', E118°15.180'	221107	1.96	25.51	Duplicate	02/01/08	Box 3 Nala	Y Weird A	Y
PM 3-8-25	N5°27.740', E118°15.180'	221107			Single	03/01/08	Box 3 Nala	Y	N
PM 3-8-26	N5°27.740', E118°15.180'	221107							N
PM 3-8-27	N5°27.740', E118°15.180'	221107							N
PM 3-8-28	N5°27.740', E118°15.180'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-8-29	N5°27.740', E118°15.180'	221107			Single	04/01/08	Box 3 Nala	Y	N
PM 3-8-30	N5°27.740', E118°15.180'	221107							N
PM 3-8-31	N5°27.740', E118°15.180'	221107							N
PM 3-8-32	N5°27.740', E118°15.180'	221107							N
PM 3-8-33	N5°27.740', E118°15.180'	221107							N
PM 3-8-34	N5°27.740', E118°15.180'	221107							N
PM 3-8-35	N5°27.740', E118°15.180'	221107							N
PM 3-8-36	N5°27.740', E118°15.180'	221107			Single	07/01/08	Box 3 Nala	Y	N
PM 3-8-37	N5°27.740', E118°15.180'	221107							N
PM 3-8-38	N5°27.740', E118°15.180'	221107							N
Mafa 3-11-1	N5°27.740', E118°15.180'	221107							N
Mafa 3-11-2	N5°27.740', E118°15.180'	221107							N
Mafa 3-11-3	N5°27.740', E118°15.180'	221107							N
Mafa 3-11-4	N5°27.740', E118°15.180'	221107							N
Mafa 3-11-5	N5°27.740', E118°15.180'	221107			Single	02/01/08	Box 2 Mafa	N	N
Mafa 3-11-6	N5°27.740', E118°15.180'	221107	1.26	39.68	Duplicate	31/12/07	Box 1 Mafa	Y (Both)	Y
Mafa 3-11-7	N5°27.740', E118°15.180'	221107			Single	02/01/08	Box 2 Mafa	N	N
PM 3-9-1	N5°27.854', E118°15.175'	221107							N
PM 3-9-2	N5°27.854', E118°15.175'	221107	1.46	34.25	Duplicate	02/01/08	Box 3 Nala	N	Y
PM 3-9-3	N5°27.854', E118°15.175'	221107			Single	07/01/08	Box 3 Nala	FEW	N
PM 3-9-4	N5°27.854', E118°15.175'	221107							N
PM 3-9-5	N5°27.854', E118°15.175'	221107							N
Mafa 4-1-1	N5°28.065', E118°15.113'	240108			Duplicate	07/02/08	Box 2 Mafa	N	N
PM 4-2-1	N5°28.020', E118°15.113'	240108							N
PM 4-2-2	N5°28.020', E118°15.113'	240108							N
PM 4-2-3	N5°28.020', E118°15.113'	240108							N
PM 4-2-4	N5°28.020', E118°15.113'	240108							N
PM 4-2-5	N5°28.020', E118°15.113'	240108							N
PM 4-2-6	N5°28.020', E118°15.113'	240108							N
PM 4-2-7	N5°28.020', E118°15.113'	240108	1.487	33.62	Duplicate	09/02/08	Box 3 Nala	Y (Both)	Y
PM 4-2-8	N5°28.020', E118°15.113'	240108							N
PM 4-2-9	N5°28.020', E118°15.113'	240108							N
PM 4-2-10	N5°28.020', E118°15.113'	240108			Duplicate	24/06/09	Box 3 Nala	Y (Both)	N
PM 4-2-11	N5°28.020', E118°15.113'	240108							N
PM 4-2-12	N5°28.020', E118°15.113'	240108							N
PM 4-2-13	N5°28.020', E118°15.113'	240108							N
PM 4-2-14	N5°28.020', E118°15.113'	240108							N
PM 4-2-15	N5°28.020', E118°15.113'	240108							N
PM 4-2-16	N5°28.020', E118°15.113'	240108	3.038	16.46	Duplicate	08/02/08	Box 3 Nala	Y (Both)	Y
PM 4-2-17	N5°28.020', E118°15.113'	240108							N
PM 4-2-18	N5°28.020', E118°15.113'	240108							N
PM 4-2-19	N5°28.020', E118°15.113'	240108							N
Mafa 4-2-1	N5°28.932', E118°14.306'	250108			Duplicate	07/02/08	Box 2 Mafa	N?	N
PM 4-3-1	N5°28.932', E118°14.306'	250108							N
PM 4-3-2	N5°28.932', E118°14.306'	250108							N
PM 4-3-3	N5°28.932', E118°14.306'	250108							N
PM 4-3-4	N5°28.932', E118°14.306'	250108							N

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 4-3-5	N5°28.932', E118°14.306'	250108							N
PM 4-3-6	N5°28.932', E118°14.306'	250108	2.514	19.89	Duplicate	08/02/08	Box 4 Nala	as 4-3-1	Y
PM 4-3-7	N5°28.932', E118°14.306'	250108							N
PM 4-4-1	N5°28.616', E118°16.249'	250108							N
PM 4-4-2	N5°28.616', E118°16.249'	250108							N
Mafa 4-3-1	N5°28.616', E118°16.249'	250108			Duplicate	07/02/08	Box 2 Mafa	N?	N
PM 4-5-1	N5°29.464', E118°16.700'	250108							N
PM 4-5-2	N5°29.464', E118°16.700'	250108			Duplicate	08/02/08	Box 3 Nala	Y (Both)	N
PM 4-6-1	N5°30.565', E118°13.775'	250108			Duplicate	08/02/08	Box 3 Nala	Y (Both)	N
PM 4-6-2	N5°30.565', E118°13.775'	250108							N
Mane 4-1-1	N5°29.880', E118°14.167'	250108							N
PM 4-7-1	N5°30.189', E118°15.649'	260108			Single	09/02/08	Box 3 Nala	Y	N
PM 4-8-1	N5°29.841', E118°15.194'	260108	3.189	15.68	Duplicate	09/02/08	Box 3 Nala	Y (Both)	Y
PM 4-9-1	N5°30.170', E118°15.595'	260108							N
PM 4-9-2	N5°30.170', E118°15.595'	260108	4.65	10.75	Duplicate	09/02/08	Box 3 Nala	Y (Both)	Y
PM 4-9-3	N5°30.170', E118°15.595'	260108							N
PM 4-9-4	N5°30.170', E118°15.595'	260108							N
PM 4-9-5	N5°30.170', E118°15.595'	260108	2.818	17.74	Duplicate	09/02/08	Box 3 Nala	Y (Both)	Y
PM 4-9-6	N5°30.170', E118°15.595'	260108							N
PM 4-9-7	N5°30.170', E118°15.595'	260108			Duplicate	24/06/09	Box 4 Nala	Y (Both)	N
PM 4-9-8	N5°30.170', E118°15.595'	260108							N
PM 3-10-1	N5°33.198', E118°20.007'	260108							N
Popy 3-2-1	N5°32.494', E118°18.563'	260108	4.807	10.40	Duplicate	08/02/08	Box 7 Mafa		Y
PM 4-10-1	N5°30.287', E118°16.366'	260108			Duplicate	09/02/08	Box 4 Nala	Y (Only B)	N
PM 4-10-2	N5°30.287', E118°16.366'	260108							N
PM 4-11-1	N5°30.272', E118°16.291'	260108	4.033	12.40	Duplicate	09/02/08	Box 4 Nala	Y (Both)	Y
PM 4-11-2	N5°30.272', E118°16.291'	260108							N
PM 4-11-3	N5°30.272', E118°16.291'	260108							N
PM 4-11-4	N5°30.272', E118°16.291'	260108							N
PM 4-11-5	N5°30.272', E118°16.291'	260108							N
PM 4-11-6	N5°30.272', E118°16.291'	260108							N
PM 4-11-7	N5°30.272', E118°16.291'	260108							N
PM 4-11-8	N5°30.272', E118°16.291'	260108							N
PM 4-11-9	N5°30.272', E118°16.291'	260108	3.626	13.79	Duplicate	09/02/08	Box 4 Nala	Y (Only B)	Y
PM 4-11-10	N5°30.272', E118°16.291'	260108			Duplicate	24/06/09	Box 4 Nala	Y (Both)	N
PM 4-11-11	N5°30.272', E118°16.291'	260108							N
PM 4-12-1	N5°29.579', E118°14.769'	260108							N
PM 4-12-2	N5°29.579', E118°14.769'	260108			Duplicate	09/02/08	Box 4 Nala	N (Few B)	N
PM 4-12-3	N5°29.579', E118°14.769'	260108							N
Mafa 4-4-1	N5°29.351', E118°17.316'	270108	1.908	26.21	Duplicate	08/02/08	Box 2 Mafa	N?	Y
Mafa 4-4-2	N5°29.351', E118°17.316'	270108			Duplicate	08/02/08	Box 2 Mafa	Y (B?)	N
Mafa 4-4-3	N5°29.351', E118°17.316'	270108			Duplicate	08/02/08	Box 2 Mafa	Y (Both)	N
Popy 4-1-1	N5°29.028', E118°16.983'	270108	3.373	14.82	Duplicate	11/02/08	Box 7 Mafa		Y
Mafa 4-5-1	N5°28.050', E118°15.022'	280108			Duplicate	08/02/08	Box 2 Mafa	Y (B?)	N
Mafa 4-5-2	N5°28.050', E118°15.022'	280108			Duplicate	08/02/08	Box 2 Mafa	N?	N
Mafa 4-5-3	N5°28.050', E118°15.022'	280108	0.902	55.43	Duplicate	08/02/08	Box 2 Mafa	N?	Y
Mafa 4-5-4	N5°28.050', E118°15.022'	280108			Duplicate	08/02/08	Box 2 Mafa	Y (Both)	N
PM 4-13-1	N5°28.050', E118°15.022'	280108			Single	09/02/08	Box 4 Nala	Y	N
PM 4-14-1	N5°29.046', E118°15.407'	280108							N
PM 4-14-2	N5°29.046', E118°15.407'	280108							N
PM 4-14-3	N5°29.046', E118°15.407'	280108							N
PM 4-14-4	N5°29.046', E118°15.407'	280108	1.173	42.63	Duplicate	11/02/08	Box 4 Nala	Y (Both)	Y
PM 4-4-3	N5°28.616', E118°16.249'	280108							N
PM 4-4-4	N5°28.616', E118°16.249'	280108							N
PM 4-4-5	N5°28.616', E118°16.249'	280108							N
PM 4-4-6	N5°28.616', E118°16.249'	280108							N
PM 4-4-7	N5°28.616', E118°16.249'	280108	1.05	47.62	Duplicate	08/02/08	Box 3 Nala	Y (Both)	Y

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 4-4-8	N5°28.616', E118°16.249'	280108							N
PM 4-4-9	N5°28.616', E118°16.249'	280108							N
PM 4-4-10	N5°28.616', E118°16.249'	280108							N
PM 4-4-11	N5°28.616', E118°16.249'	280108							N
PM 4-4-12	N5°28.616', E118°16.249'	280108							N
Mafa 4-3-2	N5°28.616', E118°16.249'	280108			Duplicate	07/02/08	Box 2 Mafa	N?	N
Mafa 4-3-3	N5°28.616', E118°16.249'	280108							N
Mafa 4-3-4	N5°28.616', E118°16.249'	280108							N
Mafa 4-3-5	N5°28.616', E118°16.249'	280108			Single	07/02/08	Box 2 Mafa	N?	N
PM 4-15-1	N5°29.220', E118°16.114'	280108							N
PM 4-15-2	N5°29.220', E118°16.114'	280108							N
PM 4-15-3	N5°29.220', E118°16.114'	280108	0.823	60.75	Duplicate	11/02/08	Box 4 Nala	Y (Both)	Y
PM 4-15-4	N5°29.220', E118°16.114'	280108							Y
PM 4-16-1	N5°28.052', E118°15.138'	280108							N
PM 4-16-2	N5°28.052', E118°15.138'	280108							N
PM 4-16-3	N5°28.052', E118°15.138'	280108							N
PM 4-16-4	N5°28.052', E118°15.138'	280108	3.551	14.08	Duplicate	11/02/08	Box 4 Nala	Y (Only B)	Y
PM 4-16-5	N5°28.052', E118°15.138'	280108			Duplicate	24/06/09	Box 4 Nala	Y (Only A)	N
Mafa 4-6-1	N5°28.052', E118°15.138'	280108			Duplicate	08/02/08	Box 2 Mafa	N?	N
Mafa 4-7-1	N5°28.538', E118°16.190'	280108			Duplicate	08/02/08	Box 2 Mafa	Y (A?)	N
PM 4-17-1	N5°28.348', E118°15.980'	280108			Duplicate	24/06/09	Box 4 Nala	Y (Only A)	N
PM 4-17-2	N5°28.348', E118°15.980'	280108							N
PM 4-17-3	N5°28.348', E118°15.980'	280108							N
PM 4-17-4	N5°28.348', E118°15.980'	280108	3.209	15.58	Duplicate	11/02/08	Box 4 Nala	Y	Y
PM 6-1-1	N5°24.592', E118°01.734'	210708	1.81	27.62	Duplicate	90908	Box 4 Nala	Y (both)	Y
PM 6-1-2	N5°24.592', E118°01.734'	210708			Duplicate	250609	Box 4 Nala	Y (both)	
PM 6-1-3	N5°24.592', E118°01.734'	210708			Duplicate	240609	Box 4 Nala	Y (Only A)	
PM 6-1-4	N5°24.592', E118°01.734'	210708			Single	120908	Box 5 Nala	Y	
PM 6-1-5	N5°24.592', E118°01.734'	210708							
PM 6-1-6	N5°24.592', E118°01.734'	210708	2.62	19.08	Duplicate	90908	Box 5 Nala	FEW	Y
PM 6-1-7	N5°24.592', E118°01.734'	210708			Duplicate	240609	Box 5 Nala	Y (both)	
PM 7-1-1	N5°23.908', E118°00.821'	210708							
PM 7-1-2	N5°23.908', E118°00.821'	210708			Single	120908	Box 5 Nala	Y	
PM 7-1-3	N5°23.908', E118°00.821'	210708	2.99	16.72	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-1-4	N5°23.908', E118°00.821'	210708							
Mafa 7-1-1	N5°24.273', E117°59.497'	210708			Single	170908	Box 3 Mafa	Y	
PM 7-2-1	N5°24.273', E117°59.497'	210708	1.84	27.17	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-2-2	N5°24.273', E117°59.497'	210708	2.07	24.15	Duplicate	90908	Box 5 Nala	Y (few A)	Y/NC
PM 7-2-3	N5°24.273', E117°59.497'	210708	1.84	27.17					
Mafa 6-11-1	N5°24.245', E117°59.923'	210708	1.08	46.30					Y/NC
Mafa 6-11-2	N5°24.245', E117°59.923'	210708	2.08	24.04					Y/NC
Mafa 6-11-3	N5°24.245', E117°59.923'	210708	2.41	20.75					Y/NC
Mafa 6-11-4	N5°24.245', E117°59.923'	210708							
Mafa 6-11-5	N5°24.245', E117°59.923'	210708							
Mafa 6-11-6	N5°24.245', E117°59.923'	210708			Single	170908	Box 4 Mafa	Y	
Mafa 6-2-1	N5°24.405', E117°58.548'	210708							
Mafa 6-2-2	N5°24.405', E117°58.548'	210708							
Mafa 6-2-3	N5°24.405', E117°58.548'	210708							
Mafa 6-2-4	N5°24.405', E117°58.548'	210708			Single	170908	Box 4 Mafa	Y	
Mafa 6-2-5	N5°24.405', E117°58.548'	210708							
Mafa 6-2-6	N5°24.405', E117°58.548'	210708	4.72	10.59	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 6-2-7	N5°24.405', E117°58.548'	210708							
Mafa 6-2-8	N5°24.405', E117°58.548'	210708			Single	170908	Box 4 Mafa	Y	
Mafa 6-2-9	N5°24.405', E117°58.548'	210708							
Mafa 6-2-10	N5°24.405', E117°58.548'	210708	1.43	34.97	Duplicate	100908	Box 3 Mafa	Y (both)	Y
Mafa 6-3-1	N5°24.895', E118°02.149	210708			Single	170908	Box 4 Mafa	Y	
PM 7-3-1	N5°24.090', E117°58.994'	220708							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 7-3-2	N5°24.090', E117°58.994'	220708	2.4	20.83	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-3-3	N5°24.090', E117°58.994'	220708							
PM 7-3-4	N5°24.090', E117°58.994'	220708							
PM 7-3-5	N5°24.090', E117°58.994'	220708			Single	120908	Box 5 Nala	Y	
PM 7-3-6	N5°24.090', E117°58.994'	220708			Single	120908	Box 5 Nala	Y	
PM 7-3-7	N5°24.090', E117°58.994'	220708	2.46	20.33	Duplicate	90908	Box 5 Nala	WEIRD	Y/NC
PM 7-3-8	N5°24.090', E117°58.994'	220708							
PM 7-3-9	N5°24.090', E117°58.994'	220708							
PM 7-3-10	N5°24.090', E117°58.994'	220708							
PM 7-3-11	N5°24.090', E117°58.994'	220708	3.12	16.03	Duplicate	90908	Box 5 Nala	WEIRD	Y/NC
PM 7-4-1	N5°24.068', E117°59.030'	220708							
PM 7-4-2	N5°24.068', E117°59.030'	220708							
PM 7-4-3	N5°24.068', E117°59.030'	220708	4.44	11.26	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-4-4	N5°24.068', E117°59.030'	220708	3.99	12.53	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-4-5	N5°24.068', E117°59.030'	220708	3.45	14.49	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-4-6	N5°24.068', E117°59.030'	220708			Single	220908	Box 5 Nala	Y	
PM 7-4-7	N5°24.068', E117°59.030'	220708							
PM 7-4-8	N5°24.068', E117°59.030'	220708							
PM 7-4-9	N5°24.068', E117°59.030'	220708							
PM 7-4-10	N5°24.068', E117°59.030'	220708							
PM 7-4-11	N5°24.068', E117°59.030'	220708	5.2	9.62	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-4-12	N5°24.068', E117°59.030'	220708							
PM 7-4-13	N5°24.068', E117°59.030'	220708	5.57	8.98	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
PM 7-4-14	N5°24.068', E117°59.030'	220708							
PM 7-4-15	N5°24.068', E117°59.030'	220708							
PM 7-4-16	N5°24.068', E117°59.030'	220708	3.7	13.51	Duplicate	90908	Box 5 Nala	Y (both)	Y/NC
Mafa 6-4-1	N5°24.118', E117°59.332'	220708			Single	170908	Box 4 Mafa	Y	
Mafa 6-4-2	N5°24.118', E117°59.332'	220708							
Mafa 6-4-3	N5°24.118', E117°59.332'	220708	3.15	15.87	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 6-4-4	N5°24.118', E117°59.332'	220708							
Mafa 6-4-5	N5°24.118', E117°59.332'	220708							
Mafa 6-4-6	N5°24.118', E117°59.332'	220708			Single	170908	Box 4 Mafa	Y	
Mafa 6-4-7	N5°24.118', E117°59.332'	220708	2.8	17.86	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 6-4-8	N5°24.118', E117°59.332'	220708							
Mafa 7-2-1	N5°24.415', E118°01.393'	220708	1.91	26.18	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 7-2-2	N5°24.415', E118°01.393'	220708							
PM 7-5-1	N5°24.415', E118°01.393'	220708							
Mafa 6-5-1	N5°24.202', E118°00.693'	220708			Single	170908	Box 4 Mafa	Y	
Mafa 6-5-2	N5°24.202', E118°00.693'	220708	0.83	60.24	Duplicate	100908	Box 3 Mafa	Y (both)	N
PM 5-1-1	N5°24.786', E118°02.798'	230708							
PM 5-1-2	N5°24.786', E118°02.798'	230708	1.95	25.64	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-1-3	N5°24.786', E118°02.798'	230708							
PM 5-1-4	N5°24.786', E118°02.798'	230708							
PM 5-1-5	N5°24.786', E118°02.798'	230708							
PM 5-1-6	N5°24.786', E118°02.798'	230708	1.88	26.60	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-1-7	N5°24.786', E118°02.798'	230708			Single	220908	Box 4 Nala	Y	
PM 5-1-8	N5°24.786', E118°02.798'	230708							
PM 5-1-9	N5°24.786', E118°02.798'	230708							
PM 5-1-10	N5°24.786', E118°02.798'	230708	2.54	19.69	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-1-11	N5°24.786', E118°02.798'	230708			Single	220908	Box 4 Nala	Y	
PM 5-1-12	N5°24.786', E118°02.798'	230708							
PM 5-1-13	N5°24.786', E118°02.798'	230708	2.23	22.42	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-2-1	N5°24.803', E118°04.547'	230708							
PM 5-2-2	N5°24.803', E118°04.547'	230708							
PM 5-2-3	N5°24.803', E118°04.547'	230708	1.57	31.85	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-2-4	N5°24.803', E118°04.547'	230708	4.14	12.08	Duplicate	90908	Box 4 Nala	Y (both)	Y/NC
PM 5-2-5	N5°24.803', E118°04.547'	230708							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 5-2-6	N5°24.803', E118°04.547'	230708							
PM 5-2-7	N5°24.803', E118°04.547'	230708							
PM 5-2-8	N5°24.803', E118°04.547'	230708			Single	220908	Box 4 Nala	Y	
PM 5-2-9	N5°24.803', E118°04.547'	230708			Single	220908	Box 4 Nala	Y	
PM 6-2-1	N5°24.895', E118°02.149'	230708							
Mafa 6-6-1	N5°24.293', E118°00.455'	230708							
Mafa 6-6-2	N5°24.293', E118°00.455'	230708							
Mafa 6-6-3	N5°24.293', E118°00.455'	230708	1.5	33.33	Single	170908	Box 4 Mafa	Y	Y/NC
Mafa 6-6-4	N5°24.293', E118°00.455'	230708							
Mafa 6-7-1	N5°24.295', E118°00.375'	230708							
Mafa 6-7-2	N5°24.295', E118°00.375'	230708			Single	170908	Box 4 Mafa	Y	
Mafa 7-3-1	N5°24.355', E118°00.357'	230708	1.32	37.88	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 7-3-2	N5°24.355', E118°00.357'	230708							
PM 7-6-1	N5°24.355', E118°00.357'	230708	1	50.00	Duplicate	110908	Box 5 Nala	?	N
Otter? 7-1-1	N5°24.355', E118°00.357'	230708	2.62	19.08					Y/NC
Mafa 6-8-1	N5°24.253', E118°00.087'	230708			Single	180908	Box 4 Mafa	Y	
Mafa 6-8-2	N5°24.253', E118°00.087'	230708	1.11	45.05	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 6-8-3	N5°24.253', E118°00.087'	230708			Single	180908	Box 4 Mafa	EMPTY	
Mafa 6-9-1	N5°24.238', E117°59.838'	230708	0.97	51.55	Duplicate	100908	Box 3 Mafa	Y (both)	Y
Mafa 6-9-2	N5°24.238', E117°59.838'	230708							
PM 6-3-1	N5°24.238', E117°59.838'	230708	1.65	30.30	Duplicate	110908	Box 5 Nala	Y Weird A	Y
Mafa 6-10-1	N5°24.788', E118°02.302'	240708							
Mafa 6-10-2	N5°24.788', E118°02.302'	240708							
Mafa 6-10-3	N5°24.788', E118°02.302'	240708							
Mafa 6-10-4	N5°24.788', E118°02.302'	240708							
Mafa 6-10-5	N5°24.788', E118°02.302'	240708	0.93	53.76	Duplicate	100908	Box 3 Mafa	Y (both)	Y
Mafa 6-10-15	N5°24.788', E118°02.302'	240708							
Mafa 6-10-16	N5°24.788', E118°02.302'	240708	4.26	11.74	Duplicate	100908	Box 3 Mafa	Y (both)	N
Mafa 6-10-17	N5°24.788', E118°02.302'	240708							
Mafa 5-1-1	N5°24.828', E118°04.476'	240708			Single	180908	Box 4 Mafa	Y	
Mafa 5-1-2	N5°24.828', E118°04.476'	240708							
Mafa 5-1-3	N5°24.828', E118°04.476'	240708							
Mafa 5-1-4	N5°24.828', E118°04.476'	240708	1.5	33.33	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 5-1-5	N5°24.828', E118°04.476'	240708			Single	180908	Box 4 Mafa	Y	
Mafa 5-2-1	N5°24.401', E118°05.069'	240708							
Mafa 5-2-2	N5°24.401', E118°05.069'	240708							
Mafa 5-2-3	N5°24.401', E118°05.069'	240708			Single	180908	Box 4 Mafa	Y	
Mafa 5-2-4	N5°24.401', E118°05.069'	240708			Single	180908	Box 4 Mafa	Y	
Mafa 5-3-1	N5°25.267', E118°02.119'	250708							
Mafa 5-3-2	N5°25.267', E118°02.119'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-3-3	N5°25.267', E118°02.119'	250708							
Mafa 5-3-4	N5°25.267', E118°02.119'	250708							
Mafa 5-3-5	N5°25.267', E118°02.119'	250708							
Mafa 5-3-6	N5°25.267', E118°02.119'	250708	2.21	22.62	Duplicate	100908	Box 3 Mafa	Y (both)	Y
Mafa 5-3-7	N5°25.267', E118°02.119'	250708							
Mafa 5-3-8	N5°25.267', E118°02.119'	250708							
Mafa 5-3-9	N5°25.267', E118°02.119'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-3-10	N5°25.267', E118°02.119'	250708	2.12	23.58	Duplicate	100908	Box 3 Mafa	Y (Only A)	Y
Mafa 5-3-11	N5°25.267', E118°02.119'	250708							
Mafa 5-4-1	N5°25.439', E118°02.504'	250708							
Mafa 5-4-2	N5°25.439', E118°02.504'	250708	0.52	96.15	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 5-4-3	N5°25.439', E118°02.504'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-4-4	N5°25.439', E118°02.504'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-4-5	N5°25.439', E118°02.504'	250708	1.97	25.38	Duplicate	100908	Box 3 Mafa	Y (both)	N
Mafa 5-5-1	N5°25.024', E118°03.427'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-5-2	N5°25.024', E118°03.427'	250708			Single	180908	Box 4 Mafa	Y	
Mafa 5-5-3	N5°25.024', E118°03.427'	250708							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 5-5-4	N5°25.024', E118°03.427'	250708							
Mafa 6-10-6	N5°24.788', E118°02.302'	250708							
Mafa 6-10-7	N5°24.788', E118°02.302'	250708	2.06	24.27	Duplicate	100908	Box 3 Mafa	Y (both)	Y/NC
Mafa 6-10-8	N5°24.788', E118°02.302'	250708							
Mafa 6-10-9	N5°24.788', E118°02.302'	250708							
Mafa 6-10-10	N5°24.788', E118°02.302'	250708	1.29	38.76	Duplicate	100908	Box 3 Mafa	Y (both)	Y
PM 7-7-1	N5°24.706', E118°01.352'	250708	2	25.00	Duplicate	110908	Box 5 Nala	Y (few A)	Y/NC
PM 7-7-2	N5°24.706', E118°01.352'	250708							
PM 7-7-3	N5°24.706', E118°01.352'	250708	1	50.00	Duplicate	110908	Box 5 Nala	Y (both)	Y/NC
PM 7-7-4	N5°24.706', E118°01.352'	250708							
PM 7-7-5	N5°24.706', E118°01.352'	250708							
PM 7-7-6	N5°24.706', E118°01.352'	250708							
PM 7-7-7	N5°24.706', E118°01.352'	250708	2	25.00	Duplicate	110908	Box 5 Nala	Y (both)	Y/NC
PM 7-8-1	N5°24.444', E118°01.370'	250708	2	25.00	Duplicate	110908	Box 6 Nala	Y (both)	Y/NC
Mafa 5-7-1	N5°27.531', E118°09.552'	290708			Single	180908	Box 4 Mafa	Y	
Mafa 5-6-1	N5°27.943', E118°08.324'	290708			Single	180908	Box 4 Mafa	Y	
Mafa 5-6-2	N5°27.943', E118°08.324'	290708							
Mafa 5-6-3	N5°27.943', E118°08.324'	290708			Single	180908	Box 4 Mafa	EMPTY	
PM 5-3-1	N5°27.569', E118°11.700'	300708			Single	220908	Box 4 Nala	Y	
PM 5-3-2	N5°27.569', E118°11.700'	300708							
PM 5-3-3	N5°27.569', E118°11.700'	300708			Single	220908	Box 4 Nala	Y	
PM 5-3-4	N5°27.569', E118°11.700'	300708							
PM 5-3-5	N5°27.569', E118°11.700'	300708							
PM 5-3-6	N5°27.569', E118°11.700'	300708							
PM 5-3-7	N5°27.569', E118°11.700'	300708	1.75	28.57	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-3-8	N5°27.569', E118°11.700'	300708							
PM 5-3-9	N5°27.569', E118°11.700'	300708							
PM 5-3-10	N5°27.569', E118°11.700'	300708							
PM 5-3-11	N5°27.569', E118°11.700'	300708	1.71	29.24	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-3-12	N5°27.569', E118°11.700'	300708							
PM 5-3-13	N5°27.569', E118°11.700'	300708							
PM 5-3-14	N5°27.569', E118°11.700'	300708							
PM 5-3-15	N5°27.569', E118°11.700'	300708			Single	220908	Box 4 Nala	Y	
PM 5-4-1	N5°27.212', E118°11.174'	300708	3.45	14.49	Duplicate	110908	Box 4 Nala	Y (few A)	Y
PM 5-4-2	N5°27.212', E118°11.174'	300708	2.27	22.03	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-4-3	N5°27.212', E118°11.174'	300708							
PM 5-4-4	N5°27.212', E118°11.174'	300708							
PM 5-4-5	N5°27.212', E118°11.174'	300708							
PM 5-4-6	N5°27.212', E118°11.174'	300708			Single	220908	Box 4 Nala	Y	
PM 5-4-7	N5°27.212', E118°11.174'	300708	4.82	10.37	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-4-8	N5°27.212', E118°11.174'	300708							
PM 5-4-9	N5°27.212', E118°11.174'	300708							
PM 5-4-10	N5°27.212', E118°11.174'	300708			Single	220908	Box 4 Nala	Y	
PM 5-4-11	N5°27.212', E118°11.174'	300708							
PM 5-4-12	N5°27.212', E118°11.174'	300708	2.52	19.84	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-4-13	N5°27.212', E118°11.174'	300708							
PM 5-4-14	N5°27.212', E118°11.174'	300708							
PM 5-4-15	N5°27.212', E118°11.174'	300708							
Mafa 3-12-1	N5°27.182', E118°10.345'	300708			Single	180908	Box 4 Mafa	Y	
PM 3-11-1	N5°27.566', E118°07.328'	300708			Single	220908	Box 3 Nala	Y	
PM 3-11-2	N5°27.566', E118°07.328'	300708	2.42	20.66	No EtOH!				Y
PM 3-11-3	N5°27.566', E118°07.328'	300708							
PM 3-11-4	N5°27.566', E118°07.328'	300708			Duplicate	110908	Box 3 Nala	Y (both)	
PM 5-5-1	N5°27.649', E118°07.338'	300708	1.86	26.88	Duplicate	110908	Box 4 Nala	Y (both)	Y
PM 5-5-2	N5°27.649', E118°07.338'	300708			Single	220908	Box 4 Nala	Y	
PM 5-5-3	N5°27.649', E118°07.338'	300708							
PM 5-5-4	N5°27.649', E118°07.338'	300708							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 5-5-5	N5°27.649', E118°07.338'	300708			Single	220908	Box 4 Nala	Y	
Mafa 6-10-11	N5°24.788', E118°02.302'	300708	1.46	34.25	Duplicate	100908	Box 3 Mafa	Y (both)	Y
Mafa 6-10-12	N5°24.788', E118°02.302'	300708							
Popy 6-1-1	N5°224.728', E118°02.398'	10808	6.49	7.70	Duplicate	120908	Box 7 Mafa		Y
Mafa 6-10-13	N5°24.788', E118°02.302'	60808	2.9	17.24					Y
Mafa 6-10-14	N5°24.788', E118°02.302'	60808	3.9	12.82					Y
Mafa 6-1-1	N5°24.607', E118°01.699'	80808			Single	180908	Box 4 Mafa	Y	
Mafa 6-1-2	N5°24.607', E118°01.699'	80808	4.1	12.20	Duplicate	100908	Box 3 Mafa	Y (both)	Y
PM 6-4-1	N5°23.970', E118°01.155'	80808			Single	220908	Box 5 Nala	Y	
PM 6-4-2	N5°23.970', E118°01.155'	80808			Single	220908	Box 5 Nala	Y	
PM 6-4-3	N5°23.970', E118°01.155'	80808			Duplicate	110908	Box 5 Nala	Y Weird A	
PM 6-4-4	N5°23.970', E118°01.155'	80808			Duplicate	240609	Box 5 Nala	Y (few A)	
PM 6-5-1	N5°24.868', E118°02.193'	80808	3.1	16.13	Duplicate	110908	Box 5 Nala	N	Y
PM 6-5-2	N5°24.868', E118°02.193'	80808	2.2	22.73	Duplicate	110908	Box 5 Nala	N (FEW B)	Y
PM 7-9-1	N5°24.514', E117°58.635'	150808							
PM 7-9-2	N5°24.514', E117°58.635'	150808			Single	220908	Box 5 Nala	Y	
PM 7-9-3	N5°24.514', E117°58.635'	150808							
PM 7-9-4	N5°24.514', E117°58.635'	150808	3.21	15.58	Duplicate	110908	Box 6 Nala	Y (both)	Y
PM 7-9-5	N5°24.514', E117°58.635'	150808	1.4	35.71	Duplicate	110908	Box 6 Nala	Y (both)	Y
Mafa 7-4-1	N5°24.942', E117°56.906'	150808							
Mafa 7-4-2	N5°24.942', E117°56.906'	150808	2.75	18.18	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 7-4-3	N5°24.942', E117°56.906'	150808			Single	180908	Box 5 Mafa	Y	
Mafa 7-4-4	N5°24.942', E117°56.906'	150808			Single	180908	Box 5 Mafa	Y	
Mafa 7-4-5	N5°24.942', E117°56.906'	150808	3.42	14.62	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 7-4-6	N5°24.942', E117°56.906'	150808			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 7-4-7	N5°24.942', E117°56.906'	150808			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 7-4-8	N5°24.942', E117°56.906'	150808							
Mafa 7-4-9	N5°24.942', E117°56.906'	150808			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 7-4-10	N5°24.942', E117°56.906'	150808	2.37	21.10	Duplicate	130908	Box 3 Mafa	Y (Both)	Y
PM 7-10-1	N5°24.744', E117°57.394'	150808							
PM 7-10-2	N5°24.744', E117°57.394'	150808							
PM 7-10-3	N5°24.744', E117°57.394'	150808			Single	220908	Box 6 Nala	FEW	
Mafa 6-12-1	N5°24.378', E117°58.273'	150808	3.13	15.97	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 6-12-2	N5°24.378', E117°58.273'	150808							
Mafa 6-12-3	N5°24.378', E117°58.273'	150808	2.89	17.30	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 6-12-4	N5°24.378', E117°58.273'	150808							
Mafa 6-12-5	N5°24.378', E117°58.273'	150808							
Mafa 6-12-6	N5°24.378', E117°58.273'	150808			Single	180908	Box 5 Mafa	Y	
Mafa 6-12-7	N5°24.378', E117°58.273'	150808	3.4	14.71	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 6-12-8	N5°24.378', E117°58.273'	150808			Single	190908	Box 5 Mafa	Y	
Mafa 6-12-9	N5°24.378', E117°58.273'	150808							
Mafa 6-12-10	N5°24.378', E117°58.273'	150808							
Mafa 6-12-11	N5°24.378', E117°58.273'	150808							
Mafa 6-12-12	N5°24.378', E117°58.273'	150808							
PM 7-11-1	N5°24.487', E117°58.588'	150808							
PM 7-11-2	N5°24.487', E117°58.588'	150808							
PM 7-11-3	N5°24.487', E117°58.588'	150808			Single	220908	Box 6 Nala	Y	
PM 7-11-4	N5°24.487', E117°58.588'	150808	2.38	21.01	Duplicate	110908	Box 6 Nala	Y (both)	Y
PM 7-11-5	N5°24.487', E117°58.588'	150808			Single	220908	Box 6 Nala	Y	
PM 7-11-6	N5°24.487', E117°58.588'	150808	5.11	9.78	Duplicate	120908	Box 6 Nala	Y (both)	Y
PM 7-11-7	N5°24.487', E117°58.588'	150808							
PM 7-11-8	N5°24.487', E117°58.588'	150808	5.15	9.71	Duplicate	120908	Box 6 Nala	Y (both)	Y
PM 7-11-9	N5°24.487', E117°58.588'	150808							
PM 7-11-10	N5°24.487', E117°58.588'	150808							
PM 7-11-11	N5°24.487', E117°58.588'	150808							
PM 7-11-12	N5°24.487', E117°58.588'	150808	5.24	9.54	Duplicate	120908	Box 6 Nala	Y (both)	Y
PM 7-11-13	N5°24.487', E117°58.588'	150808							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 7-11-14	N5°24.487', E117°58.588'	150808							
PM 7-11-15	N5°24.487', E117°58.588'	150808							
PM 7-11-16	N5°24.487', E117°58.588'	150808	5.93	8.43	Duplicate	120908	Box 6 Nala	Y (both)	Y
PM 7-11-17	N5°24.487', E117°58.588'	150808							
PM 7-11-18	N5°24.487', E117°58.588'	150808			Single	220908	Box 6 Nala	Y	
PM 7-11-19	N5°24.487', E117°58.588'	150808							
PM 7-11-20	N5°24.487', E117°58.588'	150808							
PM 7-11-21	N5°24.487', E117°58.588'	150808							
PM 7-11-22	N5°24.487', E117°58.588'	150808							
PM 7-11-23	N5°24.487', E117°58.588'	150808							
PM 7-11-24	N5°24.487', E117°58.588'	150808							
PM 7-11-25	N5°24.487', E117°58.588'	150808			Single	230908	Box 6 Nala	Y	
PM 7-11-26	N5°24.487', E117°58.588'	150808	5.84	8.56	Duplicate	120908	Box 6 Nala	Y (both)	Y
Mafa 7-5-1	N5°24.633', E117°58.801'	170808	4.7	10.64	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 7-5-2	N5°24.633', E117°58.801'	170808			Single	190908	Box 5 Mafa	Y	
Mafa 7-5-3	N5°24.633', E117°58.801'	170808							
Mafa 7-5-4	N5°24.633', E117°58.801'	170808			Single	190908	Box 5 Mafa	Y	
Mafa 7-5-5	N5°24.633', E117°58.801'	170808			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 7-5-6	N5°24.633', E117°58.801'	170808			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 7-5-7	N5°24.633', E117°58.801'	170808	5.25	9.52	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 7-5-8	N5°24.633', E117°58.801'	170808			Duplicate	300609	Box 7 Mafa	Y (both)	
PM 7-12-1	N5°24.155', E117°58.925'	170808	2.6	19.23	Single	230908	Box 6 Nala	Y	
PM 7-12-2	N5°24.155', E117°58.925'	170808			Duplicate	120908	Box 6 Nala	FEW	
Mafa 6-13-1	N5°24.043', E117°59.196'	170808	2.5	20.00	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 6-13-2	N5°24.043', E117°59.196'	170808			Single	190908	Box 5 Mafa	Y	
PM 7-13-1	N5°24.109', E117°59.184'	170808	2.33	21.46					Y
Mafa 7-6-1	N5°24.341', E118°00.555'	170808			Single	190908	Box 5 Mafa	Y	
Mafa 7-6-2	N5°24.341', E118°00.555'	170808			Single	190908	Box 5 Mafa	Y	
Mafa 7-6-3	N5°24.341', E118°00.555'	170808	1.4	35.71	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 7-6-4	N5°24.341', E118°00.555'	170808	1.69	29.59					Y
Mafa 5-8-1	N5°25.228', E118°02.055'	180808							
Mafa 5-8-2	N5°25.228', E118°02.055'	180808			Duplicate	130908	Box 3 Mafa	Y (both)	
Mafa 5-8-3	N5°25.228', E118°02.055'	180808	2.59	19.31	Single	190908	Box 5 Mafa	FEW	Y
Mafa 5-8-4	N5°25.228', E118°02.055'	180808			Single	190908	Box 5 Mafa	Y	
Mafa 5-8-5	N5°25.228', E118°02.055'	180808	2.05	24.39	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 5-8-6	N5°25.228', E118°02.055'	180808							
Mafa 5-9-1	N5°24.929', E118°02.959'	180808			Single	190908	Box 5 Mafa	Y	
Mafa 5-9-2	N5°24.929', E118°02.959'	180808							
Mafa 5-9-3	N5°24.929', E118°02.959'	180808	3.15	15.87	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 5-9-4	N5°24.929', E118°02.959'	180808			Single	190908	Box 5 Mafa	Y	
Mafa 5-9-5	N5°24.929', E118°02.959'	180808	4.29	11.66	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 5-9-6	N5°24.929', E118°02.959'	180808							
Mafa 5-9-7	N5°24.929', E118°02.959'	180808							
Mafa 5-9-8	N5°24.929', E118°02.959'	180808							
Mafa 5-10-1	N5°25.029', E118°03.418'	180808	3.61	13.85	Duplicate	130908	Box 3 Mafa	Y (both)	Y
Mafa 5-10-2	N5°25.029', E118°03.418'	180808							
Mafa 5-10-3	N5°25.029', E118°03.418'	180808	2.79	17.92	Duplicate	150908	Box 3 Mafa	Y (both)	Y
Mafa 5-10-4	N5°25.029', E118°03.418'	180808	2.93	17.06	Duplicate	150908	Box 3 Mafa	Y (both)	Y
Mafa 5-10-5	N5°25.029', E118°03.418'	180808							
Mafa 5-10-6	N5°25.029', E118°03.418'	180808							
Mafa 5-10-7	N5°25.029', E118°03.418'	180808							
PM 5-6-1	N5°25.029', E118°03.418'	180808							
PM 5-6-2	N5°25.029', E118°03.418'	180808							
PM 5-6-3	N5°25.029', E118°03.418'	180808							
PM 5-6-4	N5°25.029', E118°03.418'	180808	4.58	10.92	Duplicate	120908	Box 4 Nala	Y (both)	Y
PM 5-6-5	N5°25.029', E118°03.418'	180808			Single	230908	Box 4 Nala	Y	
PM 5-6-6	N5°25.029', E118°03.418'	180808							

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
PM 5-6-7	N5°25.029', E118°03.418'	180808			Single	230908	Box 4 Nala	Y	
PM 5-6-8	N5°25.029', E118°03.418'	180808							
PM 5-6-9	N5°25.029', E118°03.418'	180808	3.76	13.30	Duplicate	120908	Box 4 Nala	Y (both)	Y
PM 5-6-10	N5°25.029', E118°03.418'	180808	3.78	13.23	Duplicate	120908	Box 4 Nala	Y (both)	Y
PM 5-6-11	N5°25.029', E118°03.418'	180808							
PM 5-6-12	N5°25.029', E118°03.418'	180808			Single	230908	Box 4 Nala	N	
PM 5-6-13	N5°25.029', E118°03.418'	180808	4.68	10.68	Duplicate	120908	Box 4 Nala	Y (both)	Y
PM 5-6-14	N5°25.029', E118°03.418'	180808							
PM 5-6-15	N5°25.029', E118°03.418'	180808							
PM 6-6-1	N5°24.451', E118°04.711'	190808	3.32	15.06	Duplicate	120908	Box 5 Nala	Y (both)	Y
PM 6-6-2	N5°24.451', E118°04.711'	190808	3.74	13.37	Duplicate	120908	Box 5 Nala	Y (both)	Y
PM 6-6-3	N5°24.451', E118°04.711'	190808	5.02	9.96	Duplicate	120908	Box 5 Nala	Y (both)	Y
PM 6-6-4	N5°24.451', E118°04.711'	190808			Duplicate	240609	Box 5 Nala	Y (both)	
PM 6-6-5	N5°24.451', E118°04.711'	190808			Duplicate	250609	Box 5 Nala	Y (both)	
PM 6-6-6	N5°24.451', E118°04.711'	190808	3.76	13.30	Duplicate	120908	Box 5 Nala	Y (both)	Y
PM 6-6-7	N5°24.451', E118°04.711'	190808			Duplicate	250609		Y (both)	
PM 6-6-8	N5°24.451', E118°04.711'	190808							
PM 6-6-9	N5°24.451', E118°04.711'	190808			Single	230908	Box 5 Nala	Y	
PM 6-6-10	N5°24.451', E118°04.711'	190808			Single	230908	Box 5 Nala	Y	
PM 6-6-11	N5°24.451', E118°04.711'	190808			Duplicate	240609	Box 5 Nala	Y (few A)	
PM 6-7-1	N5°24.935', E118°04.086'	190808	3.95	12.66	Duplicate	120908	Box 5 Nala	N (FEW B)	Y
PM 6-7-2	N5°24.935', E118°04.086'	190808	2.57	19.46	Duplicate	120908	Box 5 Nala	few (weird B)	Y
PM 6-7-3	N5°24.935', E118°04.086'	190808			Duplicate	250609	Box 5 Nala	N	
PM 6-7-4	N5°24.935', E118°04.086'	190808			Duplicate	250609	Box 5 Nala	N	
PM 6-7-5	N5°24.935', E118°04.086'	190808			Single	230908	Box 5 Nala	Y	
Mafa 6-14-1	N5°24.641', E118°04.016'	190808	3.86	12.95	Duplicate	150908	Box 3 Mafa	Y (both)	Y
Mafa 6-14-2	N5°24.641', E118°04.016'	190808	4.44	11.26	Duplicate	150908	Box 3 Mafa	Y (both)	Y
Mafa 6-14-3	N5°24.641', E118°04.016'	190808	3.9	12.82	Duplicate	150908	Box 3 Mafa	Y (both)	Y
Mafa 6-14-4	N5°24.641', E118°04.016'	190808							
Mafa 6-14-5	N5°24.641', E118°04.016'	190808							
Mafa 6-14-6	N5°24.641', E118°04.016'	190808							
Mafa 6-14-7	N5°24.641', E118°04.016'	190808							
Mafa 6-14-8	N5°24.641', E118°04.016'	190808							
Mafa 6-15-1	N5°24.609', E118°01.837'	200808	3.94	12.69	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 6-15-2	N5°24.609', E118°01.837'	200808							
Mafa 6-15-3	N5°24.609', E118°01.837'	200808	3.6	13.89	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 6-15-4	N5°24.609', E118°01.837'	200808							
Mafa 6-15-5	N5°24.609', E118°01.837'	200808			Single	190908	Box 5 Mafa	Y	
Mafa 6-15-6	N5°24.609', E118°01.837'	200808							
Mafa 6-15-7	N5°24.609', E118°01.837'	200808	4.2	11.90	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 7-7-1	N5°24.495', E118°01.306'	200808							
Mafa 7-7-2	N5°24.495', E118°01.306'	200808	3.06	16.34	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 7-7-3	N5°24.495', E118°01.306'	200808							
Mafa 7-7-4	N5°24.495', E118°01.306'	200808							
Mafa 7-7-5	N5°24.495', E118°01.306'	200808	3.94	12.69	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 7-7-6	N5°24.495', E118°01.306'	200808			Single	190908	Box 5 Mafa	Y	
PM 7-14-1	N5°23.865', E118°00.827'	200808			Single	230908	Box 6 Nala	Y	
PM 7-14-2	N5°23.865', E118°00.827'	200808	2.12	23.58					Y
PM 7-14-3	N5°23.865', E118°00.827'	200808			Single	230908	Box 6 Nala	Y	
PM 7-14-4	N5°23.865', E118°00.827'	200808	2.79	17.92					Y
Mafa 7-8-1	N5°24.225', E118°00.772'	200808	7.36	6.79	Duplicate	150908	Box 4 Mafa	N	Y
Mafa 7-8-2	N5°24.225', E118°00.772'	200808			Single	190908	Box 5 Mafa	Y	
Mafa 7-8-3	N5°24.225', E118°00.772'	200808	5.5	9.09	Duplicate	150908	Box 4 Mafa	N (FEW B)	Y
Mafa 7-8-4	N5°24.225', E118°00.772'	200808							
Mafa 7-8-5	N5°24.225', E118°00.772'	200808							
Mafa 7-8-6	N5°24.225', E118°00.772'	200808	2.03	24.63					Y
Mafa 2-5-1	N5°31.888', E118°17.454'	280808			Single	190908	Box 5 Mafa	Y	

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 2-6-1	N5°32.252', E118°17.607'	280808	2.68	18.66	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 2-6-2	N5°32.252', E118°17.607'	280808	3.92	12.76	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 2-6-3	N5°32.252', E118°17.607'	280808	2.77	18.05	Duplicate	150908	Box 4 Mafa	Y (both)	Y
Mafa 2-6-4	N5°32.252', E118°17.607'	280808			Single	190908	Box 5 Mafa	Y	
Mafa 2-6-5	N5°32.252', E118°17.607'	280808							
Mafa 2-6-6	N5°32.252', E118°17.607'	280808							
Mafa 2-6-7	N5°32.252', E118°17.607'	280808							
Mafa 2-7-1	N5°33.004', E118°19.001'	280808	3.13	15.97	Single	190908	Box 5 Mafa	Y	Y
Mafa 2-7-2	N5°33.004', E118°19.001'	280808			Single	190908	Box 5 Mafa	Y	
Mafa 2-7-3	N5°33.004', E118°19.001'	280808			Single	190908	Box 5 Mafa	Y	
Mafa 2-7-4	N5°33.004', E118°19.001'	280808			Single	190908	Box 5 Mafa	Y	
Mafa 2-7-5	N5°33.004', E118°19.001'	280808							
Mafa 2-7-6	N5°33.004', E118°19.001'	280808	3.1	16.13	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 2-7-7	N5°33.004', E118°19.001'	280808	4.92	10.16	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 2-7-8	N5°33.004', E118°19.001'	280808							
Mafa 2-7-9	N5°33.004', E118°19.001'	280808							
Mafa 2-7-10	N5°33.004', E118°19.001'	280808							
Mafa 2-7-11	N5°33.004', E118°19.001'	280808							
Mafa 2-7-12	N5°33.004', E118°19.001'	280808			Single	190908	Box 5 Mafa	Y	
Mafa 2-7-13	N5°33.004', E118°19.001'	280808							
Mafa 2-7-14	N5°33.004', E118°19.001'	280808							
Popy 6-2-1	N5°24.846', E118°02.286'	240808	3.44	14.53	Duplicate	120908	Box 7 Mafa		Y
Mafa 1-5-1	N5°35.981', E118°20.222'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 1-5-2	N5°35.981', E118°20.222'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 1-5-3	N5°35.981', E118°20.222'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 1-6-1	N5°36.142', E118°19.597'	290808			Single	200908	Box 5 Mafa	N	
Mafa 1-6-2	N5°36.142', E118°19.597'	290808	1.75	28.57	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 2-8-2	N5°35.069', E118°19.697'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 2-8-1	N5°35.069', E118°19.697'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 1-7-1	N5°34.792', E118°19.875'	290808			Single	200908	Box 5 Mafa	Y	
Mafa 1-7-2	N5°34.792', E118°19.875'	290808			Single	200908	Box 5 Mafa		Y
Mafa 3-17-1	N5°27.830', E118°15.201'	300808							
Mafa 3-17-2	N5°27.830', E118°15.201'	300808	3.93	12.72	Duplicate	160809	Box 4 Mafa	Y (both)	Y
Mafa 3-17-3	N5°27.830', E118°15.201'	300808	6.47	7.73	Duplicate	160809	Box 4 Mafa	Y (both)	Y
Mafa 3-17-4	N5°27.830', E118°15.201'	300808							
Mafa 3-17-5	N5°27.830', E118°15.201'	300808	4.69	10.66	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 3-17-6	N5°27.830', E118°15.201'	300808							
Mafa 3-17-7	N5°27.830', E118°15.201'	300808							
Mafa 3-17-8	N5°27.830', E118°15.201'	300808			Single	200908	Box 5 Mafa	Y	
Mafa 3-17-9	N5°27.830', E118°15.201'	300808							
Mafa 3-17-10	N5°27.830', E118°15.201'	300808							
Mafa 3-17-11	N5°27.830', E118°15.201'	300808	3.44	14.53	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 3-13-1	N5°29.764', E118°12.415'	310808			Single	200908	Box 5 Mafa	Y	
Mafa 3-13-2	N5°29.764', E118°12.415'	310808			Single	210908	Box 5 Mafa	Y	
Mafa 3-13-3	N5°29.764', E118°12.415'	310808	1.94	25.77	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 3-14-1	N5°29.041', E118°12.920'	310808	3.68	13.59	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 3-15-1	N5°29.281', E118°13.456'	310808							
Mafa 3-15-2	N5°29.281', E118°13.456'	310808							
Mafa 3-15-3	N5°29.281', E118°13.456'	310808	3.59	13.93	Duplicate	160908	Box 4 Mafa	Y (both)	Y
Mafa 3-15-4	N5°29.281', E118°13.456'	310808	5.11	9.78	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 3-16-1	N5°28.361', E118°13.185'	310808			Single	210908	Box 5 Mafa	Y	
Mafa 3-16-2	N5°28.361', E118°13.185'	310808							
Mafa 3-16-3	N5°28.361', E118°13.185'	310808	3.5	14.29	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 3-16-4	N5°28.361', E118°13.185'	310808			Single	210908	Box 5 Mafa	Y	
Mafa 3-16-5	N5°28.361', E118°13.185'	310808	4.88	10.25	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 3-16-6	N5°28.361', E118°13.185'	310808							
Mafa 4-8-1	N5°29.109', E118°14.090'	310808			Single	210908	Box 5 Mafa	Y	

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 4-8-2	N5°29.109', E118°14.090'	310808			Single	210908	Box 5 Mafa	Y	
Mafa 4-8-3	N5°29.109', E118°14.090'	310808	1.91	26.18	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 4-8-4	N5°29.109', E118°14.090'	310808			Single	300609	Box 7 Mafa	Y	
Mafa 4-9-1	N5°28.051', E118°15.076'	310808			Single	210908	Box 5 Mafa	Y	
Mafa 4-9-2	N5°28.051', E118°15.076'	310808	1.5	33.33	Duplicate	170908	Box 4 Mafa	FEW	Y
Mafa 4-9-3	N5°28.051', E118°15.076'	310808			Single	210908	Box 5 Mafa		
Mafa 2-9-1	N5°30.314', E118°16.400'	10908	3.69	13.55	Duplicate	170908	Box 4 Mafa	N (FEW B)	Y
Mafa 2-9-2	N5°30.314', E118°16.400'	10908							
Mafa 4-10-1	N5°30.105', E118°15.547'	10908			Single	210908	Box 5 Mafa	Y	
Mafa 4-10-2	N5°30.105', E118°15.547'	10908			Single	210908	Box 5 Mafa	Y	
Mafa 4-10-3	N5°30.105', E118°15.547'	10908	4.21	11.88	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 4-10-4	N5°30.105', E118°15.547'	10908			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 4-10-5	N5°30.105', E118°15.547'	10908			Single	210908	Box 5 Mafa	Y	
Mafa 4-11-1	N5°29.453', E118°17.273'	10908			Duplicate	300609	Box 7 Mafa		
Mafa 4-11-2	N5°29.453', E118°17.273'	10908			Single	210908	Box 5 Mafa	Y	
Mafa 4-11-3	N5°29.453', E118°17.273'	10908	5.4	9.26	Duplicate	170908	Box 4 Mafa	Y (both)	Y
Mafa 4-11-4	N5°29.453', E118°17.273'	10908			Duplicate	300609	Box 7 Mafa	Y (both)	
Mafa 4-11-5	N5°29.453', E118°17.273'	10908			Single	210908	Box 5 Mafa	Y	
PM 10c-1-1	N5°26.542', E117°43.575'	121008			Single	281008	Box 6 Nala	Y	
PM 10c-1-2	N5°26.542', E117°43.575'	121008			Single	281008	Box 6 Nala	Y	
PM 10c-1-3	N5°26.542', E117°43.575'	121008							
PM 10c-1-4	N5°26.542', E117°43.575'	121008			Duplicate	250609	Box 6 Nala	N	
PM 10c-1-5	N5°26.542', E117°43.575'	121008	6	8.33	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-1	N5°26.436', E117°43.952'	121008			Duplicate	281008	Box 6 Nala	Y (both)	
PM 10c-2-2	N5°26.436', E117°43.952'	121008	4.05	12.35	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-3	N5°26.436', E117°43.952'	121008			Single	281008	Box 6 Nala	Y	
PM 10c-2-4	N5°26.436', E117°43.952'	121008			Duplicate	250609	Box 6 Nala	Y (both)	
PM 10c-2-5	N5°26.436', E117°43.952'	121008							
Mafa 10b-1-1	N5°26.502', E117°44.243'	131008			Single	31108	Box 5 Mafa	Y	
Mafa 10b-1-2	N5°26.502', E117°44.243'	131008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 10b-1-3	N5°26.502', E117°44.243'	131008	4.77	10.48	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10b-1-4	N5°26.502', E117°44.243'	131008	5.83	8.58	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10b-1-5	N5°26.502', E117°44.243'	131008	3.92	12.76	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10b-1-6	N5°26.502', E117°44.243'	131008			Single	31108	Box 5 Mafa	Y	
Mafa 10b-1-7	N5°26.502', E117°44.243'	131008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 10b-1-8	N5°26.502', E117°44.243'	131008							
Mafa 10b-1-9	N5°26.502', E117°44.243'	131008							
Mafa 10b-1-10	N5°26.502', E117°44.243'	131008			Single	31108	Box 5 Mafa	Y	
Mafa 10b-1-11	N5°26.502', E117°44.243'	131008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 10b-1-12	N5°26.502', E117°44.243'	131008	3.81	13.12	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10b-1-13	N5°26.502', E117°44.243'	131008							
PM 10c-3-1	N5°26.482', E117°44.215'	131008			Single	281008	Box 6 Nala	Y	
PM 10c-3-2	N5°26.482', E117°44.215'	131008	5.44	9.19	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-3-3	N5°26.482', E117°44.215'	131008				281008			
PM 10c-3-4	N5°26.482', E117°44.215'	131008							
PM 10b-1-1	N5°26.446', E117°44.960'	131008			Duplicate	250609	Box 6 Nala	Y (both)	
PM 10b-1-2	N5°26.446', E117°44.960'	131008			Single	281008	Box 6 Nala	Y	
PM 10b-1-3	N5°26.446', E117°44.960'	131008			Single	281008	Box 6 Nala	Y	
PM 10b-1-4	N5°26.446', E117°44.960'	131008			Duplicate	281008	Box 6 Nala	Y (both)	
PM 10b-1-5	N5°26.446', E117°44.960'	131008	5.17	9.67	Duplicate	281008	Box 6 Nala	Y (both)	Y
Mafa 10b-2-1	N5°26.308', E117°43.689'	131008	2.48	20.16	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 9-1-1	N5°25.380', E117°45.638'	141008							
Mafa 9-1-2	N5°25.380', E117°45.638'	141008							
Mafa 9-1-3	N5°25.380', E117°45.638'	141008							
Mafa 9-1-4	N5°25.380', E117°45.638'	141008			Single	31108	Box 5 Mafa	Y	
Mafa 9-1-5	N5°25.380', E117°45.638'	141008			Single	31108	Box 5 Mafa	Y	
Mafa 9-1-6	N5°25.380', E117°45.638'	141008	5.32	9.40	Duplicate	291008	Box 6 Mafa	Y (both)	Y

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 9-1-7	N5°25.380', E117°45.638'	141008	6.25	8.00	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 9-2-1	N5°24.795', E117°46.176'	141008							
Mafa 9-2-2	N5°24.795', E117°46.176'	141008	5.76	8.68	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 9-2-3	N5°24.795', E117°46.176'	141008	5.94	8.42	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10a-1-1	N5°24.656', E117°46.446'	141008							
Mafa 10a-1-2	N5°24.656', E117°46.446'	141008			Single	31108	Box 5 Mafa	Y	
Mafa 10a-1-3	N5°24.656', E117°46.446'	141008	2.22	22.52	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10a-1-4	N5°24.656', E117°46.446'	141008							
Mafa 10a-2-1	N5°24.567', E117°46.544'	141008	4.08	12.25	Duplicate		Box 6 Mafa	Y (both)	Y
Mafa 9-3-1	N5°24.907', E117°47.276'	141008			Duplicate		Box 6 Mafa	Y (both)	
Mafa 10a-3-1	N5°25.066', E117°47.016'	141008			Single	31108	Box 5 Mafa	Y	
Mafa 10a-3-2	N5°25.066', E117°47.016'	141008			Single	31108	Box 5 Mafa	Y	
Popy 10a-1-1	N5°24.879', E117°46.245'	141008	5.62	8.90	Duplicate	281008	Box 7 Mafa		Y
Mafa 10a-4-1	N5°25.599', E117°45.009'	141008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 10a-4-2	N5°25.599', E117°45.009'	141008			Single	31108	Box 5 Mafa	Y	
Mafa 10a-4-3	N5°25.599', E117°45.009'	141008	3.59	13.93	Duplicate		Box 6 Mafa	Y (both)	Y
Mafa 10a-4-4	N5°25.599', E117°45.009'	141008	4.31	11.60	Duplicate		Box 6 Mafa	Y (both)	Y
Mafa 10a-5-1	N5°25.141', E117°47.050'	141008			Duplicate		Box 6 Mafa	Y (both)	
Mafa 10a-6-1	N5°25.038', E117°46.996'	141008	3.8	13.16	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10a-7-1	N5°24.828', E117°46.938'	141008	4.44	11.26	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10a-7-2	N5°24.828', E117°46.938'	141008	4.37	11.44	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Popy 10a-2-1	N5°24.932', E117°46.156'	141008	6.65	7.52	Duplicate	281008	Box 7 Mafa		Y
PM 10c-4-1	N5°26.154', E117°42.986'	151008			Single	281008	Box 6 Nala	Y	
PM 10c-4-2	N5°26.154', E117°42.986'	151008			Single	281008	Box 6 Nala	Y	
PM 10c-4-3	N5°26.154', E117°42.986'	151008			Duplicate	300609	Box 6 Nala	Few both	
PM 10c-4-4	N5°26.154', E117°42.986'	151008	5.44	9.19	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-4-5	N5°26.154', E117°42.986'	151008	6.13	8.16	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-6	N5°26.436', E117°43.952'	151008			Duplicate	250609	Box 6 Nala	Few both	
PM 10c-2-7	N5°26.436', E117°43.952'	151008	4.05	12.35	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-8	N5°26.436', E117°43.952'	151008	3.54	14.12	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-9	N5°26.436', E117°43.952'	151008			Duplicate	250609	Box 6 Nala	N	
PM 10c-2-10	N5°26.436', E117°43.952'	151008			Duplicate	300609	Box 6 Nala	Y (Only A)	
PM 10c-2-11	N5°26.436', E117°43.952'	151008			Single	281008	Box 6 Nala	Y	
PM 10c-2-12	N5°26.436', E117°43.952'	151008	3.74	13.37	Duplicate	281008	Box 6 Nala	Y (both)	Y
PM 10c-2-13	N5°26.436', E117°43.952'	151008			Duplicate	281008	Box 6 Nala	Y (both)	
Mafa 10c-1-1	N5°25.930', E117°44.276'	151008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 10c-1-2	N5°25.930', E117°44.276'	151008	6.79	7.36	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10c-1-3	N5°25.930', E117°44.276'	151008	4.71	10.62	Duplicate	291008	Box 6 Mafa	Y (both)	Y
Mafa 10c-1-4	N5°25.930', E117°44.276'	151008			Single	31108	Box 5 Mafa	Y	
Mafa OutS-1-1	N5°25.425', E117°44.122'	151008							
Mafa OutS-1-2	N5°25.425', E117°44.122'	151008							
Mafa OutS-1-3	N5°25.425', E117°44.122'	151008	3.74	13.37	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa OutS-1-4	N5°25.425', E117°44.122'	151008							
Mafa OutS-1-5	N5°25.425', E117°44.122'	151008	4.76	10.50	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa OutS-1-6	N5°25.425', E117°44.122'	151008	3.29	15.20	Single	31108	Box 5 Mafa	Y	Y
PM 10c-4-6	N5°26.154', E117°42.986'	151008			Duplicate	281008	Box 6 Nala	Y (both)	
Mafa 9-4-1	N5°28.988', E117°52.491'	151008							
Mafa 9-4-2	N5°28.988', E117°52.491'	151008			Single	31108	Box 5 Mafa	Y	
Mafa 9-4-3	N5°28.988', E117°52.491'	151008	5.21	9.60	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-1-1	N5°29.933', E117°55.039'	151008			Duplicate	250609	Box 7 Mafa	Few both	N
Mafa 8-1-2	N5°29.933', E117°55.039'	151008	4.28	11.68	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-1-3	N5°29.933', E117°55.039'	151008			Duplicate	250609	Box 7 Mafa	Y	
Mafa 8-1-4	N5°29.933', E117°55.039'	151008	5	10.00	Duplicate	301008	Box 6 Mafa	Y (FEW B)	Y
Mafa 8-1-5	N5°29.933', E117°55.039'	151008							
Mafa 8-1-6	N5°29.933', E117°55.039'	151008	3.82	13.09	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-1-7	N5°29.933', E117°55.039'	151008			Duplicate	250609	Box 7 Mafa	Y (only B)	
Mafa 8-1-8	N5°29.933', E117°55.039'	151008	4.95	10.10	Duplicate	301008	Box 6 Mafa	Y (FEW B)	Y

Appendix One: *Nasalis larvatus* and *Macaca fascicularis* samples database

Name	Location GPS coordinates	Collection date	Weight	Conversion factor (to obtain epg faeces)	Extracted	Extraction date	Stored at -70°C MWB lab	(mt)DNA presence	Sucrose Flotation
Mafa 8-1-9	N5°29.933', E117°55.039'	151008			Duplicate	250609	Box 7 Mafa	Y (both)	
Mafa 8-1-10	N5°29.933', E117°55.039'	151008	4.95	10.10	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-1-11	N5°29.933', E117°55.039'	151008			Duplicate	250609	Box 7 Mafa	Y (both)	
Mafa 8-1-12	N5°29.933', E117°55.039'	151008			Duplicate	260609	Box 7 Mafa	N	
Mafa 8-2-1	N5°29.730', E117°55.211'	151008	5.16	9.69	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-2-2	N5°29.730', E117°55.211'	151008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 8-2-3	N5°29.730', E117°55.211'	151008			Single	31108	Box 5 Mafa	Y	
Mafa 8-3-1	N5°28.666', E117°54.843'	151008			Single	31108	Box 5 Mafa	Y	
Mafa 8-3-2	N5°28.666', E117°54.843'	151008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 8-3-3	N5°28.666', E117°54.843'	151008	5.54	9.03	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-3-4	N5°28.666', E117°54.843'	151008							
Mafa 8-3-5	N5°28.666', E117°54.843'	151008	5.15	9.71	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-3-6	N5°28.666', E117°54.843'	151008							
Mafa 8-3-7	N5°28.666', E117°54.843'	151008							
Mafa 8-3-8	N5°28.666', E117°54.843'	151008	5.4	9.26	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-3-9	N5°28.666', E117°54.843'	151008							
Mafa 8-3-10	N5°28.666', E117°54.843'	151008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 8-3-11	N5°28.666', E117°54.843'	151008			Duplicate	260609	Box 7 Mafa	Y (both)	
Mafa 8-4-1	N5°28.459', E117°55.129'	161008	4.67	10.71	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-5-1	N5°24.413', E117°56.287'	161008			Single	31108	Box 5 Mafa	Y	
Mafa 8-5-2	N5°24.413', E117°56.287'	161008							
Mafa 8-5-3	N5°24.413', E117°56.287'	161008	2.09	23.92	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-6-1	N5°25.990', E117°55.461'	161008	3.99	12.53	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 8-6-2	N5°25.990', E117°55.461'	161008			Single	31108	Box 5 Mafa	Y	
Mafa 9-5-1	N5°29.098', E117°55.139'	171008							
Mafa 9-5-2	N5°29.098', E117°55.139'	171008	3.66	13.66	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 9-5-3	N5°29.098', E117°55.139'	171008	3.47	14.41	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 9-5-4	N5°29.098', E117°55.139'	171008							
Mafa 9-5-5	N5°29.098', E117°55.139'	171008			Single	31108	Box 5 Mafa	Y	
Mafa 9-6-1	N5°28.857', E117°54.941'	171008	5.09	9.82	Duplicate	301008	Box 6 Mafa	Y (Only A)	Y
Mafa 9-6-2	N5°28.857', E117°54.941'	171008							
Mafa 9-6-3	N5°28.857', E117°54.941'	171008	5.54	9.03	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 9-7-1	N5°28.690', E117°54.778'	171008			Single	31108	Box 5 Mafa	Y	
Mafa 9-7-2	N5°28.690', E117°54.778'	171008							
Mafa 9-7-3	N5°28.690', E117°54.778'	171008	4.74	10.55	Duplicate	301008	Box 6 Mafa	Y (both)	Y
Mafa 9-7-4	N5°28.690', E117°54.778'	171008							
Mafa 9-7-5	N5°28.690', E117°54.778'	171008							
Mafa 9-7-6	N5°28.690', E117°54.778'	171008							
Mafa 9-7-7	N5°28.690', E117°54.778'	171008	4.45	11.24	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-7-8	N5°28.690', E117°54.778'	171008			Single	31108	Box 2 Mafa	Y	
Mafa 9-7-9	N5°28.690', E117°54.778'	171008							
Mafa 9-8-1	N5°28.178', E117°54.744'	171008	4.85	10.31	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-8-2	N5°28.178', E117°54.744'	171008							
Mafa 9-8-3	N5°28.178', E117°54.744'	171008							
Mafa 9-8-4	N5°28.178', E117°54.744'	171008	4.15	12.05	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-8-5	N5°28.178', E117°54.744'	171008			Single	31108	Box 2 Mafa	Y	
Mafa 9-8-6	N5°28.178', E117°54.744'	171008							
Mafa 9-9-1	N5°27.504', E117°55.044'	171008	3.48	14.37	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-10-1	N5°27.377', E117°55.069'	171008	5.55	9.01	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-10-2	N5°27.377', E117°55.069'	171008			Duplicate	31108	Box 5 Mafa	Y (both)	
Mafa 9-10-3	N5°27.377', E117°55.069'	171008							
Mafa 9-10-4	N5°27.377', E117°55.069'	171008							
Mafa 9-10-5	N5°27.377', E117°55.069'	171008							
Mafa 9-10-6	N5°27.377', E117°55.069'	171008	5.27	9.49	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-11-1	N5°26.293', E117°54.489'	171008	2.67	18.73	Duplicate	31108	Box 5 Mafa	Y (both)	Y
Mafa 9-11-2	N5°26.293', E117°54.489'	171008			Duplicate	31108	Box 5 Mafa	Y (both)	

APPENDIX TWO

Long-tailed macaque genotypes. Fifteen microsatellites used on the sampled populations at the north and south of the Kinabatangan River

	D8S1106	D7S2204	D12S67	DXS571	D7S503	D11S925	D6S2883	D1S548	D3S1768	D16S420	D1S207	D3S1766	D5S820	D1S550	D6S291
NORTH															
Mafa10a-2-1	140156	237245	150190	133133	137163	236236	117117	209209	189209	216216	132146	221225	192192	158158	207217
Mafa10a-4-4	136136	000000	000000	131141	141141	200200	113117	145167	193197	000000	158166	223223	000000	158158	207207
Mafa10a-6-1	144160	221237	150150	127131	135145	228228	113137	193197	193209	216216	132146	221221	188192	158162	201207
Mafa10a-7-1	148156	245245	150150	131131	135159	238238	115133	197197	213217	210224	132160	233233	196196	162162	207219
Mafa10b-1-5	140140	000000	150150	131133	135135	000000	115115	189197	000000	226232	132132	000000	000000	138138	215215
Mafa10b-2-1	156160	237249	000000	131131	139139	238238	113117	197201	209225	216216	132154	209209	192192	158166	211217
Mafa10c-1-2	152164	237241	150150	127141	135139	228234	117133	201201	201213	216220	160162	229229	192196	162162	207207
Mafa10c-1-3	152152	241241	150150	131131	147163	228228	113117	197197	189221	216222	132154	217221	192192	158162	215219
Mafa10c-1-4	140160	237237	150150	131141	137139	230234	119133	197201	193197	220220	132162	229233	180196	158166	215215
Mafa2-5-1	140156	237245	150150	127133	143147	228230	113115	197197	189209	216222	132162	221225	196196	154158	217217
Mafa2-6-1	144152	000000	198198	119131	139139	232232	109127	197197	217217	222226	156160	217221	000000	158162	000000
Mafa2-7-1	124136	241245	150150	131133	139145	230234	107115	189197	205209	222226	132156	221265	188192	154154	213217
Mafa2-7-6	144152	241245	150150	133135	143147	228236	115119	197197	209213	208224	132162	221221	192196	162162	201201
Mafa2-7-7	136140	253253	210210	131131	139145	226232	115117	197201	217221	206206	156162	000000	188188	154154	217219
Mafa2-8-1	140152	237245	150150	123131	137147	226252	109109	201201	201209	216216	158158	225229	184192	154158	201215
Mafa2-8-2	124124	221241	150150	119131	153153	238238	109113	197201	189205	216220	140160	221221	196196	150158	213213
Mafa4-10-3	160160	221221	000000	131131	145157	232232	107115	193197	209221	216216	132132	217217	188196	158162	201215
Mafa4-10-5	140152	241241	150150	127135	145147	228242	111115	197197	209221	224224	132162	217233	000000	158158	199201
Mafa4-11-1	136152	000000	000000	131131	139147	000000	107115	201201	201201	216216	158160	233233	196196	160160	203213
Mafa4-11-3	124140	249253	150150	127131	135159	238238	107115	000000	205213	216216	156160	000000	188200	160162	215217
Mafa4-4-2	144152	261261	178182	127137	145147	228228	115117	201201	219221	216224	156166	221233	188188	158158	000000
Mafa4-5-1	160168	237241	150182	127131	157161	236238	111115	197197	205205	216216	156160	229233	184196	154154	207215
Mafa4-5-4	160168	237241	150182	127131	135157	238238	107111	197201	205217	216216	132156	229233	184196	158162	211215
Mafa4-7-1	140144	221249	150150	127135	147157	228252	107131	201201	189221	216226	152158	223223	184188	158162	213215
Mafa4-8-1	156160	241245	150150	137137	137145	228238	117119	197201	189217	200202	144158	000000	180196	150158	207215
Mafa4-8-3	152156	241245	150178	135135	155185	252254	111117	197201	205205	216224	158158	221229	188196	158158	213217
Mafa4-8-4	000000	000000	000000	133135	141141	228228	107107	201201	000000	000000	146146	000000	000000	000000	201201
Mafa4-9-1	136140	241245	150180	127141	151157	232238	117139	197197	205217	204216	132132	225233	184188	158158	207211
Mafa4-9-3	152156	221245	150150	119141	139151	238252	113115	193201	197225	216224	140162	217221	192196	154162	205207
Mafa5-1-4	152156	241245	000000	133133	151165	000000	109113	197197	213221	220220	132158	221225	176184	158162	201217
Mafa5-2-3	140140	233233	150150	131143	000000	000000	113115	167167	189205	216216	150160	000000	194194	158158	201201
Mafa5-3-10	000000	000000	150150	129133	157157	000000	109115	000000	000000	242242	132132	217221	000000	154158	000000
Mafa5-3-6	144144	237237	150150	123129	135147	000000	115115	201201	000000	000000	158158	000000	192192	142158	203207
Mafa5-4-2	140140	000000	150178	129129	135147	236236	113133	000000	221221	000000	134134	000000	196196	000000	201201
Mafa5-4-3	000000	000000	150150	133137	135135	000000	113115	209209	000000	000000	140140	000000	184184	138138	211213

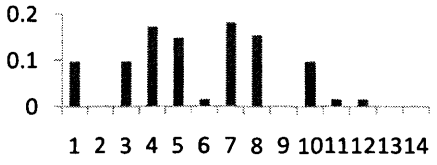
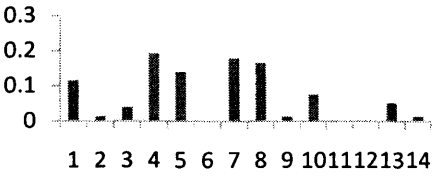
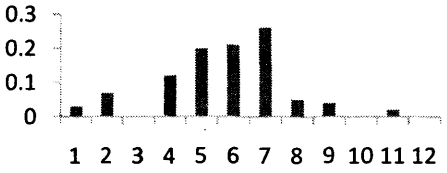
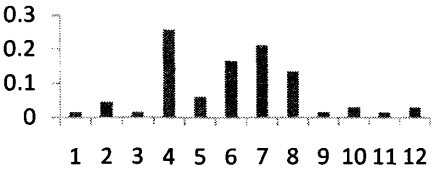
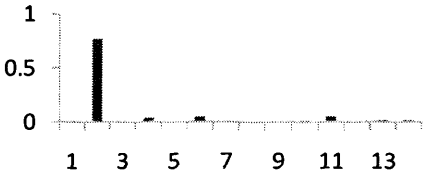
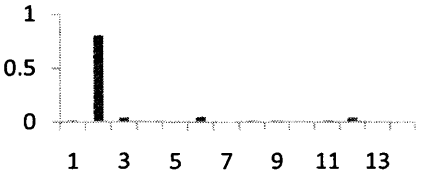
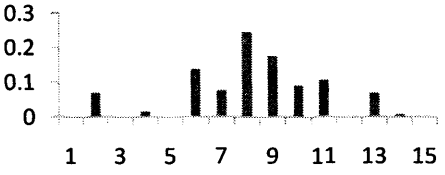
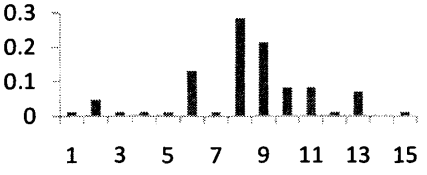
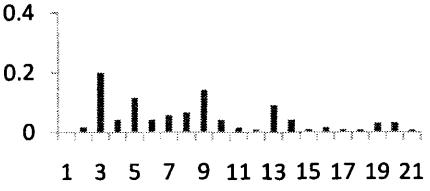
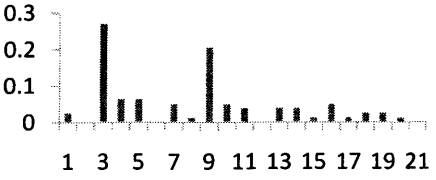
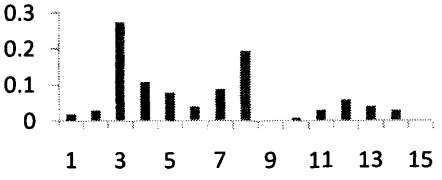
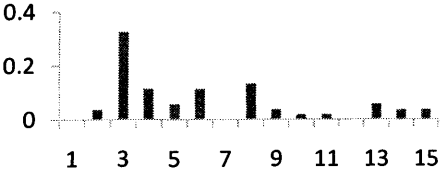
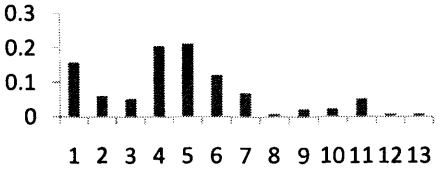
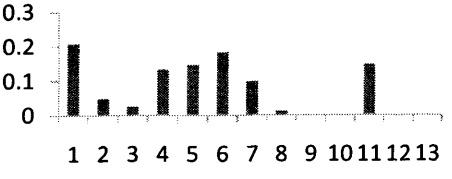
	D8S1106	D7S2204	D12S67	DXS571	D7S503	D11S925	D6S2883	D1S548	D3S1768	D16S420	D1S207	D3S1766	D5S820	D1S550	D6S291
Mafa5-6-1	140156	233245	150150	000000	000000	000000	113117	000000	189189	226226	132132	217217	200200	162162	203211
Mafa5-7-1	140156	237241	150150	133133	145175	228228	107113	197197	221225	226242	000000	217217	200200	158158	207215
Mafa5-8-2	000000	237237	000000	133133	000000	000000	107115	000000	000000	216216	152160	000000	000000	158158	201201
Mafa5-8-5	140152	237237	000000	133133	137139	228238	115133	197201	189205	224224	154158	229229	000000	158162	201201
Mafa5-9-3	144144	213233	000000	135141	179179	000000	107133	193201	213213	216220	132152	209209	188192	154162	211213
Mafa5-9-5	136136	000000	150150	127135	135135	000000	107109	000000	201201	216216	144154	221221	184184	162162	201201
Mafa7-2-1	156160	241245	150206	135135	135143	228254	107113	193197	205217	212220	140152	229229	184192	158162	201215
Mafa7-3-1	136136	000000	150150	129137	000000	236236	111115	145167	209209	000000	132160	000000	000000	000000	207207
Mafa7-4-10	136152	000000	150150	127131	151157	238246	113127	000000	189189	198220	160162	233233	188192	160160	207217
Mafa7-4-2	152160	000000	202202	141141	131139	228228	117127	197197	201209	216216	132162	221221	000000	000000	207215
Mafa7-4-5	136144	245245	150202	131133	147147	232232	113119	197197	189213	198220	160162	265265	184192	158162	201201
Mafa7-5-1	156160	241241	216216	137137	135157	236246	115117	197201	213217	224226	148158	000000	196196	154160	201203
Mafa7-5-2	144156	241253	000000	133133	147157	238252	119119	201201	205209	216244	132148	225229	188188	158158	207215
Mafa7-5-5	144156	000000	000000	133137	147157	238252	119119	197201	205209	216216	132148	000000	000000	158158	207215
Mafa7-5-7	124124	213213	150150	129133	000000	230230	111115	000000	209209	220220	132158	000000	000000	000000	207207
Mafa7-6-2	000000	233233	150150	127127	179179	226246	107115	000000	189189	210210	154154	000000	000000	158158	000000
Mafa7-6-3	124140	241241	150150	127137	000000	000000	113119	201201	221221	000000	152162	000000	000000	158158	213213
Mafa7-7-2	156156	000000	000000	119141	143147	228228	107115	197197	217223	220220	156158	000000	000000	000000	000000
Mafa7-7-5	124124	000000	150150	119119	143159	000000	107113	197197	000000	222222	132132	000000	000000	160160	201201
Mafa7-8-1	152156	237237	150150	129135	143143	000000	113117	000000	189221	226226	146158	225225	188188	000000	000000
Mafa8-1-10	144152	233233	150150	129137	135175	230230	107131	205205	189189	198198	158162	000000	184184	000000	213215
Mafa8-1-2	144152	245245	150202	129137	135175	230230	107131	197197	189189	198220	158162	000000	184184	146146	000000
Mafa8-1-8	156164	233233	150150	127129	135147	256256	111121	000000	213213	216226	140154	221229	196196	146146	209209
Mafa8-2-1	140160	221245	174186	119119	135159	228238	113133	197201	209213	220226	132132	217229	192200	154160	203217
Mafa8-3-10	152152	245245	182182	131131	139139	000000	107113	197197	189189	000000	154154	000000	000000	000000	000000
Mafa8-3-3	148152	237245	178202	131131	135167	228234	107113	197201	189201	198232	146154	225233	184188	158166	213215
Mafa8-3-5	124144	245249	000000	131137	135135	228254	113113	197197	205209	212222	154162	221229	192196	162162	201215
Mafa8-3-8	124144	245249	150150	131137	135151	228256	113113	197197	205209	212222	154162	221229	192196	162162	201215
Mafa8-4-1	140152	245245	150150	133137	131141	238254	113115	197201	189205	222222	132160	221225	192200	162162	201215
Mafa8-5-1	124156	000000	150150	127135	159175	230230	107115	197197	213221	202220	132156	000000	188200	158182	213215
Mafa8-5-3	144144	233233	150150	119127	135147	228228	109117	197197	213213	226232	132162	229229	192192	158158	201207
SOUTH															
Mafa1-3-3	144160	233233	000000	131131	135147	000000	117133	000000	000000	216216	132132	000000	000000	000000	213215
Mafa1-3-6	144160	233233	000000	131131	135135	000000	117133	201201	209221	216216	132132	000000	000000	000000	213215
Mafa1-3-8	144144	233233	150150	131131	135135	000000	117133	000000	209209	216216	132132	000000	000000	150160	000000
Mafa1-3-9	152152	233245	000000	133133	137137	000000	107119	197197	000000	226226	132162	219219	184184	154162	215215
Mafa1-5-1	156160	233245	150150	131133	135135	228230	107107	193197	205213	216220	152156	225225	196196	154162	213215
Mafa1-5-2	156160	233245	150150	131133	135135	228230	107107	193197	205213	216220	152156	225225	196196	154162	213215
Mafa1-5-3	156160	233245	150150	131133	135135	228228	107107	193197	213213	000000	152156	000000	196196	154162	213213

Appendix two: Long-tailed macaque genotypes.

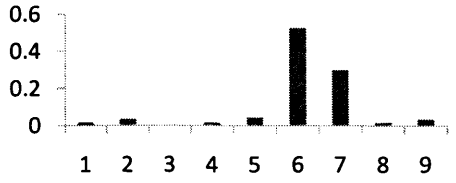
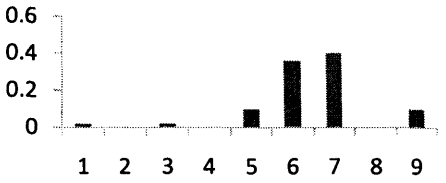
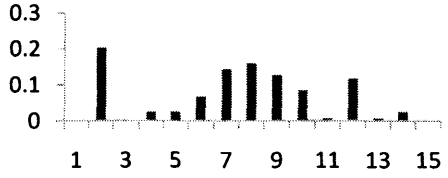
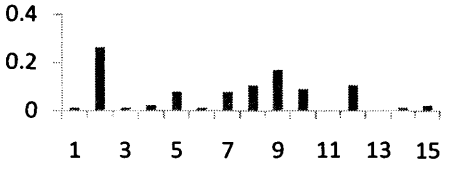
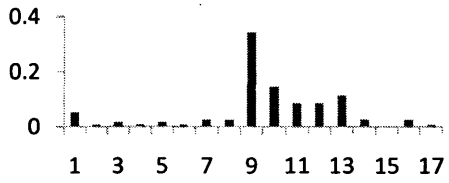
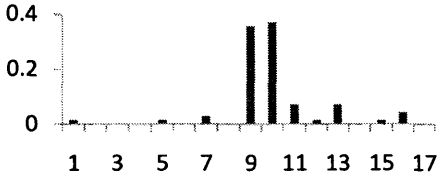
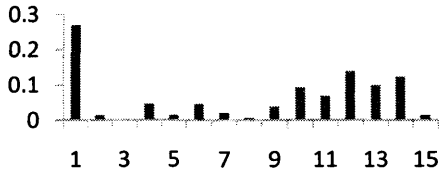
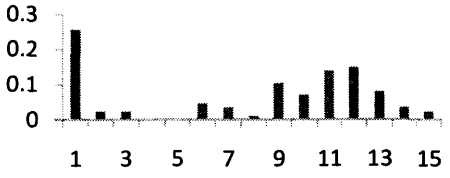
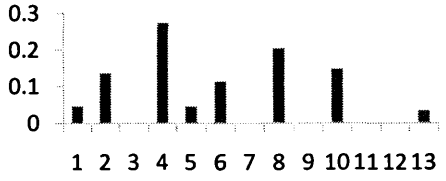
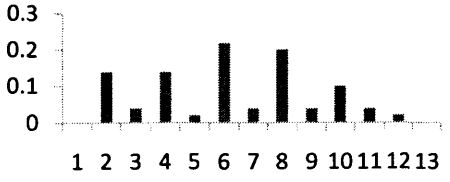
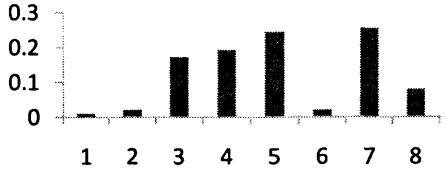
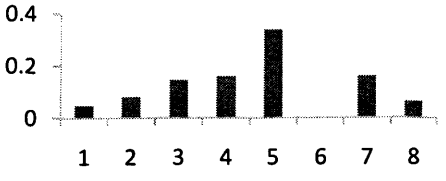
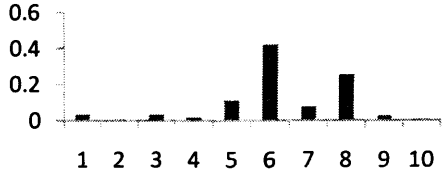
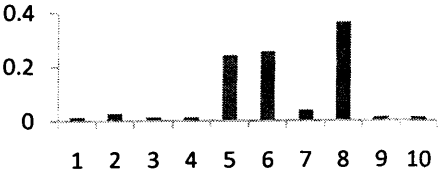
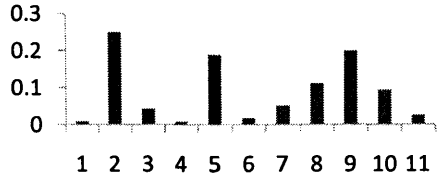
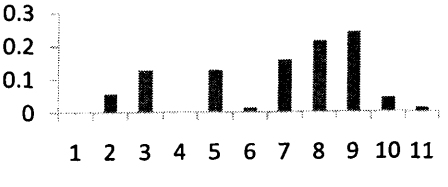
	D8S1106	D7S2204	D12S67	DXS571	D7S503	D11S925	D6S2883	D1S548	D3S1768	D16S420	D1S207	D3S1766	D5S820	D1S550	D6S291
Mafa1-6-2	136140	241249	190194	141141	135153	228228	117119	187197	197201	210216	156158	217225	196196	138158	201215
Mafa1-7-1	140152	000000	150150	133133	159159	230230	117119	000000	213213	242242	148166	225229	000000	000000	201203
Mafa3-10-6	152152	000000	150150	135135	137137	228228	117117	000000	000000	220220	132132	233233	192192	158182	203215
Mafa3-14-1	152156	237241	150150	131131	151157	256256	107133	197201	209221	216222	156160	229229	192200	158162	201211
Mafa3-15-3	140152	241245	150150	127131	139143	240240	107133	193197	189221	220226	146156	217221	184184	154162	211217
Mafa3-16-5	124124	213221	150150	127137	175175	000000	115133	209209	205209	238242	160166	000000	000000	162162	211213
Mafa3-17-2	136144	245249	202206	133141	139153	000000	115115	197197	209217	220220	132162	217223	184192	158158	207219
Mafa3-2-1	144144	000000	150150	121127	000000	000000	109115	000000	000000	000000	132132	000000	000000	162162	213215
Mafa3-4-1	124124	257257	150150	131141	000000	228228	107107	201201	189193	216216	148150	237237	000000	158158	213215
Mafa3-5-4	144156	241241	150150	127133	000000	234238	000000	201201	000000	000000	132158	000000	192192	000000	000000
Mafa3-7-2	124140	249253	138150	131139	147147	000000	115117	197197	217221	226226	158162	217233	196200	162162	000000
Mafa3-8-1	124124	000000	150150	127137	135165	226226	119133	201201	221221	210220	152154	225225	188188	158162	211211
Mafa3-9-8	000000	000000	150150	135135	135135	000000	000000	193201	189189	000000	132148	000000	000000	160160	000000
Mafa6-10-10	172172	241241	174182	131131	147147	232254	113113	201201	189209	220222	158160	231231	192192	162162	207215
Mafa6-1-2	152152	000000	000000	135135	139147	232232	107107	197197	189217	222222	154154	000000	000000	154158	000000
Mafa6-12-1	000000	000000	150150	000000	143151	000000	113115	000000	205205	000000	158158	000000	176196	146162	207207
Mafa6-13-1	144156	233245	000000	133133	151179	000000	107107	000000	197197	216216	156156	221221	192192	142154	000000
Mafa6-14-1	144156	245249	150150	127131	147147	234234	115115	197197	191209	216220	158158	217217	184192	154162	207207
Mafa6-15-1	140156	249249	206206	127133	139159	228228	113113	000000	233233	216222	132138	000000	176176	000000	211213
Mafa6-2-10	152152	237245	000000	127133	143157	000000	107117	000000	189189	220220	132154	000000	188188	154158	211213
Mafa6-4-3	156160	233241	150150	131133	135157	242254	117117	201201	189213	206216	132158	217221	188192	158158	211213
Mafa6-4-7	156174	233261	150150	131137	135161	234238	119121	197201	189213	000000	156160	000000	192192	158162	203213
Mafa6-9-1	140140	233241	150150	119125	147147	000000	109119	000000	221225	220220	158158	229229	000000	142162	203217
Mafa9-10-1	144158	000000	150150	127167	163163	000000	109113	000000	189189	220220	146152	229233	180180	158158	203207
Mafa9-1-6	000000	245245	174182	119133	129147	228230	113117	201209	197185	216220	152156	227227	184192	154154	201203
Mafa9-1-7	172172	245245	174182	133133	129151	228230	113119	209209	197197	216220	156156	000000	184192	154154	203203
Mafa9-4-3	000000	000000	150150	123129	147153	000000	133133	000000	213213	220220	138146	229229	000000	154154	213215
Mafa9-5-2	140152	249249	150150	119127	147167	000000	111117	145201	189217	216220	152158	225241	200200	158162	207211
Mafa9-5-3	152156	000000	150150	137141	143147	228238	111133	000000	189189	000000	132146	000000	180180	000000	000000
Mafa9-6-3	140156	221225	150150	117119	147147	234234	115133	000000	189221	216216	132132	000000	000000	162162	213215
Mafa9-7-3	136140	233233	150150	133141	135135	238238	109117	197201	189217	220220	152160	221233	184188	154158	203217
Mafa9-7-7	132156	221249	150182	131131	135167	260260	113119	000000	189189	216216	132154	221221	188192	162162	215215
Mafa9-8-1	124140	241245	150150	127137	137139	228228	115115	000000	205213	220224	158160	000000	188188	154162	215215
Mafa9-8-4	124152	237237	150150	137137	145163	238238	107133	000000	213217	220220	154160	229229	180192	158158	209211
Mafa9-9-1	140140	233241	150150	131135	135163	246254	115117	201201	189193	198216	132152	225225	188192	154166	207211
MafaOutS-1-3	140140	265265	000000	131131	000000	000000	113113	000000	213217	000000	134134	000000	192192	162162	000000

APPENDIX THREE

Long-tailed macaque allele frequency distribution by locus and population

Locus	North	South
D8S1106		
D7S2204		
D12S67		
DXS571		
D7S503		
D11S925		
D6S2883		

Appendix Three: Allele frequency distributions by locus and population of long-tailed macaques

Locus	North	South
D1S548		
D3S1768		
D16S420		
D1S207		
D3S1766		
D5S820		
D1S550		
D6S291		

APPENDIX FOUR

Linkage disequilibrium (LD) tests of loci screened in long-tailed macaques. LD was measured using the correlation coefficient. The significance at $P < 0.05$ by pair of loci and for each population is represented with *. NS = non significant.

Locus1	Locus 2	North	South	Locus1	Locus 2	North	South	Locus1	Locus 2	North	South
D8S1106	D7S2204	NS	*	D12S67	D3S1766	NS	*	D6S2883	D3S1768	NS	NS
D8S1106	D12S67	NS	*	D12S67	D5S820	NS	NS	D6S2883	D16S420	NS	NS
D8S1106	DXS571	NS	*	D12S67	D1S550	NS	NS	D6S2883	D1S207	NS	*
D8S1106	D7S503	*	NS	D12S67	D6S291	NS	NS	D6S2883	D3S1766	NS	NS
D8S1106	D11S925	NS	NS	DXS571	D7S503	NS	NS	D6S2883	D5S820	NS	NS
D8S1106	D6S2883	NS	NS	DXS571	D11S925	NS	NS	D6S2883	D1S550	NS	NS
D8S1106	D1S548	NS	NS	DXS571	D6S2883	NS	NS	D6S2883	D6S291	NS	NS
D8S1106	D3S1768	*	NS	DXS571	D1S548	*	NS	D1S548	D3S1768	NS	NS
D8S1106	D16S420	NS	NS	DXS571	D3S1768	NS	NS	D1S548	D16S420	NS	NS
D8S1106	D1S207	NS	NS	DXS571	D16S420	NS	NS	D1S548	D1S207	NS	NS
D8S1106	D3S1766	NS	NS	DXS571	D1S207	*	NS	D1S548	D3S1766	NS	*
D8S1106	D5S820	NS	NS	DXS571	D3S1766	NS	NS	D1S548	D5S820	NS	*
D8S1106	D1S550	NS	NS	DXS571	D5S820	NS	NS	D1S548	D1S550	NS	NS
D8S1106	D6S291	*	NS	DXS571	D1S550	NS	NS	D1S548	D6S291	NS	NS
D7S2204	D12S67	NS	NS	DXS571	D6S291	NS	NS	D3S1768	D16S420	*	NS
D7S2204	DXS571	NS	*	D7S503	D11S925	*	NS	D3S1768	D1S207	NS	*
D7S2204	D7S503	NS	*	D7S503	D6S2883	NS	NS	D3S1768	D3S1766	*	NS
D7S2204	D11S925	NS	NS	D7S503	D1S548	*	*	D3S1768	D5S820	NS	NS
D7S2204	D6S2883	NS	NS	D7S503	D3S1768	NS	NS	D3S1768	D1S550	*	*
D7S2204	D1S548	*	*	D7S503	D16S420	NS	*	D3S1768	D6S291	NS	NS
D7S2204	D3S1768	*	NS	D7S503	D1S207	NS	*	D16S420	D1S207	NS	*
D7S2204	D16S420	NS	*	D7S503	D3S1766	*	*	D16S420	D3S1766	*	*
D7S2204	D1S207	NS	*	D7S503	D5S820	*	*	D16S420	D5S820	*	NS
D7S2204	D3S1766	*	NS	D7S503	D1S550	NS	NS	D16S420	D1S550	*	NS
D7S2204	D5S820	*	*	D7S503	D6S291	NS	NS	D16S420	D6S291	NS	NS
D7S2204	D1S550	*	*	D11S925	D6S2883	NS	NS	D1S207	D3S1766	NS	NS
D7S2204	D6S291	NS	NS	D11S925	D1S548	*	NS	D1S207	D5S820	NS	*
D12S67	DXS571	NS	NS	D11S925	D3S1768	*	NS	D1S207	D1S550	NS	NS
D12S67	D7S503	NS	NS	D11S925	D16S420	*	*	D1S207	D6S291	NS	NS
D12S67	D11S925	NS	NS	D11S925	D1S207	NS	NS	D3S1766	D5S820	*	*
D12S67	D6S2883	NS	NS	D11S925	D3S1766	*	*	D3S1766	D1S550	NS	NS
D12S67	D1S548	NS	NS	D11S925	D5S820	NS	NS	D3S1766	D6S291	NS	NS
D12S67	D3S1768	NS	*	D11S925	D1S550	*	NS	D5S820	D1S550	*	NS
D12S67	D16S420	NS	NS	D11S925	D6S291	NS	NS	D5S820	D6S291	*	NS
D12S67	D1S207	NS	NS	D6S2883	D1S548	NS	NS	D1S550	D6S291	*	NS

APPENDIX FIVE

Null allele frequencies in loci screened in long-tailed macaque samples and proboscis monkeys.

Table 19 Long-tailed macaque null allele frequencies

	Population	
Locus	North	South
D8S1106	0.253	0.359
D7S2204	0.460	0.361
D12S67	0.774	0.811
DXS571	0.327	0.368
D7S503	0.264	0.380
D11S925	0.459	0.615
D6S2883	0.115	0.352
D1S548	0.600	0.560
D3S1768	0.276	0.339
D16S420	0.474	0.543
D1S207	0.231	0.302
D3S1766	0.477	0.640
D5S820	0.458	0.576
D1S550	0.561	0.401
D6S291	0.317	0.229

Table 20. Proboscis monkey null allele frequencies

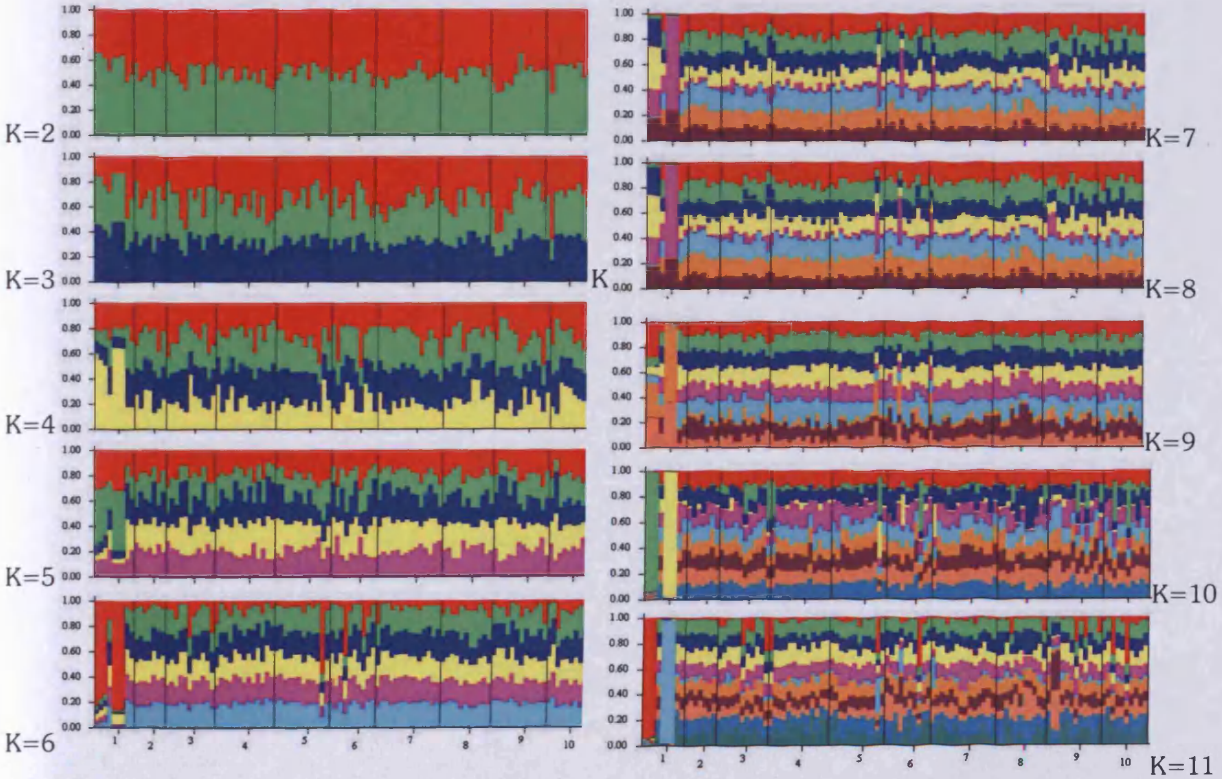
	Population	
Locus	North	South
NlP1A6	0.299	0.357
NlP3B2	0.381	0.400
NlP4C11	0.310	0.462
NlE10	0.600	0.737
NlP1C5	0.378	0.097
NlD10	0.206	0.104
NlP2D6	0.364	0.040
NlP2F3	0.500	0.368

APPENDIX SIX

Number of genetically differentiated clusters (subpopulations, K=2) of long-tailed macaques determined by STRUCTURE, comparing the 10 Lots of the LKWS as unique sampled subpopulations. K values were set between 1-11 using 5 independent runs of 100,000 iterations (plus a burn-in of 20,000 iterations) for each K value. Runs were performed with the "admixture model" and the "correlated allele frequency" model without prior information and with an initial alpha value set to 1.0. Null alleles were considered as recessive to all other alleles. Individuals were assigned to one cluster if their proportion of membership (q_i) to that cluster was equal to or larger than 0.600, the individuals with maximum inferred ancestry < 0.6 were not assigned to any group, which was the case for all sampled populations.

K	1	2	3	4	5	6	7	8	9	10	11
LnP(X|K)	-6340	***-6120***	-6168	-6503	-6228	-6322	-6476	-6698	-6423	-7109	-7681

Given population	Inferred clusters *q*I	*q*II	Given population	Inferred clusters *q*I	*q*II
Lot 1	0.415	0.585	Lot 6	0.513	0.487
Lot 2	0.508	0.492	Lot 7	0.522	0.478
Lot 3	0.491	0.509	Lot 8	0.499	0.501
Lot 4	0.532	0.468	Lot 9	0.525	0.475
Lot 5	0.478	0.522	Lot 10	0.48	0.52



APPENDIX SEVEN

Proboscis monkey genotypes. Eight microsatellites used on the sampled populations at the north and south of the Kinabatangan River.

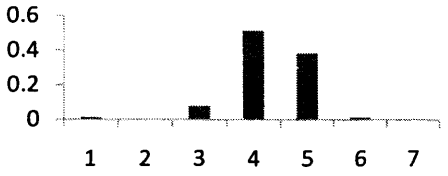
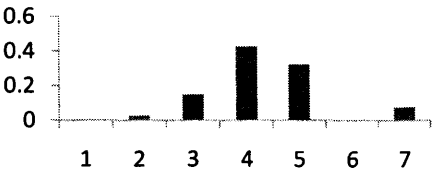
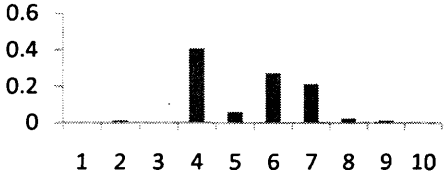
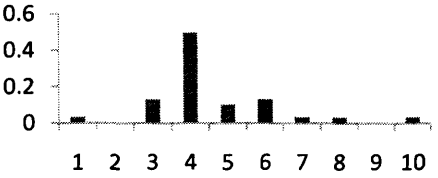
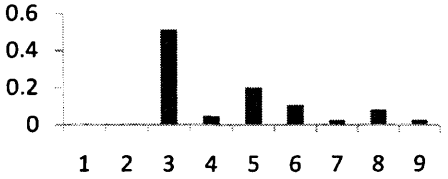
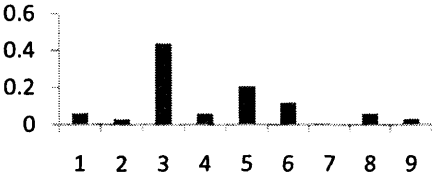
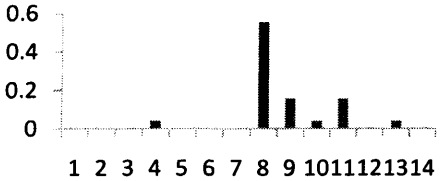
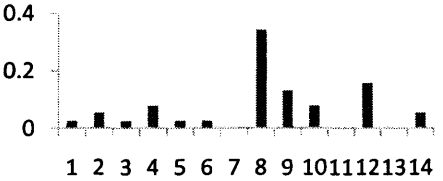
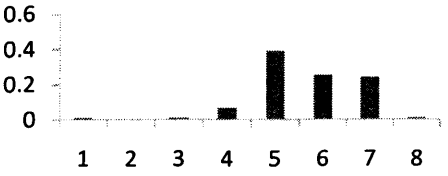
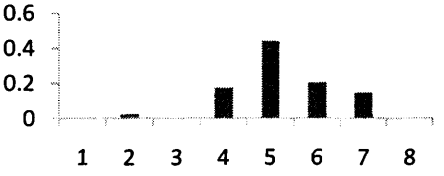
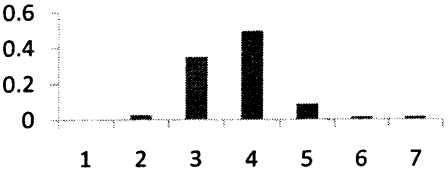
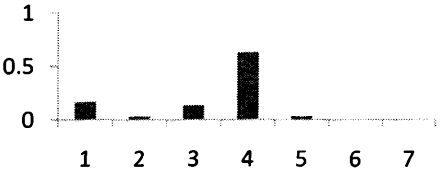
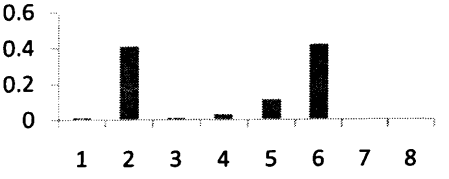
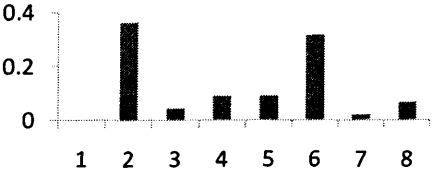
	P1A6	P3B2	P4C11	E10	P1C5	D10	P2D6	P2F3
NORTH								
PM_10b-1-1_M1-1	152155	000000	241241	000000	201221	179183	156156	177185
PM_10b-1-4_M1-1	155155	162164	241241	000000	197197	181183	160160	177185
PM_10c-2-1_M1-1	155155	162162	241241	000000	189189	179183	000000	183183
PM_10c-2-12_M1-1	155158	162166	241241	167169	197205	181183	150158	183183
PM_10c-2-13_M1-1	152155	158166	249251	167167	201205	181183	158160	183183
PM_10c-2-7_M1-1	155158	165165	241241	169169	000000	181183	160160	183185
PM_10c-2-8_M1-1	155158	162162	241241	167167	197205	181183	160160	183183
PM_10c-3-2_M1-1	000000	166166	241241	155169	197197	183183	160160	183183
PM_10c-4-6_M1-1	155158	164166	241251	155167	197205	000000	150158	177183
PM_2-5-1_M1-1	155155	162164	249253	167167	197197	000000	150156	177187
PM_4-10-1_M1-1	000000	162165	241255	167167	201201	000000	146160	177183
PM_4-11-1_M1-1	155155	162166	241255	167175	197197	181183	150160	183183
PM_4-11-9_M1-1	146158	162166	241251	175175	197201	181183	150160	183183
PM_4-1-3_M1-1	155158	162166	241251	167167	197205	000000	150150	177183
PM_4-14-4_M1-1	155158	164165	241249	167175	197205	181183	150158	177183
PM_4-15-3_M1-1	155158	162165	249251	167175	197205	181183	150158	177185
PM_4-16-4_M1-1	155155	162165	245249	205205	197201	181183	150160	183183
PM_4-17-4_M1-1	155158	162166	241245	167167	000000	181183	150160	177183
PM_4-2-16_M1-1	155158	165166	241255	167167	201205	181185	150160	177183
PM_4-2-7_M1-1	155158	000000	241255	167175	197201	000000	158160	183183
PM_4-4-7_M1-1	155158	166166	241251	167173	197201	000000	150160	177177
PM_4-8-1_M1-1	158158	165170	241249	173173	201205	183189	150150	185185
PM_4-9-5_M1-1	152158	162162	241241	169205	000000	000000	160160	000000
PM_5-1-10_M1-1	155155	165165	241251	167167	177197	183183	150150	177177
PM_5-1-2_M1-1	155158	162162	241259	000000	197201	183187	150150	177183
PM_5-1-6_M1-1	155155	165168	241255	167167	197205	181185	150160	183185
PM_5-3-11_M1-1	155158	162165	245249	167167	205205	181185	150160	185185
PM_5-3-7_M1-1	158158	165166	241249	169175	205205	181183	158160	183183
PM_5-4-7_M1-1	000000	162166	249251	167167	197197	183183	150160	183183
PM_5-5-1_M1-1	158158	162162	241241	000000	201205	181185	150158	177181
PM_5-5-5_M1-1	155158	165165	000000	000000	189189	181185	150154	000000
PM_5-6-10_M1-1	158161	000000	241241	155167	201205	181183	150150	177177
PM_5-6-13_M1-1	000000	166166	241249	000000	201201	183183	150150	177185
PM_5-6-4_M1-1	000000	165166	241241	000000	000000	181185	150160	177177
PM_5-6-9_M1-1	152158	165165	241249	000000	000000	000000	150160	177177
PM_7-11-26_M1-1	155158	162162	245251	169169	197205	181183	160160	183183
PM_7-11-4_M1-1	155155	162165	241255	167167	197205	183183	160160	177183
PM_7-11-8_M1-1	155158	162164	241259	167169	000000	000000	150160	183183
PM_7-12-1_M1-1	155158	162168	000000	167175	189197	181183	150158	177185

Appendix Seven: Proboscis monkey genotypes

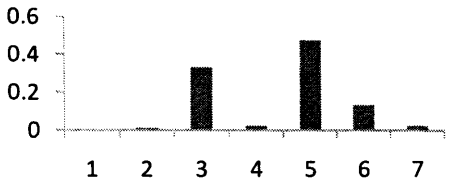
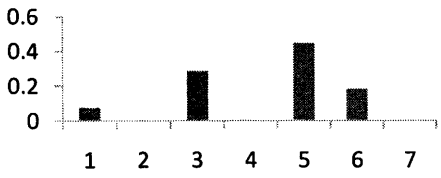
	P1A6	P3B2	P4C11	E10	P1C5	D10	P2D6	P2F3
PM_7-14-1_M1-1	155158	162165	249249	169175	187197	000000	150158	177183
PM_7-2-2_M1-1	155158	162166	253255	175175	000000	181183	150160	183183
PM_7-4-11_M1-1	152158	165165	249249	167167	197197	181183	160160	175187
PM_7-4-5_M1-1	152155	162162	000000	167169	201201	183183	160160	177181
PM_7-7-3_M1-1	000000	162162	241249	000000	201201	181183	150160	000000
PM_7-9-4_M1-1	155155	162165	241249	167167	000000	000000	150160	177183
SOUTH								
PM_1-2-13_M1-1	152158	000000	241241	000000	000000	183183	154156	177177
PM_1-4-15_M1-1	155155	165168	241251	167167	197205	183183	158160	177183
PM_1-4-16_M1-1	155158	000000	000000	211211	189189	000000	160160	177183
PM_1-4-17_M1-1	155176	165165	241241	167167	197201	000000	160160	177183
PM_1-4-8_M1-1	152158	000000	241241	167167	000000	000000	150160	177185
PM_1-5-19_M1-1	155155	164164	241241	169169	197201	000000	156160	000000
PM_3-11-1_M1-1	155158	162172	245255	167173	185205	183183	150156	177185
PM_3-11-4_M1-1	155155	160162	251255	153169	197205	181183	150160	177183
PM_3-5-18_M1-1	152155	162162	249249	169169	197197	179183	150150	183185
PM_3-5-22_M1-1	155155	160162	000000	183183	197201	181183	150150	183185
PM_3-5-37_M1-1	155158	160162	249251	151151	000000	000000	150158	183183
PM_3-5-44_M1-1	155158	162162	249249	167167	197205	181183	150150	183183
PM_3-6-5_M1-1	155155	000000	241241	173173	201201	181185	158160	185185
PM_3-8-28_M1-1	155176	162162	241241	147157	000000	000000	150156	183183
PM_6-3-1_M1-1	158158	000000	237259	000000	197197	177177	160162	000000
PM_6-4-3_M1-1	152158	162165	241249	161167	197197	177183	150150	177183
PM_6-5-1_M1-1	158158	152160	000000	000000	197201	177177	154160	000000
PM_6-6-10_M1-1	149176	162166	000000	183183	197201	183183	150160	177183
PM_6-6-2_M1-1	000000	162162	241249	155167	000000	183183	150158	183183
PM_6-6-6_M1-1	158158	162164	245251	167167	197205	183183	150160	183185
PM_6-7-2_M1-1	000000	000000	000000	183183	189189	183183	166166	173177
PM_6-7-5_M1-1	152152	000000	231231	155155	189189	000000	160166	173173

APPENDIX EIGHT

Proboscis monkey allele frequency distribution by locus and population

Locus	North	South
P1A6		
P3B2		
P4C11		
E10		
P1C5		
D10		
P2D6		

Appendix Eight: Allele frequency distributions by locus and population of proboscis monkeys

Locus	North	South
P2F3		

APPENDIX NINE

Linkage disequilibrium (LD) tests of loci screened in proboscis monkey. LD was measured using the correlation coefficient. The significance at $P < 0.05$ by pair of loci and for each population is represented with *. NS = non significant.

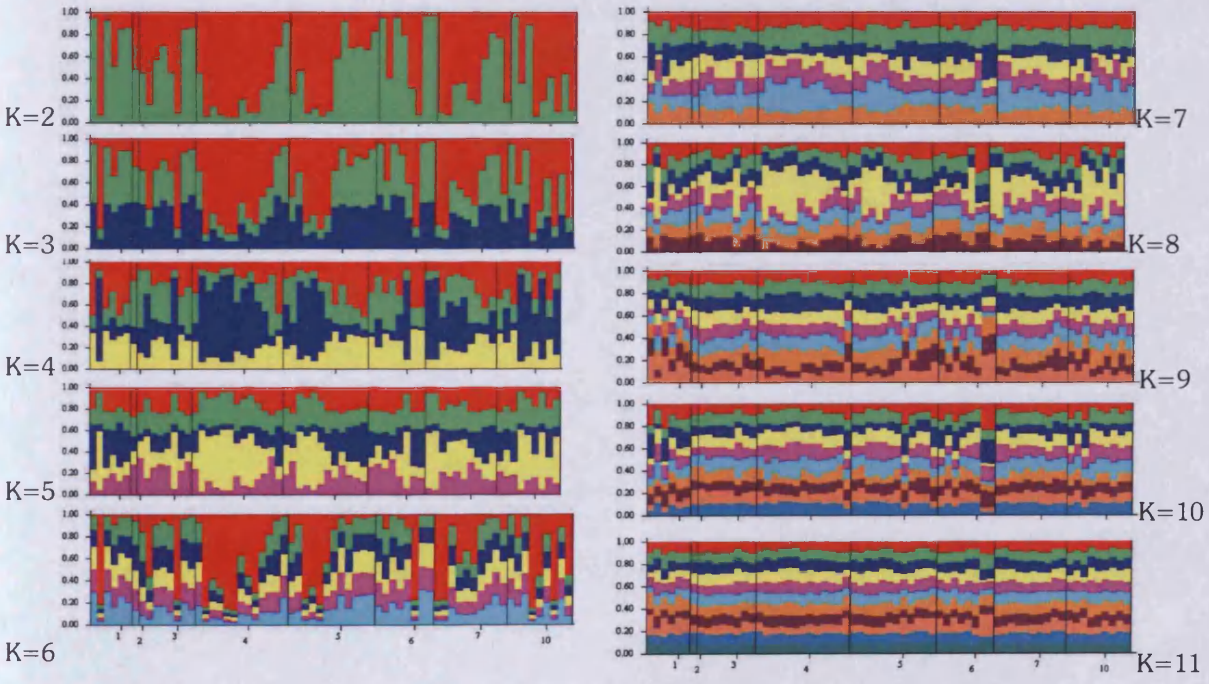
Locus 1	Locus 2	North	South
P1A6	P3B2	NS	NS
	P4C11	NS	NS
	E10	NS	NS
	P1C5	NS	NS
	D10	NS	NS
	P2D6	NS	NS
	P2F3	NS	NS
P3B2	P4C11	NS	*
	E10	NS	NS
	P1C5	NS	NS
	D10	NS	NS
	P2D6	NS	*
	P2F3	NS	*
P4C11	E10	NS	*
	P1C5	NS	*
	D10	NS	NS
	P2D6	NS	NS
	P2F3	NS	NS
E10	P1C5	NS	*
	D10	NS	*
	P2D6	NS	NS
	P2F3	NS	*
P1C5	D10	NS	NS
	P2D6	*	NS
	P2F3	NS	NS
D10	P2D6	NS	NS
	P2F3	NS	NS
P2D6	P2F3	NS	NS

APPENDIX TEN

Number of genetically differentiated clusters (subpopulations, K=2, but note value of K=3) of proboscis monkeys determined by STRUCTURE, comparing the 10 Lots of the LKWS as unique sampled subpopulations. K values were set between 1-11 using 5 independent runs of 100,000 iterations (plus a burn-in of 20,000 iterations) for each K value. Runs were performed with the "admixture model" and the "correlated allele frequency" model without prior information and with an initial alpha value set to 1.0. Null alleles were considered as recessive to all other alleles. Individuals were assigned to one cluster if their proportion of membership (q_i) to that cluster was equal to or larger than 0.600, the individuals with maximum inferred ancestry < 0.6 were not assigned to any group. Lot 1 was assigned to cluster I, while Lot 4 and Lot 10 were assigned to Cluster II.

K	1	2	3	4	5	6	7	8	9	10	11
lnP(X|K)	-1655	**-1623.7**	-1628.2	-1827.2	-1829.6	-1964.4	-1842.1	-1868.7	-1978.8	-2083.2	-2057.8

Given population	Inferred clusters qI	qII	Given population	Inferred clusters qI	qII
Lot 1	0.308	0.692	Lot 5	0.52	0.48
Lot 2	0.35	0.65	Lot 6	0.403	0.597
Lot 3	0.402	0.598	Lot 7	0.589	0.411
Lot 4	0.683	0.317	Lot 10	0.615	0.385



APPENDIX ELEVEN

Mhc-Nala-DRB sequences

>Nala-DRB*Allele1
TCCGGTCGACTGTCCCCCCAGCACGTTTCTTGAAGCAGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGAGTGCG
GTTCCTGCACAGATATATCCATAACCAAGAGGAGGACGCGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGG
AGCTGGGGCGGCCTGACGCCGAGTACTTCAACAGCCAGAAGGACATCCTGGAGCAGAGGCGGGCCGAGGTGGACACAGTG
TGCAGACACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAGGTGATCTAGAAA
>Nala-DRB*Allele2
CCGGTCGACTGTCCCCCCAGCACGTTTCTTGAAGCAGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCG
GTTCCTGCACAGATATATCCATAACCAAGAGGAGGACGCGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGG
AGCTGGGGCGGCGTACCGCCGAGTACTGGAACAGCCAGAAGGACCTCCTGGAGCAGAGGCGGGCCGTGGTGGACACAGT
GTGCAGACACAACTACGGGGTTTTTGAGAGCTTCACAGTGCAGCGGCGAGGTGATCTAGAA

>Nala-DRB*Allele3
CCGGTCGACTGTCCCCCCAGCACGTTTCTTGAAGCAGGATAAGTATGAGTGTCATTTCCTCAACGGGACGGAGCGAGTGCG
GTTCCTGCACAGATATATCCATAACCAAGAGGAGGACGCGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGG
AGCTGGGGCGGCCTGACGCCGAGTACTTCAACAGCCAGAAGGACATCCTGGAGCAGAGGCGGGCCGAGGTGGACACAGTG
TGCAGACACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAGGTGATCTAGA

>Nala-DRB*Allele4
GGTCGACTGTCCCCCCAGCACGTTTCTTGAAGCAGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGAGTGCGGTT
CCTGCACAGGTATATCCATAACCAAGAGGAGGACGCGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGC
TGGGGCGGCGTGCCGCCGAGTACTGGAACAGCCAGAAGGACATCCTGGAGCAGAGGCGGGCCGAGGTGGACACAGTGTG
CAGACACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAGGTGATCTAGA

>Nala-DRB*Allele5
CCGGTCGACTGTCCCCCCAGCACGTTTCTTGAAGCAGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCG
GTTCCTGCACAGATACATCCATAACCAAGAGGAGAACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGG
AGCTGGGGCGGCGTGCCGCCGAGTACTGGAACAGCCAGAAGGACATCCTGGAGCAGAGGCGGGCCGAGGTGGACACAGT
GTGCAGACACAACTACGGGGTTTTTGAGAGCTTCACAGTGCAGCGGCGAGGTGATCTAGAA

APPENDIX TWELVE

Mhc-Mafa-DRB sequences

>Mafa-DRB*Allele1
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAAGA
GGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGG
AACAGTCAGAAGGACATCCTGGAGCAGAAGCGGGCCGGGGTGGACAACTACTGCAGATACAACTACCGGGTTGGTGAGAGC
TTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele2
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGTCGGAGCTGGGGCGGCCTGACGCCGAGAACTG
GAACAGTCGGAAGGACATCCTGGAGCAGAGGCGGGCCGGGGTGGACAACTACTGCAGACACAACTACCGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele3
TTGGAGCAGTCTACATCTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGACACTTCTATAACCAGG
AGGAGTACGTGCGCTATGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACAGCCAGAAGGACATCCTGGAACACAGGCGGGCCGCGGTGGACACCTACTGCAGATACAACTACCGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele4
TTGGAGCAGGTTAAGTATGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGAGAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACAGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACAGCCAGAAGGACTTCCTGGAGGAGAAGCGGAGCCCGGTTGACAACTACTGCAGACACAACTACGGGGTTGTGGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele5
TTGGAGCAGGTTAAGTATGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGAGAGATACTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGTCGGAGCTGGGGCGGCGTAGCGCCCAGTCCTG
GAACGGCCAGAAGGGCATCCTGGAGCAGAGGCGGACCCAGGTGGACACCTACTGCAGACACAACTACGGGGTTGTGGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele6
TTGGAGCAGGTTAAGTATGAGTGTCGTTTCTTCAACAGGACGGAGCGGGTGCGGTTCCTGGAGAGACACTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGTGACCTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCCAGTACTTG
AACAGTCAGAAGGACCTCCTGGAAGCACAGGCGGGCCCGGGTGGACACACTAGCTGCAGACA

>Mafa-DRB*Allele7
TTGAAGCAGGTTAAACATGAGTGTCGTTTCTTCAACGGGACGGAGCGGGTGCGGTACCTGCAGAGACATTTCCATAACCAAG
AGGAGTTCGTGCACTTCGACAGTGACCTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCCAGTACTTG
AACAGTCAGAAGGACCTCCTGGAAGCACAGGCGGGCCCGGGTGGACACACTAGCTGCAGACA

>Mafa-DRB*Allele8
TTGGAGCAGGTTAAACATGAGTGTCACTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCTGAGTACTG
GAACAGCCAGAAGGACATCCTGGAAGACAGGCGGGCCCAGGTGGACACCGTGTGCAAACACAACTACGGGGTTGTGGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele9
TTAGAGTACTGTAAGTTTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCAGTACCTGATCAGATACTTCTATAACCAAGA
GGAGTACGTGCGCTACGACAGCGACGTCGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGG
AACGGCCAGAAGGACCTCCTGGAGCAGAGGCGGGCCGAGGTGGACACCGTGTGCAGACACAACTACGGGGTTGTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele10
TTGGAGGAGTGTAAATATAAATGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGTCAGATTCTTCTATAACCAAGA
GGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGG
AACGGTCAGAAGGACCTCCTGGAAGACAGGCGGGCCGAGGTGGACACCGTGTGCAGACACAACTACGGGGTTGTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele11
TTGGAGGAGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAGAGACTCTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGCGACGTCGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCCGAGAACTG
GAACAGCCGGAAGGACCTCCTGGAGCAGAGGCGGGCCCAGGTGGACACCGTGTGCAGACACAACTACGGGGTTGTTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele12

Appendix Twelve: Long-tailed macaque *Mhc-DRB* sequences

TCTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGGAGGAGTACCTGCGCT
TCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGGAACAGCCGGAAGGA
CATCCTGGAGCAGAGGCGGGCCGCGGTGGACAACTACTGCAGATACAACTACGGGGTTGGTGAGAGCTTCACAGTGCAGC
GGCGAG

>Mafa-DRB*Allele13
TCTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACATCTATAACCAGGAGGAGAACGTGCGCT
TCGACAGCGACGTGGGGGAGTTCCGGGCGGTGTCGGAGCTGGGGCGGCCTGACGCCGAGAACTGGAACAGTCAGAAGGA
CTTCCTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACCGGGTTGGTGAGAGCTTCACAGTGCAGC
GGCGAG

>Mafa-DRB*Allele14
TTGGAGCAGGTTAAACATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATTCTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCGTAACGCCGAGTACTG
GAACAGCCAGAAGGACTTTCTGGAGCACACGCGGACCGAGGTGGACACCTA

>Mafa-DRB*Allele15
TTGGAGCAGGTTAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAGAGACTCTTCTATAACCAGG
AGGAGTTCCTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTCCTG
GAACAGTCGGAAGGACCTCCTGGAAGACAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele16
TTGGAGCAGGTTAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAGAGACTCTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTCCTG
GAACAGTCGGAAGGACCTCCTGGAACACAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele17
TTGGAGCAGGCTAAACGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCCATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGTGGTGACGGAGCTGGGGCGGCGTAGCGCGGAGTACTG
GAACAGCCAGAAGGACTTTCTGGAGCAGGCGCGGACCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGGCGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele18
CTGGGGCATGCTAAGTCAGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGCAGAGATACATCTATAACCAGG
AGGAGTTGGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCCGAGTCCTG
GAACAGTCGGAAGGACTTCCTGGAGCGGGAGCGGGCCCAGGTGGACACCTGCAGACACAACTACGGGGGTGTGGAGAGCT
TCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele19
TTGTGGCAGGCTAAGGGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTACCTGGAGAGATACTTCTATAACCAG
GAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCCGAGTACT
GGAACAGCCTGAAGGACATCCTGGAGCGGAGGCGGGCCGTGGTGGACACCTGCAGATACGACTACGGGGTTGTGGAGAGC
TTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele20
TTGGAGCAGGCTAAACGTAAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGCACAGACACTTCTATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCCGAGTACTG
GAACAGCCAGAAGGACTTCCTGGAGCAGAGGCGGACCTCGGTGGACAATTACTGCAGACACAACTACAGGGTTGTGGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele21
TTGGAGCAGGCTAAGGGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTACCTGCACAGACACTTCCATAACCAG
GAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTAACGGAGCTGGGGCGGCGTAGCGCGGAGAACT
GGAACAGCCAGAAGGACTTTCTGGAGCAGGCGCGGGCCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGTGGAG
AGCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele22
GAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGGAGGAGTTCGTGCGCTTC
GACAGCGACGTGGGGGAGTTCCGGGCGGTGTCGGAGCTGGGGCGGCCTGACGCCGAGTACTGGAACAGTCAGAAGGACT
TCCTGGAGCAGAGGCGGGCCCCGGTTGACAACTACTGCAGATACAACTACCGGGTTGGTGAGAGCTTCACAGTGCAGCGGC
GAG

>Mafa-DRB*Allele23
GAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTACCTGCACAGATACTTCTATAACCAGGAGGAGTACGTGCGCTTC
GACAGCGACGTGGGGGAGTTCCGGGCGGTGTCGGAGCTGGGGCGGCCTGACGCCGAGTACTTCAACAGTCAGAAGGACTT
CCTGGAGCAGAGGCGGGCCCCGGTGGACACCTACTGCAGATACAACTACCGGGTTGGTGAGAGCTTCACAGTGCAGCGGC
GAG

>Mafa-DRB*Allele24
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGTCGGAGCTGGGGCGGCGTAGCGCCGAGTACTG

GAACGACCAGAAGGACATCCTGGAAGACACGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACCGGATTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele25
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGTCGGAGCTGGGGCGGCGTAGCGCCGAGTACTG
GAACGGCCAGAAGGACTTCCTGGAAGACAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACCGGATTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele26
TTGGAGCAGGCTAAGTCTGAGTGTCACTTCAATGGGACGGAGCGGGTGCGGTACCTGGACAGATACATCCATAACCAGGAG
GAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCCGAGTCCTGGA
ACGGCCAGAAGGACATCCTGGAGCGGAAGCGGGCCGAGGTGGACACCGTGTGCAGACACAACTACGGGGTTTTTGAGAGC
TTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele27
TTGAAGCAGGCTAAGGATGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAGAGACACTTCTATAACCAG
GAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCCCAGTACT
TCAACAGCCAGAAGGGCTACCTGGAGCACAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTGTGGAG
AGCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele28
TTGAAGCAGGTTAAGGCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGAGAGACACTTCTATAACCAG
GAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCCGAGTACT
TCAACAGCCAGAAGGGCTACCTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTGTGGAG
AGCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele29
TTCGAGTACTGTACGCATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTACCTAGTCAGATTTTTCTATAACCGAGA
GGAGTATGTGCGCTTCGACAACGACGTGGGGGAGTTCCAGGCGGTGACGGAGCTGGGGCGGCCTGAAGCCGAGTACTGGA
ACGGCCAGAAGGACCTCCTCGAGAAGGTGCGGGCCGAGGTGGACATCTGCAGACACAACTACGGTGTTGTGGAGAGCTTCA
CAGTGCAGCGGCGAG

>Mafa-DRB*Allele30
TTGGAGCAGGCTAAATGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACAGCCAGAAGGACTTTCTGGAGCAGACGCGGGCCGAGGTGGACAACTACTGCAGACACAACTACGGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele31
TTGGAGCACGCTAAATGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCCATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCGTGACGCCGAGTACTG
GAACAGCCAGAAGGACTTCCTGGAGCAGGCGCGGACCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGGCGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele32
TTGGGGCAGGGTAAACGTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAG
GAGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAACTGGGGAGGCCTGTCGCCGAGAACTT
GAACAGTCGGAAGGACTACCTGGAGCAGGCGCGGGCCGCGGTGGACACCTACTGCAGATACAACTACGGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele33
TTGGGGCAGGCTAAACGTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTTCGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGTCGCGGAGAACTG
GAACAGCCGGAAGGACTACCTGGAGCAGGCGCGGGCCGCGGTGGACAATTACTGCAGATACAACTACGGGGTTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele34
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTCGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACGGCCAGGAGGACATCCTGGAAGACGCGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTTTTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele35
TTAGAGTACTGTAAGTTTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCAGTACCTGATCAGATACTTCTATAACCAAGA
GGAGTACGTGCGCTGCGACAGCGACGTCGGGGAGTACCGGCCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTGG
AACGGCCAGAAGGACCTCCTGGAGCAGAGGCGGGCCGAGGTGGACACCGTGTGCAGACACAACTACGGGGTTGTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele36
TTGGAGCAGGCTAAACGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCCATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCGGAGTACTG
GAACAGCCAGAAGGACTTTCTGGAGCAGGCGCGGACCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGGCGAGA
GCTTCACAGTGCAGCGGCGAG

Appendix Twelve: Long-tailed macaque *Mhc-DRB* sequences

>Mafa-DRB*Allele37
TTGGAGCAGGCTAAGTCTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTACCTGCACAGATACATCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGCACCGGGCGGTGACGGAGCTGGGGCGGCCTAGCGCCGAGTACTG
GAACGGCCAGAAGGACATCCTGGAACACCAGCGGGCCGCGGTGGACAACTTCTGCAGATACAACTACCGGGTTGGTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele38
TTGGAGCAGTCTACATCTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGCACAGATACTTCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGCACCGGGCGGTGTCGGAGCTGGGGCGGCGTAGCGCCGAGAACTG
GAACAGTCGGAAGGACATCCTGGAGCACCAGCGGGCCGCGGTGGACACCTTCTGCAGATACAACTACCGGGTTTTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele39
TTGGAGTACTCTACATCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCCGAGTACTG
GAACGGCCAGAAGGACATCCTGGAAGACGCGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACCGGATTGGTGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele40
TTGTGGCAGGATAAGGCTAAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTACCTGGAGAGACACTTCTATAACCAG
GAGGAGTTAGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGTCGGAGCTGGGGCGGCCTGTCGCCGAGAACT
GGAATGGCCAGAAGGACTACCTGGAGCAGAAGCGGGCCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGGTGAG
AGCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele41
TTGTGGCAGGCTAAGGCTAAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTACCTGGAGAGACACTTCTATAACCAG
GAGGAGTTAGTGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGTCGGAGCTGGGGCGGCCTGTCGCCGAGAACT
GGAATGGCCAGAAGGACTACCTGGAGCAGAAGCGGGCCGCGGTGGACAACTACTGCAGACACAACTACGGGGTTGGTGAG
AGCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele42
TTGGAGCAGGCTAAGTCTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGATAGATACATCCATAACCAAG
AGGAGAACCTGCGCTTCGACAGCGACGTAGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACAGCCAGAAGGACTTCCTGGAGCAGAGGCGGGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGCTGTGGAGA
GCTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele43
TTGGAGTACTGTACATCTGAGTGTCACTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGAACAGATACTTCTATAACCAAG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACGGCCAGAAGGACATCCTGGAACACACGCGGGCCGAGGTGGACACCTACTGCAGACACAACTACCGGATTGTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele44
TTAGAGTACTGTACGTTTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGTCAGATACTTCTATAACCAGGA
GGAGTACGTGCGCTTCGACAGCGACGTCGGGGAGTACCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCCGAGTACTGG
AACGGCCAGAAGGACCTCCTGGAGCACGGGCGGGCCGCGGTGGACACCGTGTGCAGACACAACTACCGGATTGGTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele45
TTGGAGCAGGCTAAGTCTGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGGACAGATACATCTATAACCAGG
AGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGCACCGGGCGGTGACGGAGCTGGGGCGGCGTAGCGCCGAGTACTG
GAACGGCCAGAAGGACATCCTGGAAGACCAGCGGGCCGCGGTGGACACCTTCTGCAGATACAACTACCGGGTTTTTGAGAG
CTTCACAGTGCAGCGGCGAG

>Mafa-DRB*Allele46
TTGGAGCAGGTTAAATGTGAGTGTCATTTCTTCAATGGGACGGAGCGGGTGCGGTTCCTGGACAGATACTTCTATAACCAGG
AGGAGTACGCGCGCTTCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGACGCCGAGTACTG
GAACAGCCAGAAGGACCTCCTGGAGCAGAGGCGGGCCGAGGTGGACACCTA

APPENDIX THIRTEEN

Alignment of long-tailed macaque sequences to DRB (exon 2) consensus. Codon numbering is shown above the consensus of Mafa-DRB sequences. Identity to consensus is shown by dashes (-). An indel is shown by (\) and asterisks (*) denote unsequenced nucleotides.

```
                            10                                      20                                      30
HLA-DRB1*010101   TTG TGG CAG CTT AAG TTT GAA TGT CAT TTC TTC AAT GGG ACG GAG CGG GTG CGG TTG CTG GAA AGA TGC ATC TAT AAC
Mafa-DRB*Allele1  --- GA- T-C TC- -CA -C- --G --- --C --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele10 --- GA- G-- TG- --A -A- A-- --- --- --- --- --C --- --- --- --- --- --- --C --- -TC --- -T- T-- --- ---
Mafa-DRB*Allele11 --- GA- G-- GA- --- -A- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --G --- CT- T-- --- ---
Mafa-DRB*Allele12 *** *** *** *** *** -C- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele13 *** *** *** *** *** -C- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- --- --- ---
Mafa-DRB*Allele14 --- GA- --- G-- --A CA- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --C --- -T- T-- --- ---
Mafa-DRB*Allele15 --- GA- --- G-- --- -A- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --G --- CT- T-- --- ---
Mafa-DRB*Allele16 --- GA- --- G-- --- -A- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --G --- CT- T-- --- ---
Mafa-DRB*Allele17 --- GA- --- GC- --A CG- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- C-- ---
Mafa-DRB*Allele18 C-- G-- --T GC- --- -CA --G --- --- --- --- --C --- --- --- --- --- --- --C --- C-G --- -A- --- --- ---
Mafa-DRB*Allele19 --- --- --- GC- --- GG- --G --- --- --- --- --- --- --- --- --- --- --- -AC --- --G --- -A- T-- --- ---
Mafa-DRB*Allele2  --- GA- T-C TC- -CA -C- --G --- --C --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele20 --- GA- --- GC- --A CG- A-G --- --- --- --- --C --- --- --- --- --- --- --C --- C-C --- CA- T-- --- ---
Mafa-DRB*Allele21 --- GA- --- GC- --- GG- --G --- --- --- --- --- --- --- --- --- --- --- -AC --- C-C --- CA- T-- C-- ---
Mafa-DRB*Allele22 *** *** *** *** *** *** --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele23 *** *** *** *** *** *** --G --- --C --- --- --C --- --- --- --- --- --- -AC --- C-C --- -A- T-- --- ---
Mafa-DRB*Allele24 --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele25 --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele26 --- GA- --- GC- --- -C- --G --- --\ \\- --- --- --- --- --- --- --- --- -AC --- --C --- -A- --- C-- ---
Mafa-DRB*Allele27 --- AA- --- GC- --- GA- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --G --- CA- T-- --- ---
Mafa-DRB*Allele28 --- AA- --- G-- --- GC- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --G --- CA- T-- --- ---
Mafa-DRB*Allele29 --C GA- T-C TG- -C- CA- --G --- --- --- --- --C --- --- --- --- --- --- -AC --A -TC --- -TT T-- --- ---
Mafa-DRB*Allele3  --- GA- --- TC- -CA -C- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --C --- CA- T-- --- ---
Mafa-DRB*Allele30 --- GA- --- GC- --A -G- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele31 --- GA- --C GC- --A -G- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- C-- ---
Mafa-DRB*Allele32 --- G-- --- GG- --A CG- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele33 --- G-- --- GC- --A CG- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele34 --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele35 --A GA- T-C TG- --- --- --G --- --- --- --- --- --- --- --- --- --- -A- -AC --- ATC --- -A- T-- --- ---
Mafa-DRB*Allele36 --- GA- --- GC- --A CG- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- C-- ---
Mafa-DRB*Allele37 --- GA- --- GC- --- -C- --G --- --- --- --- --C --- --- --- --- --- --- -AC --- C-C --- -A- --- --- ---
Mafa-DRB*Allele38 --- GA- --- TC- -CA -C- --G --- --- --- --- --C --- --- --- --- --- --- --C --- C-C --- -A- T-- --- ---
Mafa-DRB*Allele39 --- GA- T-C TC- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele4  --- GA- --- G-- --- -A- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --G --- -A- T-- --- ---
Mafa-DRB*Allele40 --- --- --- GA- --- GC- A-G --- --C --- --- --C --- --- --- --- --- --- -AC --- --G --- CA- T-- --- ---
Mafa-DRB*Allele41 --- --- --- GC- --- GC- A-G --- --C --- --- --C --- --- --- --- --- --- -AC --- --G --- CA- T-- --- ---
Mafa-DRB*Allele42 --- GA- --- GC- --- -C- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --T --- -A- --- C-- ---
Mafa-DRB*Allele43 --- GA- T-C TG- -CA -C- --G --- --C --- --- --C --- --- --- --- --- --- --C --- A-C --- -A- T-- --- ---
Mafa-DRB*Allele44 --A GA- T-C TG- -C- --- --G --- --- --- --- --- --- --- --- --- --- --- --C --- -TC --- -A- T-- --- ---
Mafa-DRB*Allele45 --- GA- --- GC- --- -C- --G --- --- --- --- --C --- --- --- --- --- --- --C --- --C --- -A- --- --- ---
Mafa-DRB*Allele46 --- GA- --- G-- --A -G- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele5  --- GA- --- G-- --- -A- --G --- --- --- --- --- --- --- --- --- --- --- --C --- --G --- -A- T-- --- ---
```

```
                                10                                  20                                  30
HLA-DRB1*010101   TTG TGG CAG CTT AAG TTT GAA TGT CAT TTC TTC AAT GGG ACG GAG CGG GTG CGG TTG CTG GAA AGA TGC ATC TAT AAC
Mafa-DRB*Allele6  --- GA- --- G-- --- -A- --G --- -G- --- --- --C A-- --- --- --- --- --- --C --- --G --- CA- T-- --- ---
Mafa-DRB*Allele7  --- AA- --- G-- --A CA- --G --- -G- --- --- --C --- --- --- --- --- --- -AC --- C-G --- CAT T-- C-- ---
Mafa-DRB*Allele8  --- GA- --- G-- --A CA- --G --- --C --- --- --- --- --- --- --- --- --- --C --- --C --- -A- T-- --- ---
Mafa-DRB*Allele9  --A GA- T-C TG- --- --- --G --- --- --- --- --- --- --- --- --- --- --- -A- -AC --- ATC --- -A- T-- --- ---
Mafa-DRB4*0110    --- --- --- GG- --- GC- --G --- --C --- --- --C --- --- --- --- --- --- -AC --- ATC --- -A- --- --- ---

                                            40                                  50
HLA-DRB1*010101   CAA GAG GAG TCC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC CGG GCG GTG ACG GAG CTG GGG CGG CCT GAT GCC GAG
Mafa-DRB*Allele1  --- --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele10 --- --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele11 --G --- --- -T- --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele12 --G --- --- -A- C-- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele13 --G --- --- AA- --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele14 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- -G- A-C --- ---
Mafa-DRB*Allele15 --G --- --- -T- C-- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele16 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele17 --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- -T- --- --- --- --- --- --- -G- AGC --G ---
Mafa-DRB*Allele18 --G --- --- -TG --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele19 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele2  --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- --C --- ---
Mafa-DRB*Allele20 --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele21 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --A --- --- --- --- --- --- -G- AGC --G ---
Mafa-DRB*Allele22 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele23 --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele24 --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele25 --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele26 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele27 --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- -G- AGC --- C--
Mafa-DRB*Allele28 --G --- --- -A- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele29 -G- --- --- -AT --- --- --- --- -A- --- --- --- --- -T- -A- --- --- --- --- --- --- --- --- --A --- ---
Mafa-DRB*Allele3  --G --- --- -A- --- --- -AT --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele30 --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele31 --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- --C --- ---
Mafa-DRB*Allele32 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --A --- --- A-- --- -TC --- ---
Mafa-DRB*Allele33 --G --- --- -T- --- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- -TC --G ---
Mafa-DRB*Allele34 --G --- --- -A- --- --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele35 --- --- --- -A- --- --- -G- --- --- --- --C --- --- --- --- C-- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele36 --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- -G- AGC --G ---
Mafa-DRB*Allele37 --G --- --- -A- --- --- --- --- --- --- --- --- --- C-- --- --- --- --- --- --- --- --- --- AGC --- ---
Mafa-DRB*Allele38 --G --- --- -A- --- --- --- --- --- --- --- --- --- C-- --- --- --- T-- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele39 --G --- --- -A- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele4  --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- --- --C --- ---
```

Appendix Thirteen: Alignment of long-tailed macaque sequences to DRB (exon 2) consensus.

```
                                      40                                                  50
HLA-DRB1*010101    CAA GAG GAG TCC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC CGG GCG GTG ACG GAG CTG GGG CGG CCT GAT GCC GAG
Mafa-DRB*Allele40 --G --- --- -TA --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele41 --G --- --- -TA --- --- --- --- --- --- --- --- --- -T- --- --- --- T-- --- --- --- --- --- -TC --- ---
Mafa-DRB*Allele42 --- --- --- AA- C-- --- --- --- --- --- --A --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele43 --- --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele44 --G --- --- -A- --- --- --- --- --- --- --C --- --- --- --- --- --- --- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele45 --G --- --- -A- --- --- --- --- --- --- --- --- --- C-- --- --- --- --- --- --- --- --- -G- AGC --- ---
Mafa-DRB*Allele46 --G --- --- -A- -C- --- --- --- --- --- --- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB*Allele5  --G --- --- -T- --- --- --- --- --- --- --- --- --- --- --- --- --- T-- --- --- --- --- -G- AGC --- C--
Mafa-DRB*Allele6  --G --- --- -T- --- --- --- --- --T --- C-- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- C--
Mafa-DRB*Allele7  --- --- --- -T- --- -A- --- --- --T --- C-- --- --- -T- --- --- --- --- --- --- --- --- --- --C --- C--
Mafa-DRB*Allele8  --G --- --- -A- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --T ---
Mafa-DRB*Allele9  --- --- --- -A- --- --- -A- --- --- --- --C --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---
Mafa-DRB4*0110    --G --- --- AA- -C- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --C --- ---

                   60                                      70                                                  80
HLA-DRB1*010101    TAC TGG AAC AGC CAG AAG GAC CTC CTG GA\ GCA GAG GCG GGC CGC GGT GGA CAC \CT A\C TGC AGA CAC AAC TAC GGG
Mafa-DRB*Allele1  --- --- --- --T --- --- --- A-- --- --\ --- --A --- --- --G --- --- --A \-- -\- --- --- T-- --- --- C--
Mafa-DRB*Allele10 --- --- --- G-T --- --- --- --- --- --\ AG- C-- --- --- --A --- --- --- \-G T\G --- --- --- --- --- ---
Mafa-DRB*Allele11 A-- --- --- --- -G- --- --- --- --- --\ --- --- --- --- -CA --- --- --- CG- G\\ --- --- --- --- --- ---
Mafa-DRB*Allele12 --- --- --- --- -G- --- --- A-- --- --\ --- --- --- --- --- --- --- --A \-- -\- --- --- T-- --- --- ---
Mafa-DRB*Allele13 A-- --- --- --T --- --- --- T-- --- --\ --- --- --- --- --- --- --- --- \-- -\- --- --- --- --- --- C--
Mafa-DRB*Allele14 --- --- --- --- --- --- --- T-T --- --\ --- C-C --- -A- --A --- --- --- \-- -\* *** *** *** *** *** ***
Mafa-DRB*Allele15 -C- --- --- --T -G- --- --- --- --- --A -\- C-- --- --- --- --- --- --- \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele16 -C- --- --- --T -G- --- --- --- --- --A C\- C-- --- --- --- --- --- --- \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele17 --- --- --- --- --- --- --- T-T --- --\ --- -GC --- -A- --- --- --- --A \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele18 -C- --- --- --T -G- --- --- T-- --- --\ --G -GA --- --- -CA --- --- --- \-\ \\\ --- --- --- --- --- ---
Mafa-DRB*Allele19 --- --- --- --- -T- --- --- A-- --- --\ --G --- --- --- --T --- --- --- \-- \\\ \-- --- T-- G-- --- ---
Mafa-DRB*Allele2  A-- --- --- --T -G- --- --- A-- --- --\ --- --- --- --- --G --- --- --A \-- -\- --- --- --- --- --- C--
Mafa-DRB*Allele20 --- --- --- --- --- --- --- T-- --- --\ --- --- --- -A- -T- --- --- --A \T- -\- --- --- --- --- --- A--
Mafa-DRB*Allele21 A-- --- --- --- --- --- --- T-T --- --\ --- -GC --- --- --- --- --- --A \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele22 --- --- --- --T --- --- --- T-- --- --\ --- --- --- --- -C- --- T-- --A \-- -\- --- --- T-- --- --- C--
Mafa-DRB*Allele23 --- -TC --- --T --- --- --- T-- --- --\ --- --- --- --- -C- --- --- --- \-- -\- --- --- T-- --- --- C--
Mafa-DRB*Allele24 --- --- --- GA- --- --- --- A-- --- --\ AG- C-C --- --- --- --- --- --- \-- -\- --- --- --- --- --- C--
Mafa-DRB*Allele25 --- --- --- G-- --- --- --- T-- --- --\ AG- C-- --- --- --- --- --- --- \-- -\- --- --- --- --- --- C--
Mafa-DRB*Allele26 -C- --- --- G-- --- --- --- A-- --- --\ --G --A --- --- --A --- --- --- \-G T\G --- --- --- --- --- ---
Mafa-DRB*Allele27 --- -TC --- --- --- --- -G- TA- --- --\ --- C-- --- --- --- --- --- --- \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele28 --- -TC --- --- --- --- -G- TA- --- --\ --- --- --- --- --- --- --- --- \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele29 --- --- --- G-- --- --- --- --- --C --\ -A- -GT --- --- --A --- --- --T \-\ \\\ --- --- --- --- --- --T
Mafa-DRB*Allele3  --- --- --- --- --- --- --- A-- --- --\ A-- C-- --- --- --- --- --- --- \-- -\- --- --- T-- --- --- C--
Mafa-DRB*Allele30 --- --- --- --- --- --- --- T-T --- --\ --- --C --- --- --A --- --- --A \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele31 --- --- --- --- --- --- --- T-- --- --\ --- -GC --- -A- --- --- --- --A \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele32 A-- -T- --- --T -G- --- --- TA- --- --\ --- -GC --- --- --- --- --- --- \-- -\- --- --- T-- --- --- ---
```

Appendix Thirteen: Alignment of long-tailed macaque sequences to DRB (exon 2) consensus.

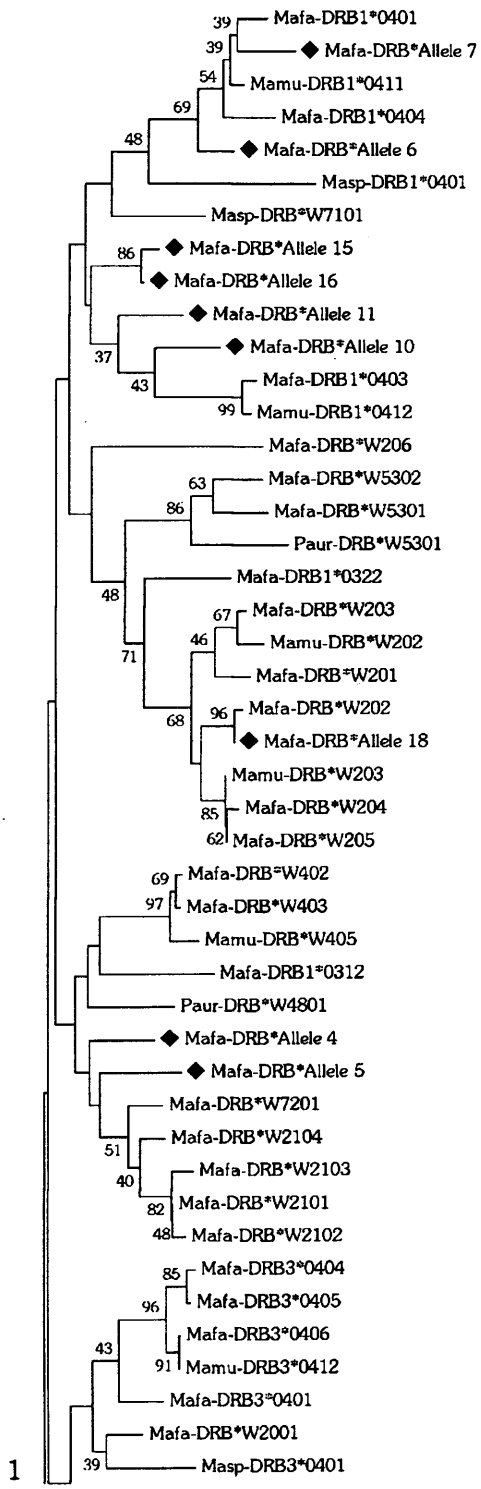
```
                                60                                  70                                  80
HLA-DRB1*010101   TAC TGG AAC AGC CAG AAG GAC CTC CTG GA\ GCA GAG GCG GGC CGC GGT GGA CAC \CT A\C TGC AGA CAC AAC TAC GGG
Mafa-DRB*Allele33 A-- --- --- --- -G- --- --- TA- --- --\ --- -GC --- --- --- --- --A \T- -\- --- --- T-- --- --- ---
Mafa-DRB*Allele34 --- --- --- G-- --- G-- --- A-- --- --\ AG- CGC --- --- --- --- --- \-- -\- --- --- --- --- --- ---
Mafa-DRB*Allele35 --- --- --- G-- --- --- --- --- --- --\ --- --- --- --- --A --- --- --- \-G T\G --- --- --- --- ---
Mafa-DRB*Allele36 --- --- --- --- --- --- --- T-T --- --\ --- -GC --- -A- --- --- --A \-- -\- --- --- --- --- ---
Mafa-DRB*Allele37 --- --- --- G-- --- --- --- A-- --- --\ A-- CCA --- --- --- --- --A \-- T\- --- --- T-- --- C--
Mafa-DRB*Allele38 A-- --- --- --T -G- --- --- A-- --- --\ --- CCA --- --- --- --- --- \-- T\- --- --- --- --- C--
Mafa-DRB*Allele39 --- --- --- G-- --- --- --- A-- --- --\ AG- CGC --- --- --- --- --- \-- -\- --- --- --- --- C--
Mafa-DRB*Allele4  --- --- --- --- --- --- --- T-- --- --\ -G- --A --- -AG -C- --- T-- --A \-- -\- --- --- --- --- ---
Mafa-DRB*Allele40 A-- --- --- --T G-- --- --- TA- --- --\ --- --A --- --- --- --- --A \-- -\- --- --- --- --- ---
Mafa-DRB*Allele41 A-- --- --- --T G-- --- --- TA- --- --\ --- --A --- --- --- --- --A \-- -\- --- --- --- --- ---
Mafa-DRB*Allele42 --- --- --- --- --- --- --- T-- --- --\ --- --- --- --- --- --- --- \-- -\- --- --- --- --- ---
Mafa-DRB*Allele43 --- --- --- G-- --- --- --- A-- --- --\ A-- C-C --- --- --A --- --- \-- -\- --- --- --- --- C--
Mafa-DRB*Allele44 --- --- --- G-- --- --- --- --- --- --\ --- CG- --- --- --- --- --- \-G T\G --- --- --- --- C--
Mafa-DRB*Allele45 --- --- --- G-- --- --- --- A-- --- --\ AG- CCA --- --- --- --- --- \-- T\- --- --- T-- --- C--
Mafa-DRB*Allele46 --- --- --- --- --- --- --- --- --- --\ --- --- --- --- --A --- --- \-- -** *** *** *** *** *** ***
Mafa-DRB*Allele5  -C- --- --- G-- --- --- -G- A-- --- --\ --- --- --- -A- -CA --- --- \-- -\- --- --- --- --- ---
Mafa-DRB*Allele6  --- -T- --- --T --- --- --- --- --- --A --- C-- --- --- --- -CG --- --- A-- -G- --- --- --* *** *** ***
Mafa-DRB*Allele7  --- -T- --- --T --- --- --- --- --- --A --- C-- --- --- --- -CG --- --- A-- -G- --- --- --* *** *** ***
Mafa-DRB*Allele8  --- --- --- --- --- --- --- A-- --- --\ AG- C-- --- --- --- -CA --- --- \-G T\G --- -A- --- --- ---
Mafa-DRB*Allele9  --- --- --- G-- --- --- --- --- --- --\ --- --- --- --- --A --- --- \-G T\G --- --- --- --- ---
Mafa-DRB4*0110    --- --- --- G-- --- --- --- --- --- --\ --G --- --- --- --A --- --- \-- -\- --- --- T-- --- ---

                                90
HLA-DRB1*010101   GTT GGT GAG AGC TTC ACA GTG CAG CGG CGA G
Mafa-DRB*Allele1  --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele10 --- -T- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele11 --- -T- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele12 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele13 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele14 *** *** *** *** *** *** *** *** *** *** *
Mafa-DRB*Allele15 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele16 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele17 --- --C --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele18 -G- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele19 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele2  --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele20 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele21 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele22 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele23 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele24 A-- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele25 A-- --- --- --- --- --- --- --- --- --- -
```

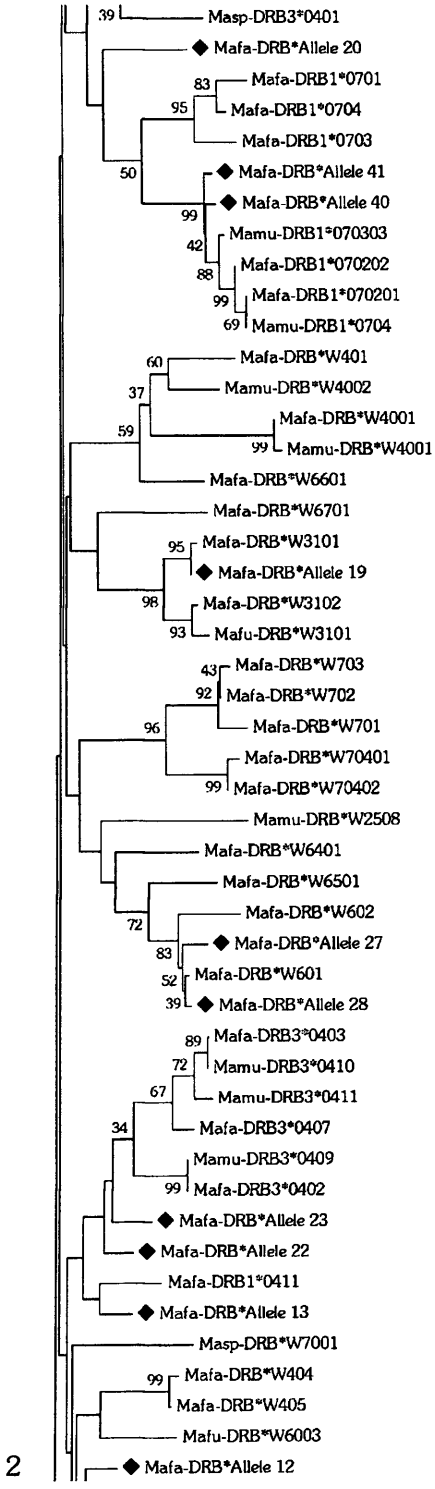
Appendix Thirteen: Alignment of long-tailed macaque sequences to DRB (exon 2) consensus.

```
                                  90
HLA-DRB1*010101   GTT GGT GAG AGC TTC ACA GTG CAG CGG CGA G
Mafa-DRB*Allele26 --- TT- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele27 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele28 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele29 --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele3  --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele30 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele31 --- --C --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele32 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele33 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele34 --- TT- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele35 --- -T- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele36 --- --C --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele37 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele38 --- TT- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele39 A-- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele4  --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele40 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele41 --- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele42 -C- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele43 A-- -T- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele44 A-- --- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele45 --- TT- --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele46 *** *** *** *** *** *** *** *** *** *** *
Mafa-DRB*Allele5  --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele6  *** *** *** *** *** *** *** *** *** *** *
Mafa-DRB*Allele7  *** *** *** *** *** *** *** *** *** *** *
Mafa-DRB*Allele8  --- -TG --- --- --- --- --- --- --- --- -
Mafa-DRB*Allele9  --- -T- --- --- --- --- --- --- --- --- -
Mafa-DRB4*0110    --- --- --- --- --- --- --- --- --- --- -
```

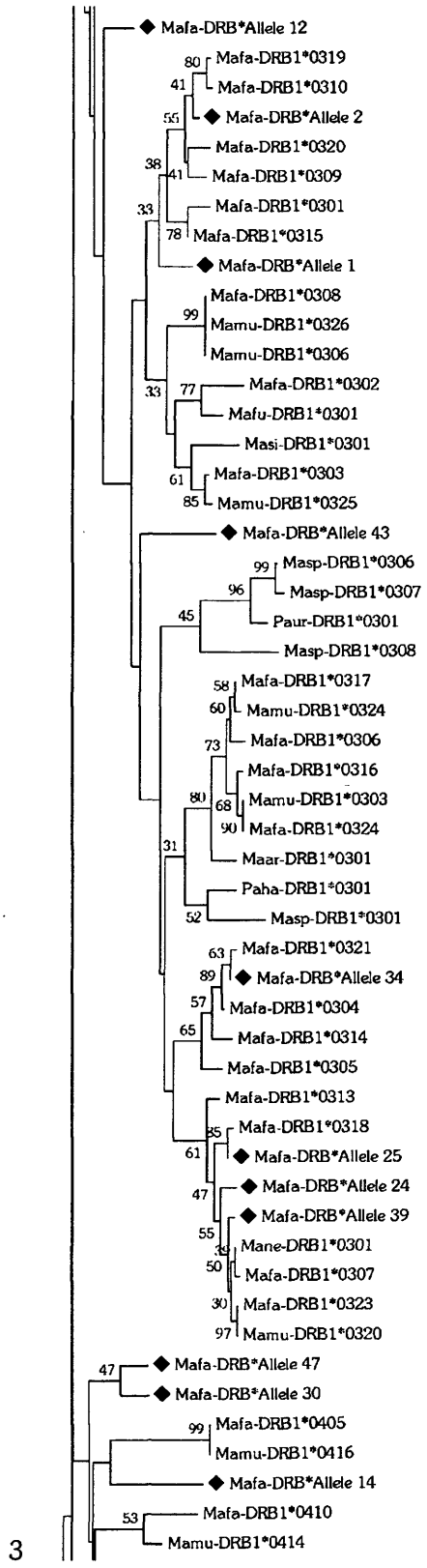
APPENDIX FOURTEEN

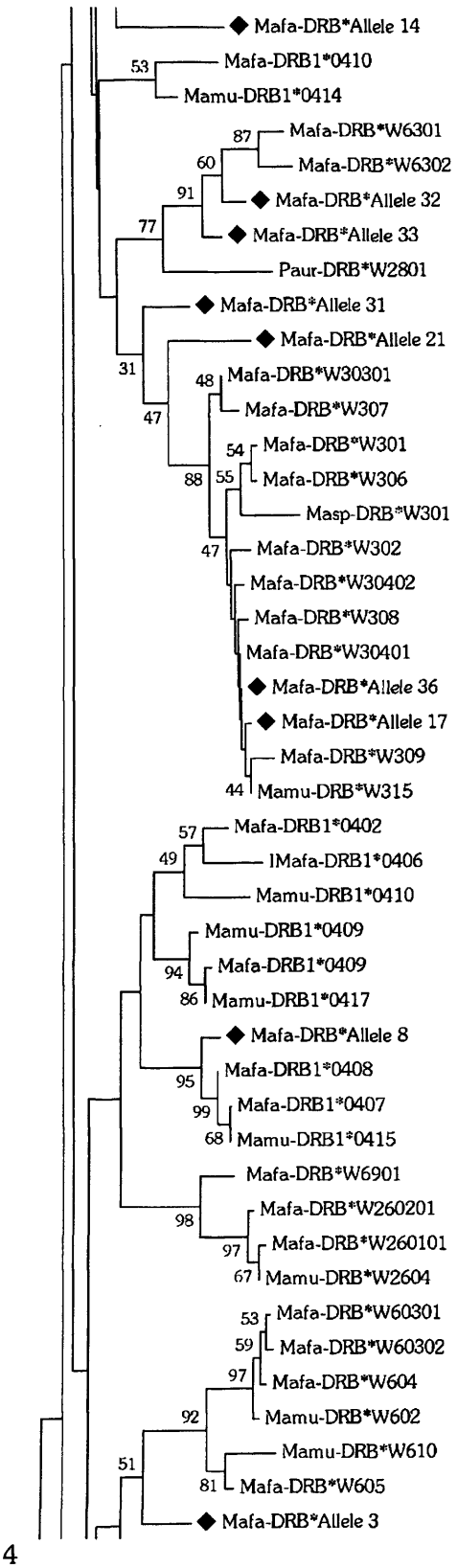
Phylogenetic tree of 47 LKWS *Mafa-* (◆) and other published *Mafa-* and *Macaca* sp. *-DRB* sequences. The tree configuration was derived from 305 nucleotide sequences sequences using the neighbour-joining and Jukes-Cantor methods in MEGA 5.0 Beta. Numbers on the branches refer to confidence probability values.



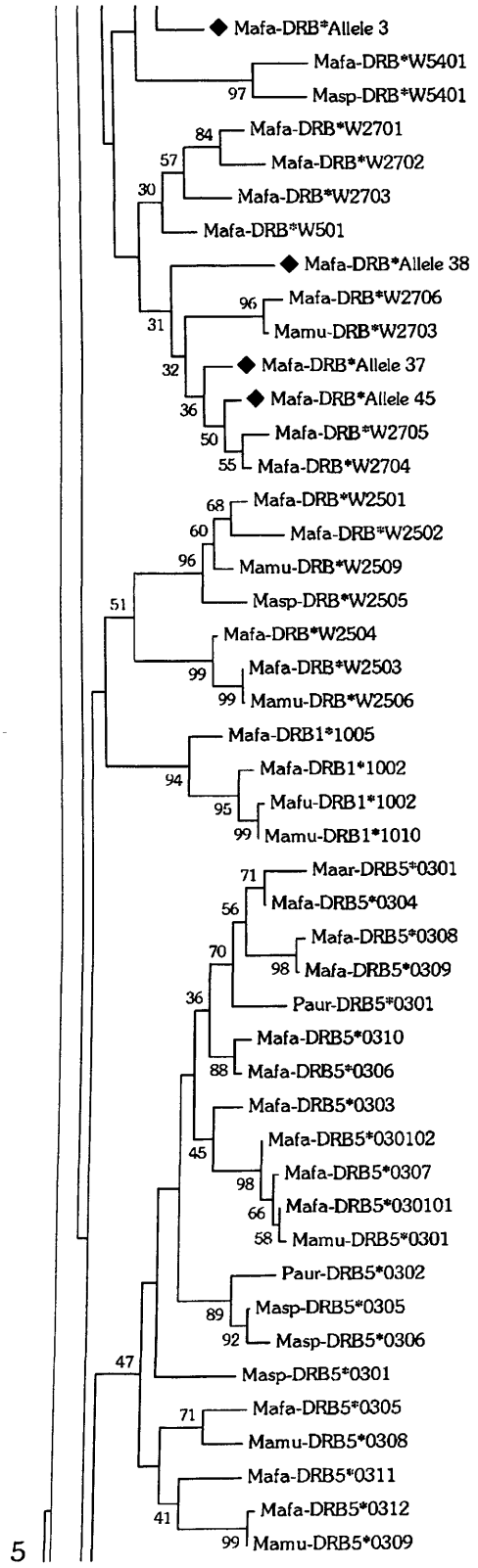


Appendix Fourteen: Phylogenetic tree of LKWS *Mafa-DRB* sequences

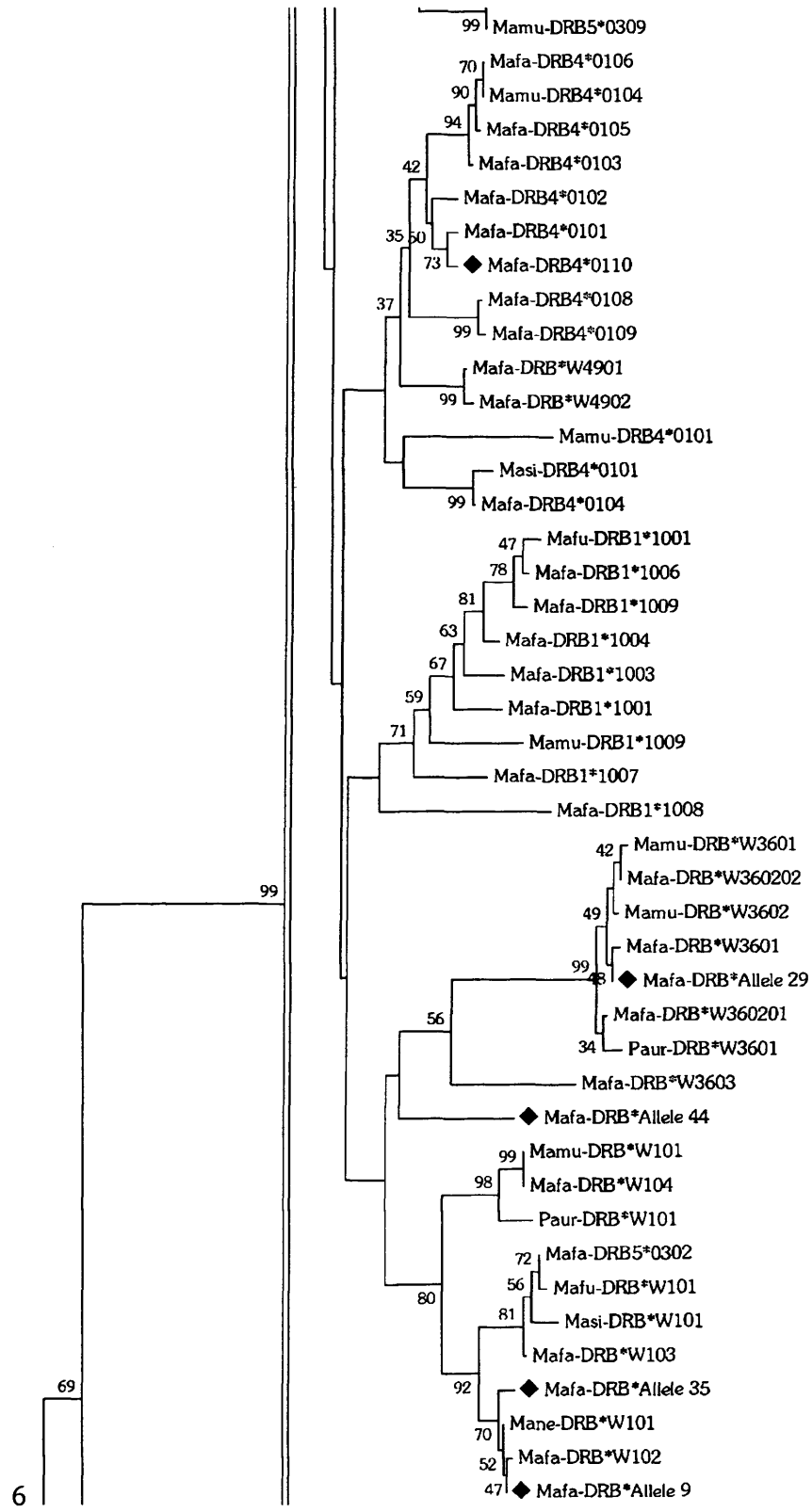




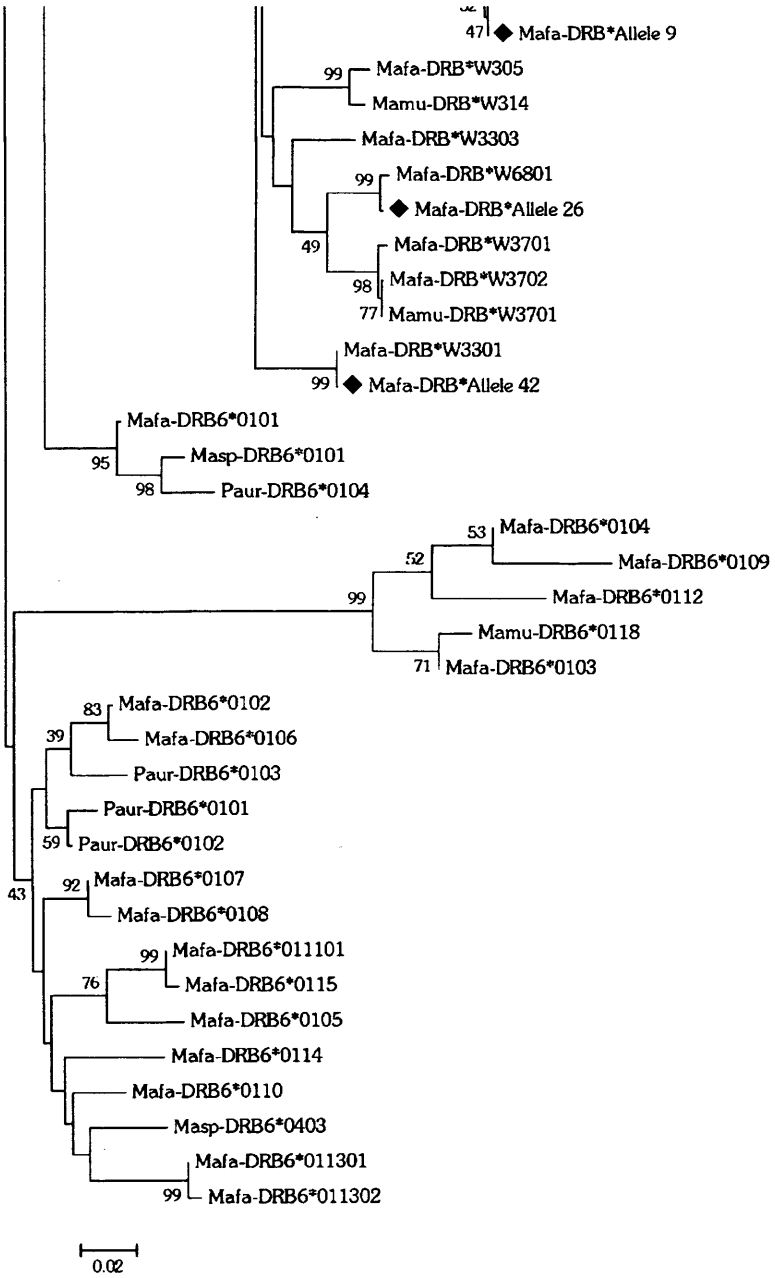
Appendix Fourteen: Phylogenetic tree of LKWS *Mafa-DRB* sequences



Appendix Fourteen: Phylogenetic tree of LKWS *Mafa-DRB* sequences



Appendix Fourteen: Phylogenetic tree of LKWS *Mafa-DRB* sequences

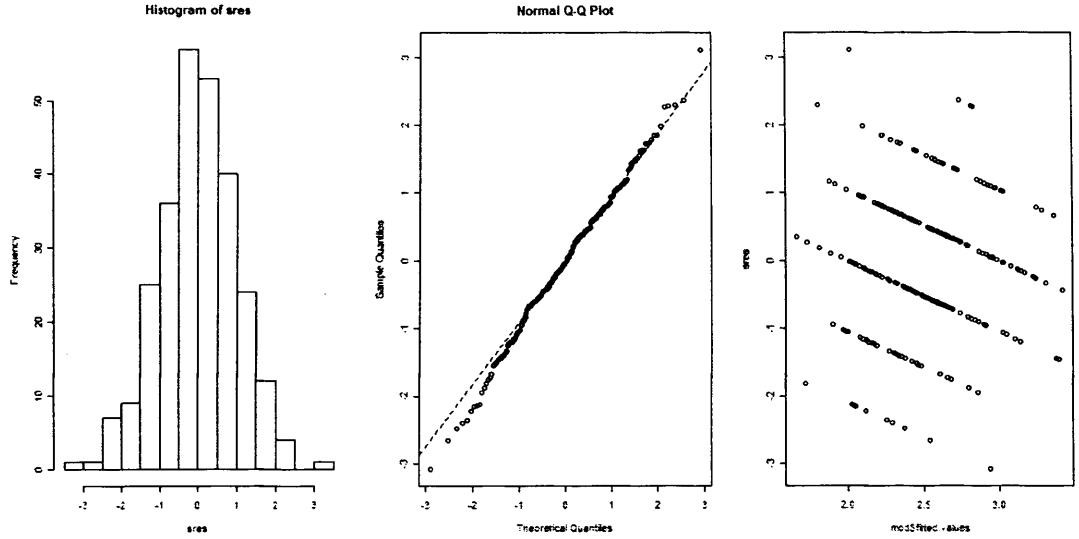


7

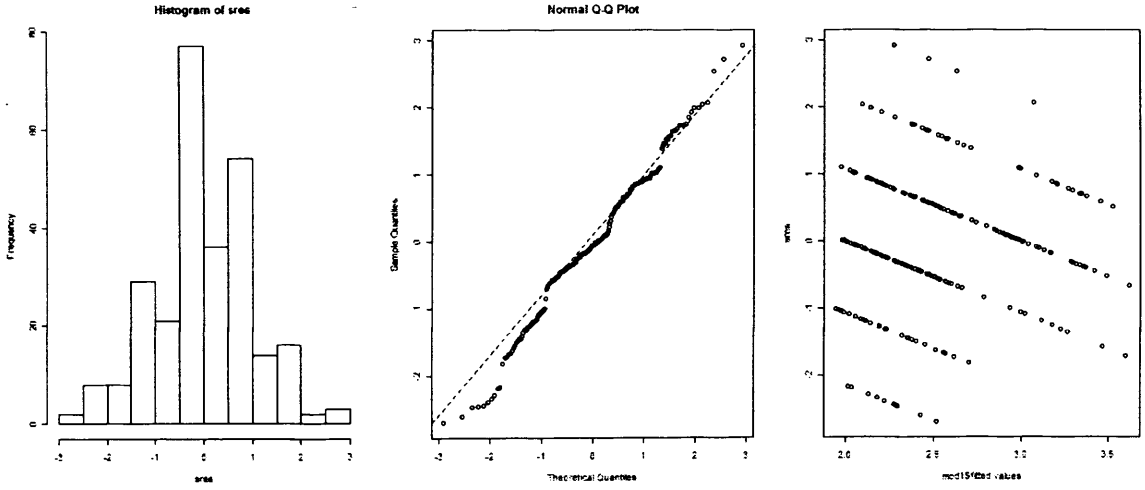
APPENDIX FIFTEEN

Examples of GLM standardized residuals.

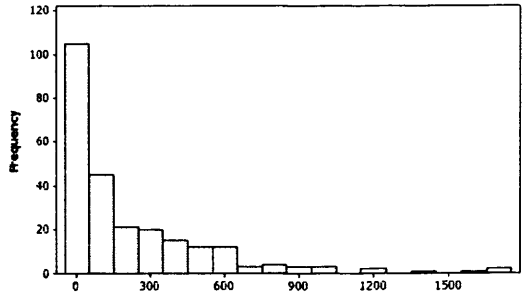
Richness Model A (Locality: North and South of the Kinabatangan River



Richness Model B (Locality: 3 different areas in the LKWS)



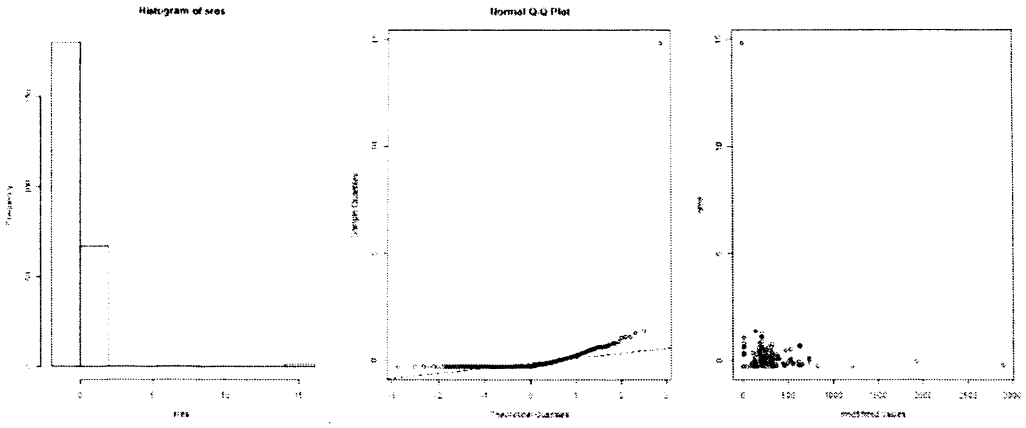
Strongylid (hookworms) frequency data



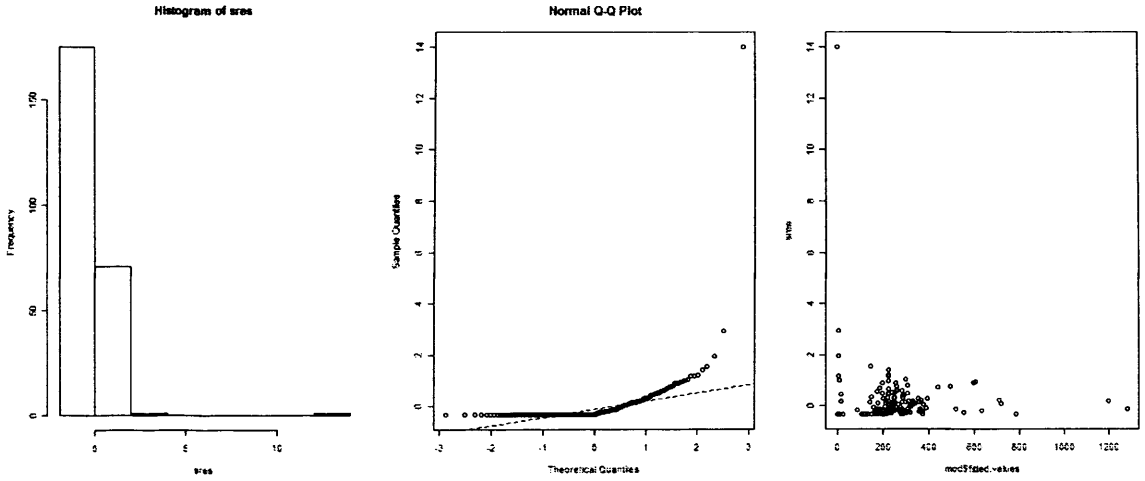
(EPG faeces)

Appendix Fifteen: GLM standardized residuals

Strongylid environmental contamination Model A (with a negative binomial error)

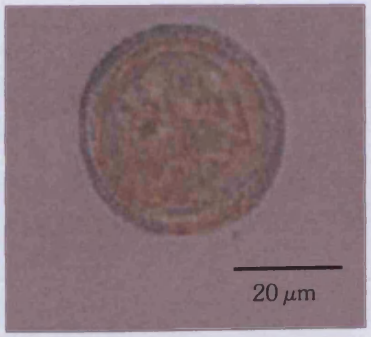
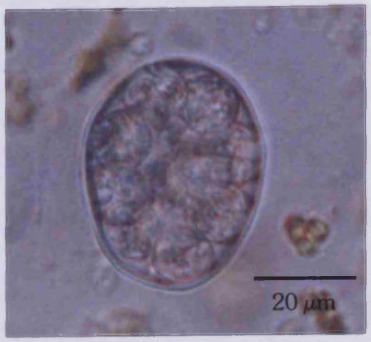


Strongylid environmental contamination Model B (with a quasipoisson error)

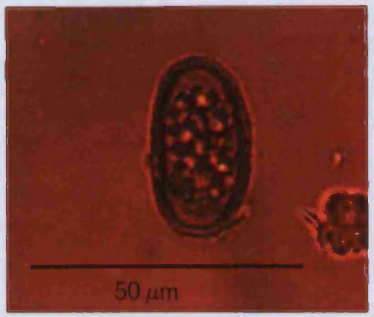
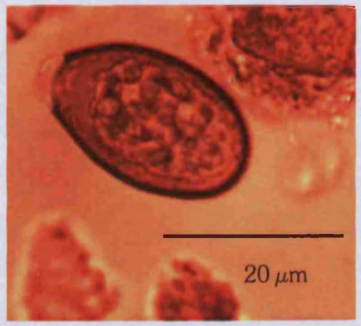
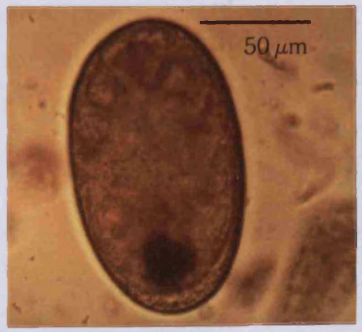


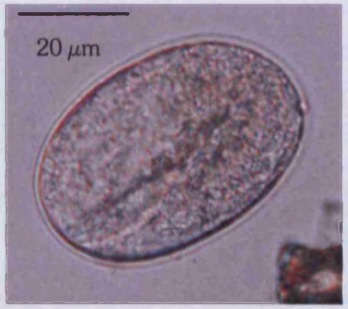
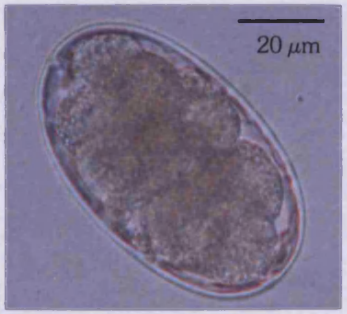
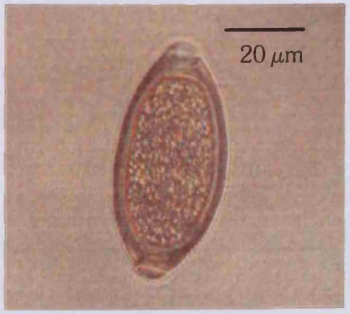
APPENDIX SIXTEEN

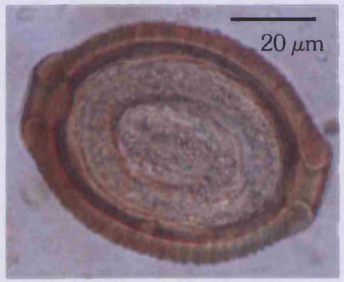
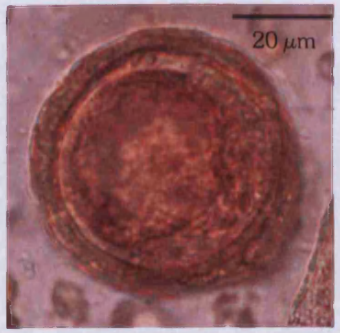
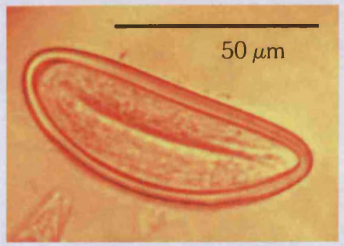
Gastrointestinal helminth parasites found in faeces of proboscis monkeys (*Nasalis larvatus* [PM]) and long-tailed macaques (*Macaca fascicularis* [LTM]) of the Lower Kinabatangan Wildlife Sanctuary

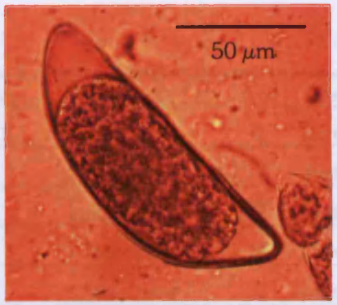
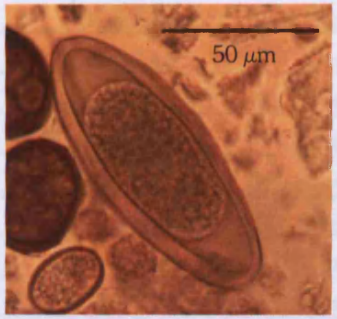
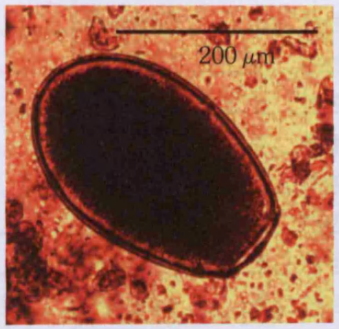
Taxonomic group	Parasite type	Description	Experimental observation	FEC range
Cestoda	*Taenia* sp.	*Taenia* sp. eggs were identified based on size and morphology (spherical with thick striated shell) and measured (44.9 ± 0.4 x 42.9 ± 0.2 μm [n=2]). They were found in both primate species but prevalence was higher in proboscis monkeys than in long-tailed macaques (χ^2= 34.7692, df = 1, P < 0.001) (Table 5.3).		PM:1-133 LTM: 2-11
	Dipylidium-like	Eggs that resemble *Dipylidium* sp. (spherical or oval with 5-15 eggs (or more) enclosed in a sac or capsule) were found in both primate species and measured (40.8 ± 1.0 x 32.2 ± 1.0 μm [n=7]). The size of the packet is smaller than the one reported in the literature (capsule usually ranges in size from 58 μm to 60 μm x 170 μm – DPDx), hence the classification of this morph (*Dipylidium*-like) is based solely on the shape of the eggs and must be taken cautiously. Prevalence of *Dipylidium*-like morphs did not differ among primate species (P>0.1) (Table 5.3)		PM:3-38 LTM: 1-863

Appendix Sixteen: Parasites found in long-tailed macaque and proboscis monkey stool samples.

Trematoda			
	Dicrocoeliid	A dicrocoeliid liver fluke was identified based on egg size and morphology (thick shelled, ellipsoid and operculated). Eggs were found in the faeces of both primate species (38.2 ± 1.4 x 20.9 ± 2.3 μm [n=3]), however these eggs may represent *Dicrocoelium* sp., *Brodenia* sp., and/or *Concinnum* sp. since they cannot be differentiated to the genus level based on egg morphology alone. This was the only trematode shared among primate species and it presented a similar prevalence between them (P>0.5) (Table 5.3).	 PM:1-1 LTM: 30-341
	Clonorchis sp.	*Clonorchis* sp. was identified based on egg morphology and size (small, ovoidal or elongated, with broad rounded posterior end and a convex operculum, "knob" on posterior end was observed occasionally). This trematode was only found in long-tailed macaques (measuring 24.4 ± 1.7 x 16.5 ± 1.2 μm [n=2]) and prevalence was very low (1.4 %) (Table 5.2).	 PM:NP LTM: NP-15
	Fasciola sp.	Egg morphology and size (large size, broadly oval, thin shell, with small, indistinct operculum) were used to identify *Fasciola* sp.. This species was also only found in *Macaca fascicularis* (measuring 140.5 ± 2.6 x 77.4 ± 5.3 μm [n=2]) and prevalence was slightly higher (4.9 %) than *Clonorchis* sp. (Table 5.3).	 PM:NP LTM: NP-2

Nematoda				
	Strongyloides sp.	*Strongyloides* sp. was identified based on egg size and morphology (ellipsoid, with a thin wall resembling strongyle eggs in appearance but slightly smaller in size). Eggs were present in both primate species (measuring 51.8 ± 0.8 x 31.4 ± 1.9 μm [n=5]), and the presence of rhabditoid larvae was also used for inference on the presence of this parasite. Prevalence was higher on long-tailed macaques than on proboscis monkeys ($\chi2$= 8.138, df = 1, P< 0.005) (Table 5.3).		PM:1-15 LTM: 1-42
	Strongylid	Unidentified strongylid eggs (measuring 74.3 ± 1.7 x 52.4 ± 2 μm [n=10]), were found in the faeces of both primate species. These strongyles may represent *Necator* sp., *Ancylostoma* sp., and/or *Oesophagostomum* sp.; however, coprocultures were not performed, limiting the identification of these parasites to the genus level. Prevalence of unidentified strongyles was high on both primate species, but it was higher on macaques than proboscis monkeys ($\chi2$= 56.0885, df = 1, P< 0.001) (Table 5.3).		PM:1-50 LTM: 1-73
	Trichuris sp.	*Trichuris* sp. was identified based on egg size and morphology (barrel-shape, yellow-brown coloration, and bipolar plugs). Eggs measured 56.7 ± 1.4 x 26.2 ± 0.4 μm [n=3] and were present in both primate species. Prevalence of *Trichuris* sp. was significantly higher in proboscis monkeys (> 90%) than in macaques (53.8%, $\chi2$= 67.6835, df = 1, P<0.001) (Table 5.3).		PM:1-143 LTM: 1-53

Nematoda			PM:1-3 LTM: 1-1
Anatrichosoma sp.	Eggs resembling *Trichuris* sp. and *Capillaria* sp. were identified as *Anatrichosoma* sp. based on size and morphology (barrel shaped with bipolar plugs, containing a larva, thick and golden-brown shell, with a roughened or textured surface). Present in both monkey species, eggs measured 75.6 ± 0.8 x 51.6 ± 2.5 μm [n=3]. No primate species had a higher prevalence of this parasite than the other (P>0.5) (Table 5.3).		
Ascaris sp.	*Ascaris* sp. was identified based on egg size and morphology (round or oval, thick-shelled and mammillated albuminous covering). Eggs were found in faeces of both primates and measured 59.3 ± 3.5 x 53.9 ± 2.6 μm [n=9]. Prevalence of *Ascaris* sp. was over 60% in both primate species but no different between each other (P>0.1) (Table 5.3).		PM:1-30 LTM: 1-50
Oxyurid	Identification of *Oxyuridae*. was based on size and morphology of eggs (smooth and thin eggshell, elongated and asymmetrical with one side flattened, other side convex). Eggs measured 72.3 ± 0.8 x 29.8 ± 1.0 μm [n=4] and were equally prevalent in both types of monkeys (P>0.1) (Table 5.3).		PM:1-1 LTM: 1-1

Nematoda	Oxyurid-like	Eggs that resemble *Oxyuridae* were found in long-tailed macaque samples. However, size ranges doubled those reported for the aforementioned species (measuring 124.2 ± 3.8 x 47.8 ± 1.8 μm [n=3]). Prevalence was overall low (2.8%) (Table 5.3).		PM:NP LTM: 1-1
Acanthocephala	"Acanthocephala"	Eggs of what could be a member of the Phylum Acanthocephala were identified based on size and morphology (elongated, oval, with a thick clear shell which appeared textured on different focal planes). Since only two eggs (measuring 98.8 ± 1.2 x 50.8 ± 2.1 μm) were found in all the samples (n=290), identification of this parasite-like object must be taken cautiously. Prevalence did not differ between the two primate species (P>0.1) (Table 5.3).		PM:1-1 LTM: 1-1
Unidentified	Unknown Morph	Unidentified objects, resembling parasite-like eggs were found in faeces of long-tailed macaques. "Eggs" measured 270 ± 1.3 x 170.8 ± 2.5 [n=3] and were dark, oval, with a capped end and the shell was thin and smooth. Appearing in only one sample, prevalence was very low (1.4%) (Table 5.3).		PM:NP LTM: 3-3

APPENDIX SEVENTEEN

Long-tailed macaque samples analysed for microsatellites (STR), MHC or parasites.

Sample ID	Lot	Side	STR	MHC	Parasites
Mafa 1-2-4	1	North			✓
Mafa 1-3-3	1	North	✓		
Mafa 1-3-6	1	North	✓		
Mafa 1-3-7	1	North			✓
Mafa 1-3-8	1	North	✓		
Mafa 1-3-9	1	North	✓		
Mafa 1-4-3	1	North			✓
Mafa 1-5-1	1	North	✓		
Mafa 1-5-2	1	North	✓		
Mafa 1-5-3	1	North	✓		
Mafa 1-6-2	1	North	✓		✓
Mafa 1-7-1	1	North	✓		
Mafa 2-1-1	2	South			✓
Mafa 2-1-10	2	South			✓
Mafa 2-1-4	2	South			✓
Mafa 2-1-5	2	South			✓
Mafa 2-1-6	2	South			✓
Mafa 2-1-7	2	South			✓
Mafa 2-1-8	2	South			✓
Mafa 2-1-9	2	South			✓
Mafa 2-2-1	2	South			✓
Mafa 2-3-8	2	South			✓
Mafa 2-4-1	2	South			✓
Mafa 2-4-4	2	South			✓
Mafa 2-5-1	2	South	✓	✓	

Sample ID	Lot	Side	STR	MHC	Parasites
Mafa 4-5-3	4	North			✓
Mafa 4-5-4	4	North	✓	✓	
Mafa 4-7-1	4	North	✓	✓	
Mafa 4-8-1	4	North	✓		
Mafa 4-8-3	4	North	✓		✓
Mafa 4-8-4	4	North	✓		
Mafa 4-9-1	4	North	✓	✓	
Mafa 4-9-2	4	North			✓
Mafa 4-9-3	4	North	✓		
Mafa 5-10-1	5	North			✓
Mafa 5-10-3	5	North			✓
Mafa 5-10-4	5	North			✓
Mafa 5-1-4	5	North	✓	✓	✓
Mafa 5-2-3	5	North	✓		
Mafa 5-3-10	5	North	✓		✓
Mafa 5-3-6	5	North	✓		✓
Mafa 5-4-2	5	North	✓		✓
Mafa 5-4-3	5	North	✓		
Mafa 5-6-1	5	North	✓		
Mafa 5-7-1	5	North	✓		
Mafa 5-8-2	5	North	✓		
Mafa 5-8-3	5	North			✓
Mafa 5-8-5	5	North	✓		✓
Mafa 5-9-3	5	North	✓		✓
Mafa 5-9-5	5	North	✓		✓

Sample ID	Lot	Side	STR	MHC	Parasites
Mafa 7-5-7	7	North	✓		✓
Mafa 7-6-2	7	North	✓		
Mafa 7-6-3	7	North	✓		✓
Mafa 7-6-4	7	North			✓
Mafa 7-7-2	7	North	✓		✓
Mafa 7-7-5	7	North	✓		✓
Mafa 7-8-1	7	North	✓		✓
Mafa 7-8-3	7	North			✓
Mafa 7-8-6	7	North			✓
Mafa 8-1-10	8	North	✓		✓
Mafa 8-1-2	8	North	✓		✓
Mafa 8-1-4	8	North			✓
Mafa 8-1-6	8	North			✓
Mafa 8-1-8	8	North	✓		✓
Mafa 8-2-1	8	North	✓	✓	✓
Mafa 8-3-10	8	North	✓		
Mafa 8-3-3	8	North	✓	✓	✓
Mafa 8-3-5	8	North	✓		✓
Mafa 8-3-8	8	North	✓		✓
Mafa 8-4-1	8	North	✓	✓	✓
Mafa 8-5-1	8	North	✓		
Mafa 8-5-3	8	North	✓	✓	✓
Mafa 8-6-1	8	North			✓
Mafa 9-10-1	9	South	✓		✓
Mafa 9-10-2	9	South		✓	

Appendix seventeen: Long-tailed macaque samples analysed in this study.

Sample ID	Lot	Side	STR	MHC	Parasites
Mafa 2-6-1	2	South	✔		✔
Mafa 2-6-2	2	South			✔
Mafa 2-6-3	2	South			✔
Mafa 2-7-1	2	South	✔	✔	✔
Mafa 2-7-6	2	South	✔		✔
Mafa 2-7-7	2	South	✔		✔
Mafa 2-8-1	2	South	✔	✔	
Mafa 2-8-2	2	South	✔		
Mafa 2-9-1	2	South			✔
Mafa 3-10-6	3	South	✔		✔
Mafa 3-11-6	3	South			✔
Mafa 3-13-3	3	South			✔
Mafa 3-14-1	3	South	✔	✔	✔
Mafa 3-15-3	3	South	✔	✔	✔
Mafa 3-15-4	3	South			✔
Mafa 3-16-3	3	South			✔
Mafa 3-16-5	3	South	✔		✔
Mafa 3-17-11	3	South			✔
Mafa 3-17-2	3	South	✔	✔	✔
Mafa 3-17-3	3	South			✔
Mafa 3-17-5	3	South			✔
Mafa 3-2-1	3	South	✔		✔
Mafa 3-3-1	3	South			✔
Mafa 3-4-1	3	South	✔		✔
Mafa 3-5-4	3	South	✔		✔
Mafa 3-7-2	3	South	✔	✔	
Mafa 6-10-10	6	South	✔	✔	✔
Mafa 6-10-11	6	South			✔
Mafa 6-10-13	6	South			✔
Mafa 6-10-14	6	South			✔
Mafa 6-10-5	6	South			✔
Mafa 6-10-7	6	South			✔
Mafa 6-11-1	6	South			✔
Mafa 6-11-2	6	South			✔
Mafa 6-11-3	6	South			✔
Mafa 6-1-2	6	South	✔		✔
Mafa 6-12-1	6	South	✔		✔
Mafa 6-12-3	6	South			✔
Mafa 6-12-7	6	South			✔
Mafa 6-13-1	6	South	✔		✔
Mafa 6-14-1	6	South	✔	✔	✔
Mafa 6-14-2	6	South			✔
Mafa 6-14-3	6	South			✔
Mafa 6-15-1	6	South	✔		✔
Mafa 6-15-3	6	South			✔
Mafa 6-15-7	6	South			✔
Mafa 6-2-10	6	South	✔		✔
Mafa 6-2-6	6	South			✔
Mafa 6-4-3	6	South	✔	✔	✔
Mafa 6-4-7	6	South	✔	✔	✔
Mafa 6-6-3	6	South			✔
Mafa 6-8-2	6	South			✔
Mafa 9-10-6	9	South			✔
Mafa 9-11-1	9	South			✔
Mafa 9-1-6	9	South	✔	✔	✔
Mafa 9-1-7	9	South	✔	✔	✔
Mafa 9-2-2	9	South			✔
Mafa 9-2-3	9	South			✔
Mafa 9-4-3	9	South	✔		✔
Mafa 9-5-2	9	South	✔	✔	✔
Mafa 9-5-3	9	South	✔		✔
Mafa 9-6-1	9	South			✔
Mafa 9-6-3	9	South	✔		✔
Mafa 9-7-3	9	South	✔	✔	✔
Mafa 9-7-7	9	South	✔	✔	✔
Mafa 9-8-1	9	South	✔	✔	✔
Mafa 9-8-4	9	South	✔		✔
Mafa 9-9-1	9	South	✔	✔	✔
Mafa OutS-1-3		South	✔		✔
Mafa OutS-1-5		South			✔
Mafa OutS-1-6		South			✔
Mafa 10a-1-3	10	North			✔
Mafa 10a-2-1	10	North	✔		✔
Mafa 10a-4-3	10	North			✔
Mafa 10a-4-4	10	North	✔		✔
Mafa 10a-6-1	10	North	✔	✔	✔
Mafa 10a-7-1	10	North	✔		✔
Mafa 10a-7-2	10	North			✔

Appendix seventeen: Long-tailed macaque samples analysed in this study.

Sample ID	Lot	Side	STR	MHC	Parasites
Mafa 3-8-1	3	South	✓	✓	
Mafa 3-9-8	3	South	✓		✓
Mafa 4-10-3	4	North	✓		✓
Mafa 4-10-5	4	North	✓		
Mafa 4-11-1	4	North	✓		
Mafa 4-11-3	4	North	✓	✓	✓
Mafa 4-4-1	4	North			✓
Mafa 4-4-2	4	North	✓		
Mafa 4-5-1	4	North	✓		
Mafa 6-9-1	6	South	✓		✓
Mafa 7-2-1	7	North	✓	✓	✓
Mafa 7-3-1	7	North	✓		✓
Mafa 7-4-10	7	North	✓		✓
Mafa 7-4-2	7	North	✓		✓
Mafa 7-4-5	7	North	✓	✓	✓
Mafa 7-5-1	7	North	✓	✓	✓
Mafa 7-5-2	7	North	✓		
Mafa 7-5-5	7	North	✓		
Mafa 10b-1-12	10	North			✓
Mafa 10b-1-3	10	North			✓
Mafa 10b-1-4	10	North			✓
Mafa 10b-1-5	10	North	✓		✓
Mafa 10b-2-1	10	North	✓	✓	✓
Mafa 10c-1-2	10	North	✓		✓
Mafa 10c-1-3	10	North	✓	✓	✓
Mafa 10c-1-4	10	North	✓		

APPENDIX EIGHTEEN

Proboscis monkey samples analysed for microsatellites (STR), MHC or parasites.

Sample ID	Lot	Side	STR	MHC	Parasites
PM 1-1-1	1	South			✓
PM 1-1-2	1	South			✓
PM 1-1-3	1	South			✓
PM 1-1-4	1	South			✓
PM 1-1-5	1	South			✓
PM 1-1-6	1	South			✓
PM 1-1-7	1	South			✓
PM 1-1-8	1	South			✓
PM 1-2-13	1	South	✓		✓
PM 1-2-14	1	South			✓
PM 1-2-4	1	South			✓
PM 1-2-5	1	South			✓
PM 1-2-7	1	South			✓
PM 1-2-8	1	South			✓
PM 1-3-1	1	South			✓
PM 1-4-15	1	South	✓		
PM 1-4-16	1	South	✓		
PM 1-4-17	1	South	✓		
PM 1-4-8	1	South	✓		✓
PM 1-5-10	1	South			✓
PM 1-5-19	1	South	✓		
PM 2-1-1	2	North			✓
PM 2-1-2	2	North			✓
PM 2-1-3	2	North			✓
PM 2-1-4	2	North			✓
PM 2-1-5	2	North			✓
PM 3-3-2	3	South			✓
PM 3-3-3	3	South			✓
PM 3-3-4	3	South			✓
PM 3-4-1	3	South			✓
PM 3-4-2	3	South			✓
PM 3-5-1	3	South			✓
PM 3-5-18	3	South	✓	✓	
PM 3-5-22	3	South	✓		
PM 3-5-37	3	South	✓		
PM 3-5-44	3	South	✓		✓
PM 3-6-18	3	South			✓
PM 3-6-5	3	South	✓		✓
PM 3-7-14	3	South			✓
PM 3-7-2	3	South		✓	✓
PM 3-7-34	3	South			✓
PM 3-8-13	3	South			✓
PM 3-8-24	3	South			✓
PM 3-8-28	3	South	✓		
PM 3-9-2	3	South			✓
PM 4-10-1	4	North	✓		
PM 4-11-1	4	North	✓		✓
PM 4-11-9	4	North	✓	✓	✓
PM 4-1-2	4	North			✓
PM 4-1-3	4	North	✓		✓
PM 4-14-4	4	North	✓	✓	✓
PM 4-15-3	4	North	✓		✓
PM 6-4-3	6	South	✓	✓	
PM 6-5-1	6	South	✓		✓
PM 6-5-2	6	South			✓
PM 6-6-1	6	South			✓
PM 6-6-10	6	South	✓		
PM 6-6-2	6	South	✓	✓	✓
PM 6-6-3	6	South			✓
PM 6-6-6	6	South	✓	✓	✓
PM 6-7-1	6	South			✓
PM 6-7-2	6	South	✓		✓
PM 6-7-5	6	South	✓		
PM 7-11-12	7	North			✓
PM 7-11-16	7	North			✓
PM 7-11-26	7	North	✓	✓	✓
PM 7-11-4	7	North	✓		✓
PM 7-11-6	7	North			✓
PM 7-11-8	7	North	✓		✓
PM 7-12-1	7	North	✓		
PM 7-1-3	7	North			✓
PM 7-13-1	7	North			✓
PM 7-14-1	7	North	✓		
PM 7-14-2	7	North			✓
PM 7-14-4	7	North			✓
PM 7-2-1	7	North			✓
PM 7-2-2	7	North	✓		✓
PM 7-3-11	7	North			✓

Sample ID	Lot	Side	STR	MHC	Parasites
PM 2-2-1	2	North			✔
PM 2-2-2	2	North			✔
PM 2-2-3	2	North			✔
PM 2-2-4	2	North			✔
PM 2-3-1	2	North			✔
PM 2-3-2	2	North			✔
PM 2-3-3	2	North			✔
PM 2-4-1	2	North			✔
PM 2-4-2	2	North			✔
PM 2-4-3	2	North			✔
PM 2-4-4	2	North			✔
PM 2-4-5	2	North			✔
PM 2-5-1	2	North	✔		✔
PM 2-5-4	2	North			✔
PM 2-6-40	2	North			✔
PM 2-6-9	2	North			✔
PM 2-7-6	2	North			✔
PM 2-9-5	2	North			✔
PM 3-1-1	3	South			✔
PM 3-11-1	3	South	✔		
PM 3-11-2	3	South			✔
PM 3-11-4	3	South	✔	✔	
PM 3-1-2	3	South			✔
PM 3-1-3	3	South			✔
PM 3-1-4	3	South			✔
PM 3-1-5	3	South			✔
PM 3-1-6	3	South			✔
PM 3-1-7	3	South			✔
PM 4-15-4	4	North			✔
PM 4-16-4	4	North	✔	✔	✔
PM 4-17-4	4	North	✔	✔	✔
PM 4-2-16	4	North	✔		✔
PM 4-2-7	4	North	✔		✔
PM 4-3-6	4	North			✔
PM 4-4-7	4	North	✔		✔
PM 4-8-1	4	North	✔		✔
PM 4-9-2	4	North			✔
PM 4-9-5	4	North	✔		✔
PM 5-1-10	5	North	✔		✔
PM 5-1-13	5	North			✔
PM 5-1-2	5	North	✔		✔
PM 5-1-6	5	North	✔	✔	✔
PM 5-2-3	5	North			✔
PM 5-2-4	5	North			✔
PM 5-3-11	5	North	✔		✔
PM 5-3-7	5	North	✔		✔
PM 5-4-1	5	North			✔
PM 5-4-12	5	North			✔
PM 5-4-2	5	North			✔
PM 5-4-7	5	North	✔		✔
PM 5-5-1	5	North	✔		✔
PM 5-5-5	5	North	✔		
PM 5-6-10	5	North	✔		✔
PM 5-6-13	5	North	✔		✔
PM 5-6-4	5	North	✔		✔
PM 5-6-9	5	North	✔		✔
PM 7-3-2	7	North			✔
PM 7-3-7	7	North			✔
PM 7-4-11	7	North	✔		✔
PM 7-4-13	7	North			✔
PM 7-4-16	7	North			✔
PM 7-4-3	7	North			✔
PM 7-4-4	7	North			✔
PM 7-4-5	7	North	✔		✔
PM 7-7-1	7	North			✔
PM 7-7-3	7	North	✔		✔
PM 7-7-7	7	North			✔
PM 7-8-1	7	North			✔
PM 7-9-4	7	North	✔		✔
PM 7-9-5	7	North			✔
PM 10b-1-1	10	North	✔		
PM 10b-1-4	10	North	✔		
PM 10b-1-5	10	North			✔
PM 10c-1-5	10	North			✔
PM 10c-2-1	10	North	✔		
PM 10c-2-12	10	North	✔	✔	✔
PM 10c-2-13	10	North	✔		
PM 10c-2-2	10	North			✔
PM 10c-2-7	10	North	✔		✔
PM 10c-2-8	10	North	✔		✔
PM 10c-3-2	10	North	✔		✔
PM 10c-4-4	10	North			✔
PM 10c-4-5	10	North			✔
PM 10c-4-6	10	North	✔		

Appendix eighteen: Proboscis monkey samples analysed in this study.

Sample ID	Lot	Side	STR	MHC	Parasites	Sample ID	Lot	Side	STR	MHC	Parasites	Sample ID	Lot	Side	STR	MHC	Parasites
PM 3-2-1	3	South			✔	PM 6-1-1	6	South			✔	SK019		Klias		✔	
PM 3-2-7	3	South			✔	PM 6-1-6	6	South			✔	T1S2		Klias		✔	
PM 3-3-1	3	South			✔	PM 6-3-1	6	South	✔		✔						