Comparative phylogeography of three primate species in the Lower Kinabatangan Wildlife Sanctuary, Sabah, Malaysia

MOHD. FAIRUS B. JALIL

Thesis submitted for the degree of Doctor of Philosophy

Biodiversity and Ecological Processes Research Group
School of Biosciences, Cardiff University

December 2006

UMI Number: U215944

UMI

Dissertation Publishing

UMI U215944

ProQuest

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

Abstract

This study investigates the population genetic structure of three primate species living in forest fragments of the Lower Kinabatangan Wildlife Sanctuary (LKWS), Sabah, Malaysia. The sanctuary is surrounded by oil palm plantations and human settlements but still retains high diversity of both flora and fauna. LKWS is famous for its orang-utan and proboscis monkey populations but also supports Sabah's eight other primate species. The current study investigated the effects of forest fragmentation and geographical barriers, especially the Kinabatangan River, on three species of primates with different social systems and dispersal abilities. The orang-utan is a large bodied, solitary ape that is incapable of swimming whereas the proboscis monkey and the long-tailed macaque, are smaller bodied, live in large groups and are good swimmers. Using non-invasive samples (faeces), we sequenced approximately 100 individuals from each of these three primates using the left domain (and right domain for long-tailed macaques) of the mitochondrial control region. High levels of genetic diversity were detected in the proboscis monkey and long-tailed macaque, but lower levels were detected in the orang-utan. Statistical analyses (haplotype and nucleotide diversities, mismatch distributions and neutrality tests) indicate that the orang-utan and the proboscis monkey have experienced population bottlenecks, which for the orang-utan supports our earlier studies using microsatellites. Long-tailed macaques, on the other hand, show evidence of population stability. As predicted from the known mobility of these primates, the Kinabatangan River did not appear to impede the movement of proboscis monkeys and long-tailed macaques, but did act as a geographical barrier for orang-utans. Furthermore, re-analysis of the current data with previously published sequences of orang-utans collected throughout Borneo revealed the likely importance of rivers in differentiating between populations that corresponded closely to currently described subspecies (with the exception of the subspecies *morio*).

There are four general conclusions from the current study. Firstly, non-invasive faecal samples are viable for large scale studies on these wild primate populations. Secondly, mitochondrial DNA is an informative marker for population studies due to its high levels of polymorphism over small spatial scales (with the left domain of the control region providing better resolution than the right domain). Thirdly, the social structure of primate species profoundly influences patterns of mitochondrial genetic diversity. Finally, dispersal patterns greatly influence the mitochondrial genetic structure of these populations. The implications of these findings for the future of Borneo primates and conservation of Lower Kinabatangan Wildlife Sanctuary and Sabah are discussed.

Acknowledgements

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

This study examines the comparative population structure of the long-tailed macaque, the orang-utan and the proboscis monkey in the Kinabatangan Wildlife Sanctuary, using mitochondrial DNA sequences. To introduce the study area, the geological history and geography of South East Asia and the impact of habitat degradation on biodiversity are described. The biology of the three primates is reviewed, including previous genetic studies. In the context of phylogeography and conservation genetics, the utility of different molecular markers and the suitability of non-invasive sources of DNA are considered. Finally, specific details of the geography and biodiversity of the study site are described.

1.1   HISTORICAL BIOGEOGRAPHY OF SOUTH EAST ASIA

From the early Eocene (~50 million years ago, MYA) to the late Oligocene (~25 MYA), South East Asia formed a continental block with Indochina (Hall, 1998; Wilson & Moss, 1999). Between 20 and 10 MYA, Borneo is reported to have experienced a counter clockwise rotation of 45° based on Hall's (1996) reconstruction. However, this rotation hypothesis for Borneo is highly contentious (see Wilson & Moss, 1999). By the early Pliocene (~5 MYA), the region had essentially attained its modern formation, although the exact margins and low-lying areas at this time are unknown (Fig. 1.1, Hall, 1998; Voris, 2000).

During the Pleistocene, climate changes caused the sea level to drop, which exposed land bridges among islands. Until around 9,500 years ago, all three of the major Sunda Islands remained connected to one another and the mainland Indochina, via the Malay Peninsula, forming the Sunda Shelf (Fig.1.2; Voris, 2000; Inger & Voris, 2001). During Pleistocene glacial periods, the central range of New Guinea experienced a lowering of air temperatures

by up to 6°C, which in turn lowered the snowline by approximately 1000 m, influencing the treeline and other altitudinal vegetation (Verstappen, 1997). In Borneo, the Kinabalu glaciation was limited to just a few square kilometres above 3,800 m. Fluctuations in sea levels would have governed coastal evolution during the Pleistocene, when the extensive shelf areas emerged in the glacial period and a drainage network developed (Verstappen, 1997).

Figure 1.1. Map showing the modern form of South East Asia.
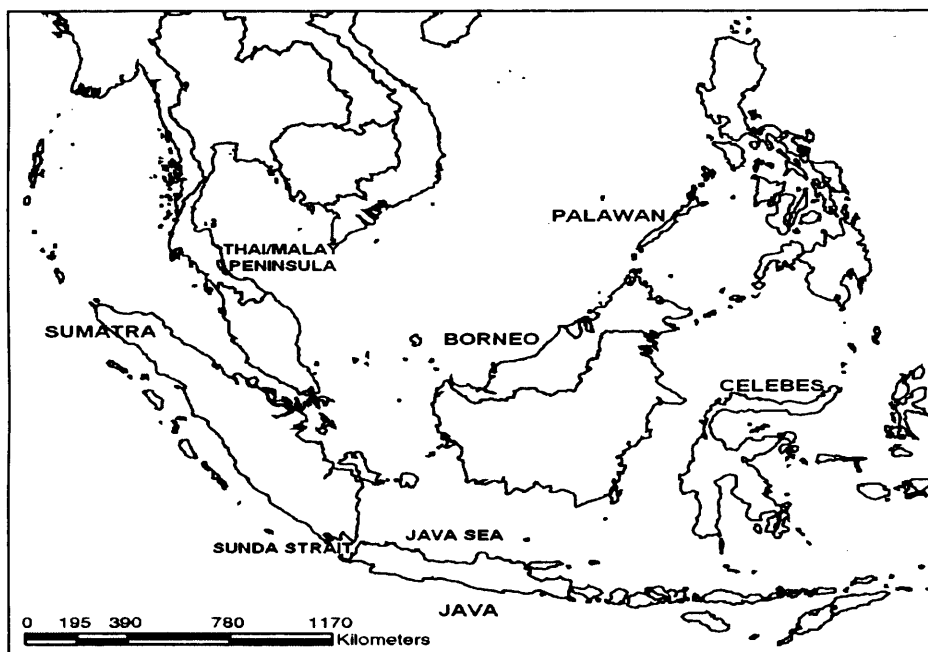
The ability of rainforest animals to move across the exposed shelf in the Pleistocene would have been influenced by topographical and ecological features of the landscape (Tougard, 2001). There is fossil evidence from *Pongo pygmaeus* and *Ailuropoda melanoleuca* that the colonization of large mammals across Sundaland followed migration routes from Indochina to the Malay Peninsula and Sumatra, and to Java and Borneo (Tougard, 2001). The fossils found in mainland Asia in the Late Middle Pleistocene already featured modern characteristics, whereas the fossils found in Java still contained archaic forms indicating that Java was isolated compared to mainland Asia (Tougard, 2001).

Figure 1.2 Voris' (2000) maps of South East Asia and Australasia illustrating depth contours of 10, 20, 30, 40, 50, 75, 100 and 120 m below present level. On the lower left corner of each map, a horizontal bar graph provides estimates of the percentage of time that the sea level was at or below the level illustrated on the map during the past 17,000, 150,000 and 250,000 years.
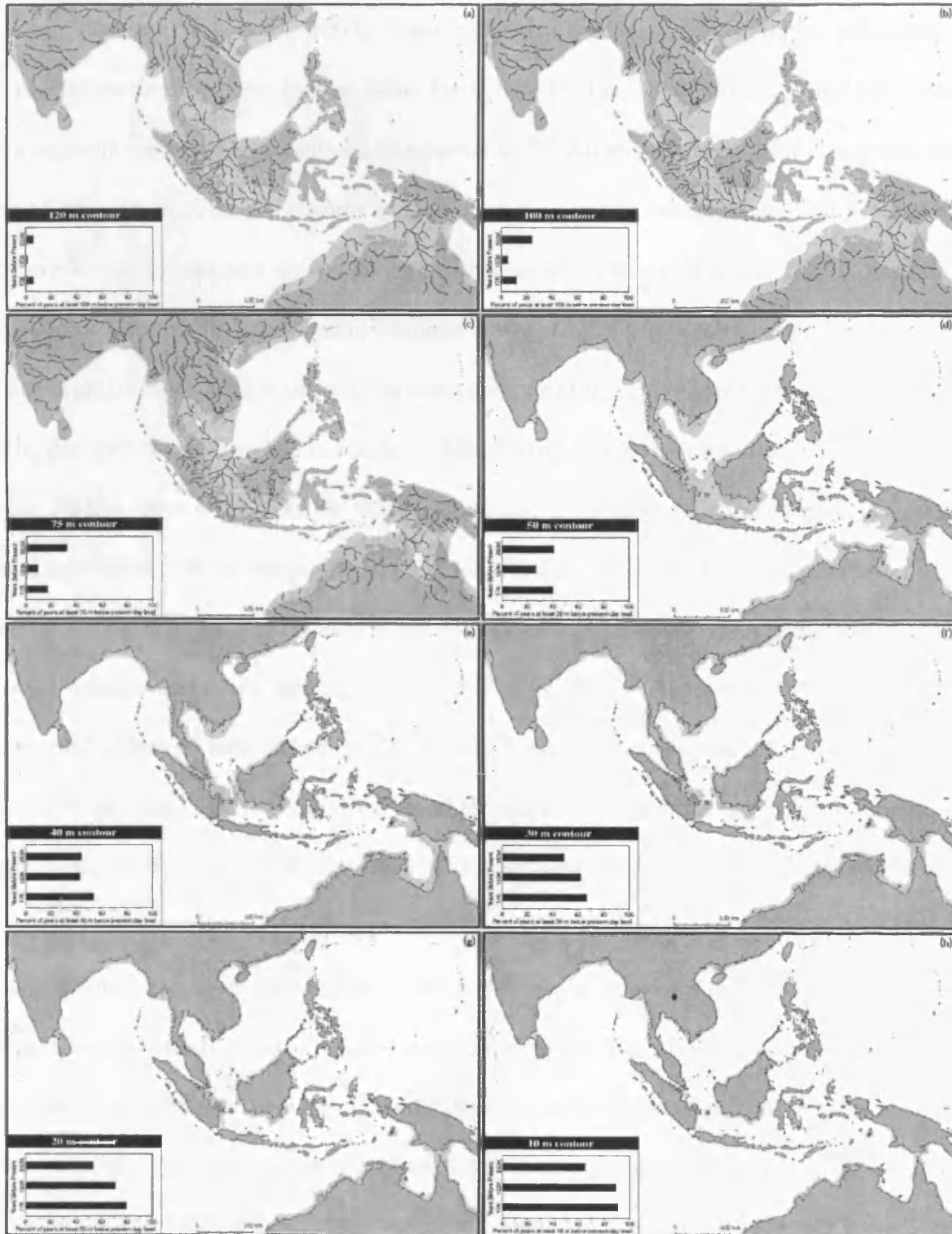
Many studies have revealed a predominance of the grassland characteristic of a savanna (Verstappen, 1997; Morley, 1998; Harrison *et al.*, 2006) on the Sunda shelf during the last glaciation whereas others indicate tropical rainforest and mangrove swamp habitats were common (e.g. Sun *et al.*, 2000). Based on geomorphology, biogeography, palynology and vegetation modelling for insular South East Asia, Bird *et al.* (2005) suggested that there was a savanna corridor from southern Sundaland to the equator. This corridor connected the areas of open vegetation at the north and south of the equator and separated forest areas of unknown extent to east and west at times of lowered sea level (Bird *et al.*, 2005). Northern Borneo, like Sumatra, probably remained moist due to rainfall from moisture coming from the sea and high altitudes, which allowed the interception of the weak summer monsoon and fog, forcing precipitation (Newsome & Flenley, 1988; Stuijts, 1993; Gathorne-Hardy *et al.*, 2002). Fossils studies have shed light on the dispersal and movement of animals colonising the Sundaland region (van de Bergh *et al.*, 2001; Harrison *et al.*, 2006). Extant animal and plant diversity found on the Sunda Islands provides clues about these historical, spatial and temporal events, especially for the so-called 'Sunda Shelf' (Harrison *et al.*, 2006). The presence of different habitats on this historical land mass determined the presence and distribution of characteristic fauna when the shelf was flooded after the last glaciation (van de Bergh *et al.*, 2001; Meijaard & Groves, 2006). Gathorne-Hardy *et al.* (2002) hypothesized, based on termite assemblage distributions, that there were Pleistocene refugia in Northern and Eastern Borneo and in Sumatra. The location of two of these refugia was supported by Meijaard's (2003) study on geographical analysis of habitat specific species, which indicated rainforest cover in Northwest Borneo, West Sumatra, but they also identified potential refugia in the Malacca Straits and around Palawan. Other areas (on and to the east of Malay/Thai Peninsula and the Java Sea, including the Sunda Strait and Eastern Borneo) may have been covered by more open vegetation types, such as savanna or open deciduous forest (Meijaard, 2003). In addition, extant and paleo plant diversity can provide clues about habitat distribution

and the nature of the associated climate during the Pleistocene and Holocene (Cannon & Manos, 2003; Brown *et al.*, 2006). The high diversity of vascular plants (~5,000 species) on Mount Kinabalu suggests that this area retained its rainforest cover during the last glacial maxima, thereby acting as a refugium (Cannon & Manos, 2003; Brown *et al.*, 2006; Grytnes & Beaman, 2006).

1.2   TROPICAL RAINFOREST AND DEFORESTATION

Tropical rainforests in South East Asia have developed during the past 65 million years or more (Heaney, 1991). During the Miocene, rainforest extended as far north as Japan but slowly contracted due to changes in climate (Heaney, 1991). The Food and Agriculture Organization of the United Nations (FAO) estimates that over the last decade more than 2.4 million ha of tropical forests have been lost each year (FAO, 2001; Stibig & Malingreau, 2003). Sodhi *et al.* (2004) predicted that South East Asia could lose three quarters of its original forests by 2100 and up to 42% of its biodiversity. The forest of Sundaland represents some of the world most valuable and productive tropical forests (Myers *et al.*, 2000; Brooks *et al.*, 2002). Borneo, as part of Sundaland, is one of the most important biodiversity hotspots in the world (Myers *et al.*, 2000; Brooks *et al.*, 2002). The rainforest of Borneo is dominated by Dipterocarpaceae (i.e. *Shorea* and *Dryobalanops* spp.), commercially valuable trees (Ahston, 1988), but is also exceptionally diverse, for example 1,175 tree species were recorded in a 0.52 km^2 survey area in Lambir National Park alone (Davies *et al.*, 1998) and in Mount Kinabalu National Park around 5,000 species of vascular plants were documented from an area of 1,200 km^2 (Beaman, 2005). Unfortunately, the logging and deforestation rate of South East Asia (0.91% per year) is three times that of South America (0.38%) (Dennis & Colfer, 2006). The tropical rainforest in Borneo has been reduced considerably due to logging, conversion to plantation, forest fires and population pressure (Fig. 1.3; Stibig & Malingreau, 2003).

Figure 1.3. Reduction in forest cover on Peninsular Malaysia, Sumatra and Borneo from (A) middle of 1980s (B) to 2000 (adapted from Stibig & Malingreau, 2003).
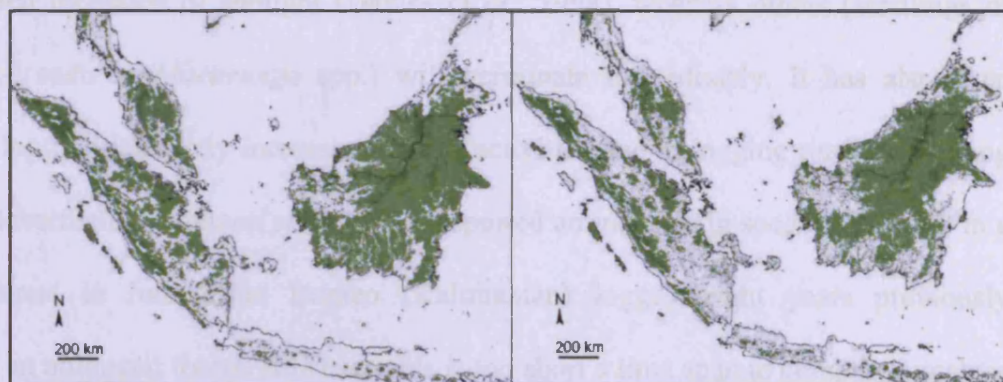
In Sabah, deforestation has reduced the proportion of forested areas from 86% in 1953 to 63% in 1984 (McMorrow & Talip, 2001) and to 59.7% in 2001 (Jomo *et al.*, 2004). Again, the main reasons for such a drastic reduction are logging and forest conversion to agriculture resulting in further isolation of forest fragments by cultivated lands. This was clearly shown by McMorrow & Talip (2001) on a land use map analysis of eastern Sabah (Fig. 1.4).

Figure 1.4. Reduction in natural forest cover in Sabah from 1975 to 1995 (adapted from McMorrow & Talip, 2001).

	Mangrove forests	Transitional, beach and swamp forests	Undisturbed high forests	Montane forest	Disturbed forests
1995 area ('000 ha)	317	193	300	700	2,799
1975 area ('000 ha)	366	203	2,800	712	1,399
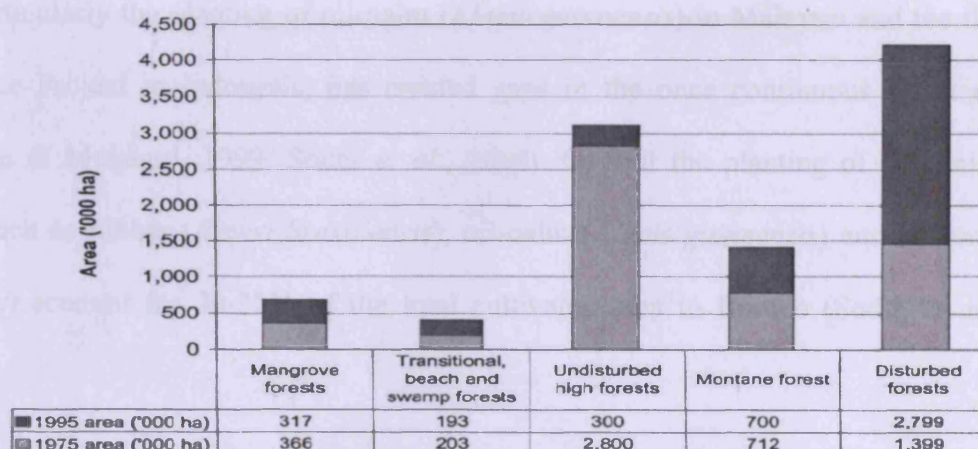
1.2.1 Logging

Recently, Curran *et al.* (2004) reported the loss of more than 56% of valuable protected forest in West Kalimantan between 1985 and 2001 due to illegal logging. The logging processes have not only increased the vulnerability of certain species by reducing

their habitat, but also altered inter-specific interactions. For example, some plants die following sudden increases of sunlight (Davies *et al.*, 1998), whereas others (seedlings or pioneer species, such as *Macaranga* spp.) will germinate immediately. It has also been suggested that logging indirectly increases hunting activities due to logging roads improving accessibility. Nevertheless, Cannon *et al.* (1998) reported an increase in species diversity in a regenerating forest in Indonesian Borneo (Kalimantan) logged eight years previously compared with an unlogged forest. However, this is too short a time span to compare a mature forest to a newly regenerating forest as many factors might have been overlooked, such as logging increasing the heterogeneity of forest microhabitats and providing niches for colonization by immigrant species. Such colonizers tend to be good dispersers and "disturbance dependent", and they replace the vulnerable taxa characteristic of primary forest (Sheil *et al.*, 1999).

1.2.2. Agriculture

Some forests have been cut down either for their hardwood or agricultural purposes. Agriculture, particularly the planting of oil-palm (*Elaeis guineensis*) in Malaysia and the ill-fated Mega Rice Project in Indonesia, has created gaps in the once continuous forest of Borneo (Rijksen & Meijaard, 1999; Sodhi *et al.*, 2004). Overall the planting of perennial export crops, such as rubber (*Hevea brasiliensis*), oil-palm (*Elaeis guineensis*) and coconut (*Cocos nucifera*) account for 20-30% of the total cultivated area in Borneo (Sodhi *et al.*, 2004).

1.2.3 Forest fires

Until recently, Borneo had only experienced two major forest fires in living memory, the first in 1982-83 and the second in 1997-98 (Dennis & Colfer, 2006). The third one is currently on-going as this PhD is being written. Based on time series satellite imagery from

1983 to 2000, Dennis & Colfer (2006) studied the impact of forest fires on land use in East Kalimantan. After the first forest fire, the authors observed a recovery in burnt forest but the effects of the second forest fire coupled with changes of land use completely degraded the recovering habitat. A second study on forest cover change also using satellite imagery was carried out by Trigg *et al.* (2006) in West Kalimantan in Gunung Palung National Park. Using data obtained from 1988 to 2002, they revealed accelerating deforestation, with 70% of the buffer zones surrounding the national park being cleared (logged) over a period of 14 years.

1.2.4 Impact of deforestation on biodiversity in Borneo

Deforestation creates a mosaic of habitats which includes fragments of forest and open scrub-land, resulting in a matrix of habitats isolated or semi-isolated from each other (Fahrig, 2003; Wright, 2005). The impact of habitat fragmentation varies widely in faunal assemblages (Turner & Corlett, 1996; Fahrig, 2003; Kinnaird *et al.*, 2003; Wright, 2005). Kinnaird *et al.*'s (2003) long term study on large mammals in Sumatra showed that animals with large home ranges, particularly those which avoided human contact, were forced into suboptimal conditions as the forest dwindled in area, sometimes leading to their demise. Reduction in habitat size increases the chances of human-animal conflict, especially where animals raid food crops to supplement their diets. Opportunistic and/or versatile species tend to be less affected than niche specialists (Webb *et al.*, 2002; Pribil & Houlahan, 2003). Studies assessing the impact of forest disturbance on biodiversity are seldom conclusive (Fahrig, 2003; Turner & Corlett, 1996; Kinnaird *et al.*, 2003; Wright, 2005; Meijaard *et al.*, 2006), nevertheless over time a decline in species richness deterministically accumulates in fragments isolated from continuous forest (Turner & Corlett, 1996).

Recently, Meijaard *et al.* (2006) reviewed the impact of timber concessions on Bornean wildlife and found that different animals respond differently to logging activities.

Some ungulates, including the rare banteng (*Bos javanicus*), seem to benefit from road building activity, being attracted to grazing on the herbaceous vegetation growing along roadsides (Meijaard *et al.*, 2006). The Malay civet (*Viverra tangalunga*) in Danum Valley, Sabah, was not drastically affected by logging, but the densities of the animals were higher in unlogged compared to logged forest (Colón, 2002). Other species, such as hornbills, experience reduced survival in logged areas due to loss of breeding sites (holes in large trees) and food resources (Meijaard *et al.*, 2006). The survival of two mouse deer species (*Tragulus tragulus* and *T. napu*) was also affected by logging activities (Heydon & Bulloh, 1997), and both *Hylobates lar* and *Presbytis melalophos* decreased their level of activity possibly to conserve energy (Johns, 1986). For primates, which need high quality food, logging usually reduces the diversity of food resources available, making foraging more difficult. However, most of these studies have only considered short-term effects (up to three years), seldom are the long-term effects of logging considered (Meijaard *et al.*, 2006).

Habitat loss and fragmentation has resulted in the restriction of many species to small habitat patches separated by a matrix of inhospitable environments. Adverse impacts on genetic diversity include a reduction in local population size, reduced migration, stochastic population dynamics (genetic drift) and inbreeding depression (Avise *et al.*, 1987; Frankham *et al.*, 2002; Avise, 2004; Frankham, 2005). Short term effects of habitat fragmentation and isolation are influenced by effective population sizes and reproductive success, controlled by factors such as the availability of mates and food, and predation and parasitism pressure. Long term consequences may include genetic drift, mutation and fixation of certain haplotypes.

1.3   PRIMATES OF SUNDALAND

There are nine genera (i.e. *Pongo, Hylobates, Macaca, Presbytis, Trachypithecus, Nasalis, Simias, Tarsius* and *Nycticebus*) of extant primates found in Sundaland of which 13

species are found in Borneo (Harrison *et al.*, 2006) and ten of these occur in the Lower Kinabatangan Wildlife Sanctuary in Sabah, Malaysia. Amongst these ten Kinabatangan species, three were selected for the current study, an ape (orang-utan, *Pongo pygmaeus*), a colobine (proboscis monkey, *Nasalis larvatus*) and a cercopithecine (long-tailed macaque, *Macaca fascicularis*). The orang-utan is a large-bodied, long-lived, slow reproducing animal that usually occurs at low population densities and is thus potentially more prone to extinction due to demographic factors than more rapidly reproducing species (Harcourt, 1999; Purvis *et al.*, 2000). In contrast, long-tailed macaques (and to a lesser extent, proboscis monkeys), occur at higher densities, have shorter life spans and a higher rate of reproduction and therefore are potentially able to adapt faster to environmental changes. These three species are predicted to show different responses to forest fragmentation, particularly as orang-utans and proboscis monkeys are unable to tolerate human proximity whereas long-tailed macaques are known to exploit humans for food (i.e. by crop-raiding). Based on mobility, since orang-utans cannot swim and both long-tailed macaques and proboscis monkeys (especially) can swim, the Kinabatangan River and its tributaries potentially provide barriers to movement between forested areas. Deforestation also degrades the quality of the forest affected thus influencing food availabilities for these primates (Johns, 1986; Bawa & Seidler, 1998; Chapman *et al.*, 2000).

1.3.1 The orang-utan

The orang-utan, the largest arboreal ape, is found on Sumatra and Borneo in South East Asia. Traditionally, the orang-utan was classified as two subspecies, *P. pygmaeus pygmaeus* in Borneo and *P. pygmaeus abelii* in Sumatra. However, recent molecular data has led to the re-classification of the orang-utan into two distinct species, *P. pygmaeus* and *P. abelii* (see Xu & Arnason, 1996, but see Muir *et al.*, 1998; 2000; Zhang *et al.*, 2001). Furthermore, based on mitochondrial control region DNA data, Warren *et al.* (2001)

suggested there are four distinct phylogenetic groupings of the Bornean orang-utans, corresponding to (i) Sabah, (ii) Sarawak and Northwest Kalimantan, (iii) Southwest and Central Kalimantan, and (iv) East Kalimantan populations.

A recent Sabah-wide survey indicated that there are about 11,000 orang-utans in the state with most individuals recorded outside protected areas (Ancrenaz *et al.*, 2005; Fig.1.5). There are two major concentrations of orang-utans in Sabah, namely Segama (4,500 individuals) and on the north-side of the upper Kinabatangan River (1,700 individuals); both areas comprise of logged, commercial forest reserves (except the Danum Valley Conservation Area, which lies within Segama). Other noteworthy populations are in the Tabin Wildlife Reserve (1,400 individuals), Kinabatangan Wildlife Sanctuary (1,100 individuals, see also Ancrenaz *et al.*, 2004) and Kulamba Wildlife Reserve (500 individuals) (Ancrenaz *et al.*, 2005).

Figure 1.5 Distribution of the five major orang-utan populations in Sabah (A - Segama; B - Upper Kinabatangan; C - Tabin Wildlife Reserve; D - Lower Kinabatangan Wildlife Sanctuary; E - Kulamba Wildlife Reserve (modified from Ancrenaz *et al.*, 2005).
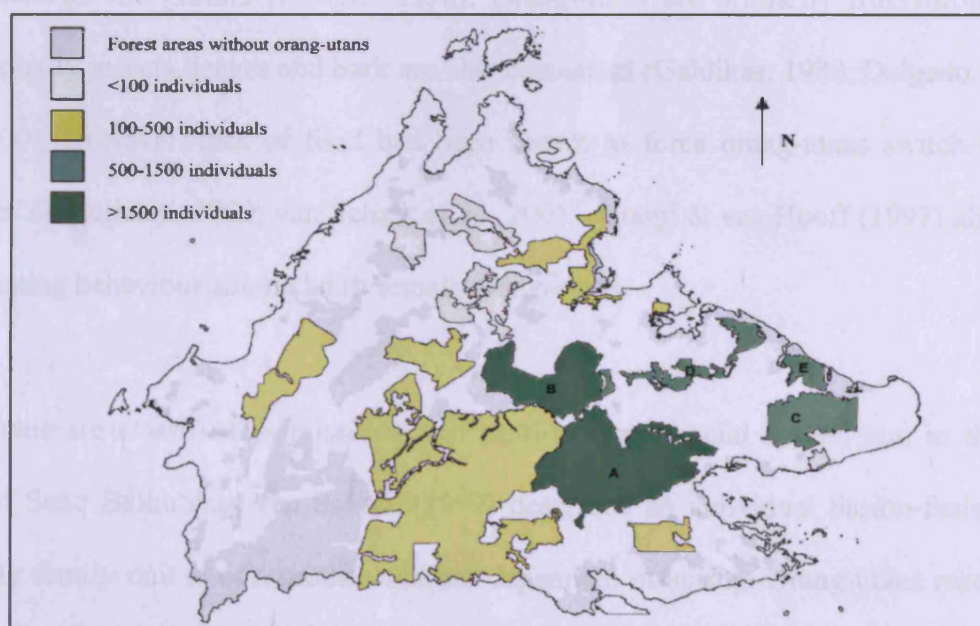
There is a high degree of sexual dimorphism in orang-utans. Average body mass for an adult orang-utan is 86 kg for males and 38 kg for females (Delgado & van Schaik, 2000). Recent studies show that sexually mature males have two distinct physical morphs (bimaturism), flanged and unflanged (Delgado & van Schaik, 2000; Utami et al., 2002). Flanged males have throat sacs which they use to emit 'long calls' to advertise their presence, assert their dominance and territorial rights (Delgado & van Schaik, 2000). Male orang-utans are assumed to be larger than females because of male-male competition and sexual selection. Large, flanged males displace smaller conspecifics for access to females for mating, although forced copulation has been occasionally observed by unflanged males (Rodman & Mitani, 1987; Delgado & van Schaik, 2000). Utami et al. (2002) and Goossens et al. (2006a) showed that both flanged and unflanged males are reproductively successful.

Orang-utans are diurnal, building nests made of branches and leaves each night (MacKinnon & MacKinnon, 1974). During the day, they move through the forest using inter-connected branches or by bending a branch using their weight to form an inter-crown pathway (Fleagle, 1998; Felton et al., 2003). Only when there is a large gap between trees will an orang-utan descend to the ground (Fleagle, 1998). Orang-utans are primarily frugivorous, although occasionally insects, leaves and bark are also consumed (Galdikas, 1988; Delgado & van Schaik, 2000). However, lack of food had been shown to force orang-utans switch to folivory (Rijksen & Meijaard, 1999; van Schaik et al., 2001). Utami & van Hooff (1997) also recorded meat eating behaviour among adult female orang-utans.

Orang-utans are classified as possessing a "semi-solitary" social system but, in the swamp forest of Suaq Balimbing, van Schaik (1999) described an individual fission-fusion system. The only family unit is of a mother and her dependent offspring. Orang-utans reach sexual maturity after 8 to 10 years for males and 11 to 15 years for females (Delgado & van

Schaik, 2000). Orang-utans have a gestation period of almost nine months and give birth to a single offspring every five to eight years. Females will often settle in an area near their mother's range (philopatry) whereas males tend to disperse (Rodman & Mitani, 1987; Delgado & van Schaik, 2000, but see Goossens *et al.*, 2006a). Male territories overlap with one or more females. Home range sizes range from 64 - 600 ha (Borneo) and 300 - 900 ha (Sumatra) in females and 500 - 600 ha (Borneo) and 800 - 4000 ha (Sumatra) in males (Delgado & van Schaik, 2000). Orang-utans were thought to live up to 45 years in the wild (Delgado & van Schaik, 2000) but Wich *et al.* (2004) recently estimated that Sumatran male and female orang-utans can live at least 58 and 53 years, respectively.

A dramatic population decline in Sabah's orang-utans from ca. 315,000 individuals in 1900 to 27,000 in 1997 has been estimated by Rijksen & Meijaard (1999), while Wich *et al.* (2003) indicated the number of wild Sumatran orang-utans was as low as 3,500 by the end of 2002. Major threats to orang-utans include habitat destruction, particularly due to logging and anthropogenic disturbance (e.g. hunting and the pet trade) (Rijksen & Meijaard, 1999; Robertson & van Schaik, 2001; van Schaik *et al.*, 2001). Habitat degradation following logging activities has several detrimental effects on orang-utan survival. Destruction of canopy areas used by orang-utans as their pathway throughout the forest restricts their movement and increases their vulnerability to predators (Rijksen & Meijaard, 1999; van Schaik *et al.*, 2001; Felton *et al.*, 2003). Large scale commercial logging activities with conversion to agricultural estates (particularly oil-palm plantations) further reduces the orang-utan's food resources. It also increases contact with humans, and orang-utans that forage in plantation and cultivated lands are often killed (Rijksen & Meijaard, 1999). Logging and land conversion also opens up adjoining forest areas to increased illegal hunting and poaching due to the greater accessibility offered by logging roads (Rijksen & Meijaard, 1999; van Schaik *et al.*, 2001; Felton *et al.*, 2003). Finally, forest fires caused either by prolonged dry seasons

(predominantly due to the El Niño Southern Oscillation) or from the clearing of forest patches for small scale agriculture also play an important role in orang-utan population declines as these slow moving apes are unable to escape rapidly moving fires. For example, a recent study showed a high mortality rate for orang-utans during the 1997-1998 forest fires in Kalimantan (Rijksen & Meijaard, 1999).

Genetic studies on orang-utans to date have been rather limited, with most studies focusing on systematics (Xu & Arnason, 1996; Muir *et al.*, 2000; Zhang *et al.*, 2001; Zhi *et al.*, 1996); rather than genetic structure (Warren *et al.*, 2000; 2001; Goossens *et al.*, 2005; Kanthaswamy *et al.*, 2006). Furthermore, a large proportion of these studies have utilised invasive samples from zoo and rehabilitation centres, compromising the geo-referencing of data produced, with very limited non-invasive sampling (i.e. hair and faeces) from wild animals. Only Goossens *et al.* (2005, 2006b), studying a fragmented orang-utan population in the Lower Kinabatangan floodplain, completely utilised non-invasive samples from wild orang-utans.

1.3.2 The long-tailed macaque

The genus *Macaca* is the most widespread of the Cercopithecinae (Tosi *et al.*, 2000; Abegg & Thierry, 2002). Members of this genus can be found from North Africa (i.e. *Macaca sylvanus*) to Japan (i.e. *Macaca fuscata*). Of the 20 recognised species, 19 occur in Asia. The divergence of the genus *Macaca* from the tribe Papionini dates back seven MYA based on fossil and molecular data (Tosi *et al.*, 2000; Abegg & Thierry, 2002) and major diversification of this genus began around five MYA, resulting in three lineages, *silenus-sylvanus*, *sinica-artoides* and *fascicularis* (see Abegg & Thierry, 2002). The long-tailed macaque (*Macaca fascicularis*) is grouped within the *fascicularis* lineage which includes three other species, *M.*

mulatta (rhesus macaque), *M. fuscata* (Japanese macaque) and *M. cyclopis* (Formosan macaque) (Tosi *et al.*, 2000; Abegg & Thierry, 2002).

Long-tailed macaques are found from Burma and Indochina in the north, to Bali on the extreme east, and Sumatra and Nicobar on the extreme west. They are neither rare nor threatened in their native range (Abegg & Thierry, 2002; Umapathy *et al.*, 2003). In fact, long-tailed macaques 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 (Lowe *et al.*, 2000; Long, 2003).

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), often raiding crops where they are classified as a pest (van Schaik *et al.*, 1996; Abegg & Thierry, 2002). In Sumatra, long-tailed macaques occur at high densities in selectively logged forest, secondary forest and cultivated land (Supriatna *et al.*, 1996). They are omnivorous and opportunistic feeders. The average body mass for an adult male and female long-tailed macaque is around 6 and 3.3 kg, respectively (Harcourt & Schwartz, 2001). They are diurnal and can be totally or semi- arboreal. They move through the forest canopy and on land quadrupedally, but are also considered good swimmers (Richard, 1985).

The social system of the long-tailed macaque is multi-male, multi-female with an average group size of about 30 individuals (de Ruiter, 1994; Harcourt & Schwartz, 2001). Female macaques usually remain in their natal group (philopatry) and males disperse (de Ruiter, 1994). There is a hierarchical system amongst group members based upon matrilines (de Ruiter & Geffen 1998). Dispersal of female long-tailed macaques occurs most commonly

by group fission, where low ranking females split off to form a new group. Mating is promiscuous and dominated by the alpha male (de Ruiter, 1994, de Ruiter & Geffen 1998).

Macaques in general have been widely studied, rhesus macaques being particularly useful in medical research. The population genetic structure of long-tailed macaques has been studied mostly in Indonesia. de Ruiter (1994), studying the social aspect of Sumatran long-tailed macaques in Ketambe, found that large differentiation between adjacent social groups was caused by social structure rather than distances. However, for sites other than the main study population in Ketambe, distance plays an important role for differentiation. Perwitasari-Farajallah *et al.* (1999) studying the Javanese long-tailed macaques found high genetic diversity among local populations and, like de Ruiter (1994), found a significant positive correlation with geographic distance.

1.3.3 The proboscis monkey

The proboscis monkey (*Nasalis larvatus*), known locally as Bangkatan or the Dutch man, is endemic to Borneo with no evidence that it formerly occurred elsewhere (Payne & Francis, 1998; Harcourt & Schwartz, 2001). This primate is adapted to nipa-dominated mangrove, mangrove, peat swamp, riverine and lowland forest (Kawabe & Mano, 1972; Meijaard & Nijman, 2000). Currently, the proboscis monkey is threatened by habitat destruction and hunting (Meijaard & Nijman, 2000) 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). Hunting is much in evidence in Sarawak and Kalimantan (Meijaard & Nijman, 2000). The proboscis monkey is currently classified by IUCN as 'vulnerable' (IUCN, 2000) and is protected by law throughout its range (Meijaard & Nijman, 2000).

Karyological studies split the genus *Nasalis* (2n = 48) from all other Asian colobines

(2n = 44) (Disotell, 1996; Page *et al.,* 1999; Bigoni *et al.,* 2003). However, this observation is

incongruent with morphological and molecular analyses which group *Nasalis* with all other

Asian colobines (Collura *et al.,* 1996; Page *et al.,* 1999; Bigoni *et al.,* 2003). For example,

based on a γ-Globin DNA phylogeny, *Nasalis* occurs in the same clade as *Trachypithecus*

obscurus with estimated divergence times between *Nasalis-Trachypithecus* of 5-6 million

years (Page *et al.,* 1999). Recently, Sterner *et al.* (2006) estimated the divergence time

between African and Asian colobines to be 10.8 MYA with Asian colobines diversifying at

about 6.7 MYA.

The proboscis monkey has been well documented in terms of its behavioural ecology

and conservation needs within its range in the Malaysian states of Sarawak and Sabah

(Kawabe & Mano, 1972; Salter *et al.,* 1985; Salter & MacKenzie, 1985; Bennett & Sebastian,

1988; Bernard, 1997; Boonratana, 2002; 2003; Murai, 2004; 2006, Murai *et al.,* 2006), Brunei

(Macdonald, 1982; Yeager, 1995) and Kalimantan, Indonesia (Yeager, 1989; 1992; Meijaard

& Nijman, 2000). There are two elements to the social system; harems (consisting of one

adult male and several females) and all-male groups (consisting of juveniles, adolescents and

adult males) (Bennett & Sebastian, 1988; Yeager, 1995; Boonratana, 2002; Murai, 2004).

Proboscis monkeys live in groups of three to 32 individuals. Males as young as 18 months

will leave their natal group and join an all male group (Bennett & Sebastian, 1988). 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 *et al.,* 2006). In the

Kinabatangan, Murai (2004) reported that females can join all-male groups temporarily and

will copulate with them. Murai (2006) also found that females are promiscuous. Both 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). Despite considerable knowledge of this species, there are currently no population genetic studies of this endangered colobine.

1. 4 PHYLOGEOGRAPHY AND CONSERVATION GENETICS

Phylogeographic studies examine biogeographic patterns and infer population processes from an evolutionary (historical) perspective (Avise *et al.*, 1987). Biogeographic barriers, such as rivers (Telfer *et al.*, 2003; Eriksson *et al.*, 2004; Goossens *et al.*, 2005) and mountain ranges (Hewitt, 2000), may disrupt gene flow, thus moulding genetic structure. Nevertheless for widespread species, genetic divergence may simply be a function of distance due to limited gene flow (Avise *et al.*, 1987). At a regional level, comparative phylogeography can be used for conservation purposes to localize areas of high diversity and hence high conservation value (Moritz & Faith, 1998). Intraspecific phylogeography can contribute to our understanding of fragmentation by analysing evolutionary relationships between haplotypes and geographical distribution of haplotype variants over a variety of timescales. Historical demography influences genetic structure and hence intra-specific phylogenies (Avise, 2004).

The application of molecular markers in a conservation context ('conservation genetics') aims to describe and understand the origin and maintenance of genetic variation and to use this information to minimize extinction risks due to genetic factors (Frankham, 1996; 2005). Large-bodied, slow-reproducing, solitary animals, such as the orang-utan which occurs in low densities, are predicted to be much more vulnerable to extinction compared to smaller bodied, rapidly reproducing species, such as long-tailed macaques (Frankham, 1996; 2005). At low densities, organisms may experience a reduction in genetic variability and thus be less capable of adapting to changes in environment (i.e. climate change, exposure to disease, etc). Genetic variability may be lost due to a combination of founder effects,

inbreeding, genetic drift and population bottlenecks. However, populations with a large effective population size are usually safeguarded against these factors (Frankham, 1996; 2005).

1.5   MOLECULAR MARKERS

Mitochondrial (mt) DNA has proved a reliable workhorse for intraspecific phylogeographic studies (Avise, 2000, 2004). However, a variety of mitochondrial and nuclear markers are available for phylogeography and conservation genetic studies depending on questions that need to be answered (Wan *et al.*, 2004). Here, I discuss mtDNA and nuclear markers in light of their suitability for the current study.

1.5.1   mtDNA

The use of mtDNA has become standard in phylogenetic and population genetic studies, first encouraged by developments in methodology for mtDNA isolation and the use of restriction enzymes to detect nucleotide polymorphisms (Lansman *et al.*, 1981), and subsequently by the development of PCR methodology and the application of 'universal' PCR primers (Kocher *et al.*, 1989) for amplification of mtDNA. Much of this interest is related to the rapid rate of evolution of mtDNA compared to nuclear genes. Higher numbers of nucleotide substitutions accumulate in the mitochondrial compared to the nuclear DNA genome (Brown, 1979) probably due to inefficiency of DNA repair and replication errors. MtDNA is haploid and (almost) exclusively maternally inherited. Compared to diploid nuclear autosomal genes with biparental transmission, the effective population size of mtDNA is approximately 25% (Moore, 1995). Therefore, mtDNA phylogeny is more likely to be congruent with a species phylogeny due to a high probability of coalescence (convergence of lineages to a point in the past i.e. the most recent common ancestor) even when speciation events have occurred within short time-periods. Mitochondrial genes are inherited as a single

linkage group in the absence of recombination (Hayashi *et al.*, 1985; Hoech *et al.*, 1991). Animal mtDNA is a circular molecule of 15-20 kb and in vertebrates contains genes for 22 transfer-RNAs, 2 ribosomal-RNAs and 13 messenger-RNAs coding for proteins involved in electron transport and oxidative phosphorylation (Fig 1.6; Moritz *et al.*, 1987; Sorensen *et al.*, 1999; Ballard & Rand, 2005). The only major non-coding region of the mtDNA is the control region, typically 1 kb in length, involved in the regulation and initiation of mtDNA replication and transcription (Moritz *et al.*, 1987; Ballard & Rand, 2005).
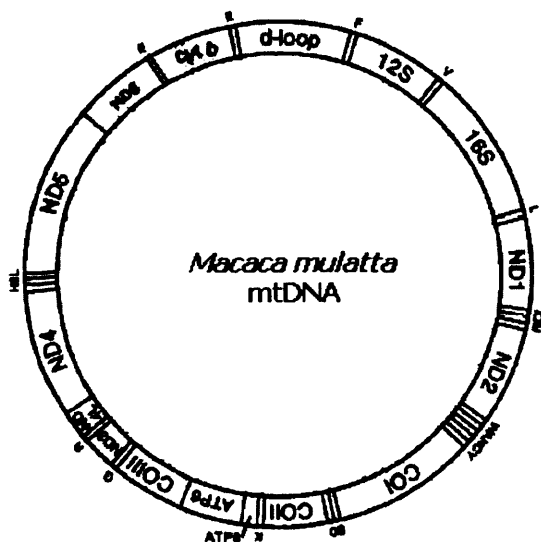
Figure 1.6. Schematic diagram of *Macaca mulatta* mtDNA showing gene organisation (adapted from Sorensen *et al.*, 1999).

The vertebrate control region (typically around 1 kb) is involved in the regulation and initiation of mtDNA replication and does not code for proteins and has only a few short sequence blocks conserved among taxa (Fumagalli *et al.*, 1996; Avise, 2004). The human mtDNA control region contains three segments, HVI, HVII and HVIII. HVI (otherwise known as the left domain) is more polymorphic than HVII (right domain), with HVIII showing the lowest polymorphism (Lutz *et al.*, 1998). The regions separating the segments are relatively conserved. For example, in the black howler monkey (*Alouatta caraya*) the control region contains two highly divergent peripheral (left and right) domains flanking a conserved region (Ascunce *et al.*, 2003). Within HVII and HVIII, a short conserved block reduces the polymorphism of these domains. The left domain is always more variable than the

right domain (e.g. Ascunce *et al.*, 2003). However, for this study, I test the utility of using the left and right hypervariable domains of the mtDNA control region. I will test this on long-tailed macaque (see Chapter 3) and based on the results will influence the choice of domains for subsequent studies.

1.5.2 Microsatellites

Microsatellites are nuclear markers, composed of 1-6 base pairs arranged as tandem repeated motifs (Tauntz & Renz, 1984). They are most abundant in non-coding regions of the genome and possess a high mutation rate (Hancock, 1999; Zane *et al*, 2002). Since their first description (Tautz, 1989; Weber & May, 1989), microsatellites have been detected in both eukaryotic and prokaryotic genomes (Zane *et al.*, 2002) and been widely used as genetic markers in different kinds of studies. For example, Goossens *et al.* (2000, 2005, 2006a,b) used human-derived microsatellites in their studies on orang-utans in Sumatra, Indonesia and Sabah, Malaysia. The wide applicability of microsatellite markers spans areas such linkage mapping (Straub *et al.*, 1993), population and conservation genetics (Goossens *et al.*, 2005, 2006b), forensic DNA studies (Halos *et al.*, 1999), paternity analyses (Launhardt *et al.*, 2001), identifying individuals (Maudet *et al.*, 2002), human origin reconstruction and hybridization (e.g. Goldstein & Schlotterer, 1999; Tautz, 1989; Weber & May, 1989; Zane *et al.*, 2002). Major advantages of microsatellites include their high allelic variability and the fact that primers may cross amplify between different species (Primmer *et al.*, 1996), sometimes even between species diverging millions of years ago (Zane *et al.*, 2002).

1.6 NON-INVASIVE SAMPLES FOR ENDANGERED MAMMALS

Studying endangered species requires an approach to minimize the handling of animals in order to obtain samples. Non-invasive sampling techniques provide alternative methods for sample collection, ideal for studying endangered species, and have already

proved successful for studying vulnerable populations, such as the giant panda (Zhan *et al.*, 2006) and the bonobos (Eriksson *et al.*, 2004). A range of different source materials collected by non-invasive sampling have been successfully tested, including plucked hair (Li *et al.*, 2005), shed hair (Morin *et al.* 1993 in chimpanzees), shed feathers (Taberlet & Bouvet, 1991), sloughed/shed skin (Baker *et al.*, 1991; Hauser *et al.*, 2000), faeces (Roeder *et al.*, 2004) and egg shell (Strausberger & Ashley, 2001) (for more details see Taberlet & Luikart, 1999; Avise, 2004).

Goossens *et al.* (2005, 2006a,b) demonstrated the utility of non-invasive samples whilst studying a wild orang-utan population in Sabah. A large sample size (>200 samples) was obtained without disturbance to the Kinabatangan orang-utan population. The results of the study suggested that the Kinabatangan River has had a major impact on gene flow and dispersal behaviour of individuals (Goossens *et al.*, 2006a). Despite, the relatively small census size of the remaining population (approximately 1,050 individuals), Kinabatangan orang-utans retain relatively high levels of genetic diversity (Goossens *et al.*, 2005) although analysis of microsatellite allelic spectra also shows that that the population recently underwent a collapse (Goossens *et al.*, 2006b) and this variation might not be retained in the long-term unless conservation measures are immediately put in place.

Studying endangered species (such as the orang-utan) requires different approaches to studying common species. Firstly, endangered species are usually found in small populations and are easily affected by human disturbance (i.e. Amur tiger, Kerley *et al.*, 2002). Secondly, endangered species are protected by law, thus invasive techniques for obtaining samples is not an option and CITES regulations prevent the samples being exported to other countries. Non-invasive samples circumnavigate this problem (Zhan *et al.*, 2006). This innovation has prompted an increase in the study of protected animals using non-invasive samples not

protected by local and international law. Thus, studies on endangered species, such as gorillas, chimpanzees, orang-utans, giant pandas, and whales, have been made possible with minimal disturbance to the individual (Valsecchi *et al.*, 1998; Morin *et al.*, 2001; Clifford *et al.*, 2004; Goossens *et al.*, 2005; Zhan *et al.*, 2006). Previously, animals had to be tracked, darted and/or handled to obtain biopsy or blood samples (Nyakaana & Arctander, 1999). The current trends of using non-invasive samples have advanced so much that as little as 100 g of faeces is now a valuable source of genetic material. The improvement of techniques for extracting DNA using commercial kits (e.g. the QIAGEN DNA stool mini kit) have also increased the popularity of non-invasive samples. There are many methods used to extract DNA from different kinds of samples and the success rate depends on the suitability of methods used (see Waits & Paetkau, 2005). Jeffery (2003) investigated the utility of shed/plucked hair samples for gorilla conservation in Gabon. She discovered that whilst plucked hairs have a higher success rate than shed hairs, the success rate for the latter was halved from 56% for freshly shed hair (within 24h) to 26% for 3 day old hair (Jeffery, 2003). Recently, an astonishing new non-invasive technique, using a blood sucking bug (*Dipetalogaster maximus*), has been used to obtain high quality DNA from birds (Becker *et al.*, 2006). The bugs (second and third instar) were housed in an empty shell and were collected after the parent birds vacated the nest (Becker *et al.*, 2006) (Fig. 1.7). This has transformed the possibility of obtaining non-invasive samples to include DNA rich blood samples without invasive procedures by sampling ectoparasitic insects.

Figure. 1.7. Semi-invasive method for obtaining avian blood sample. The dummy eggs in the bird nest containing an instar of *Dipetalogaster maximus*. The *Dipetalogaster maximus* after the blood meal (Becker *et al.*, 2006).
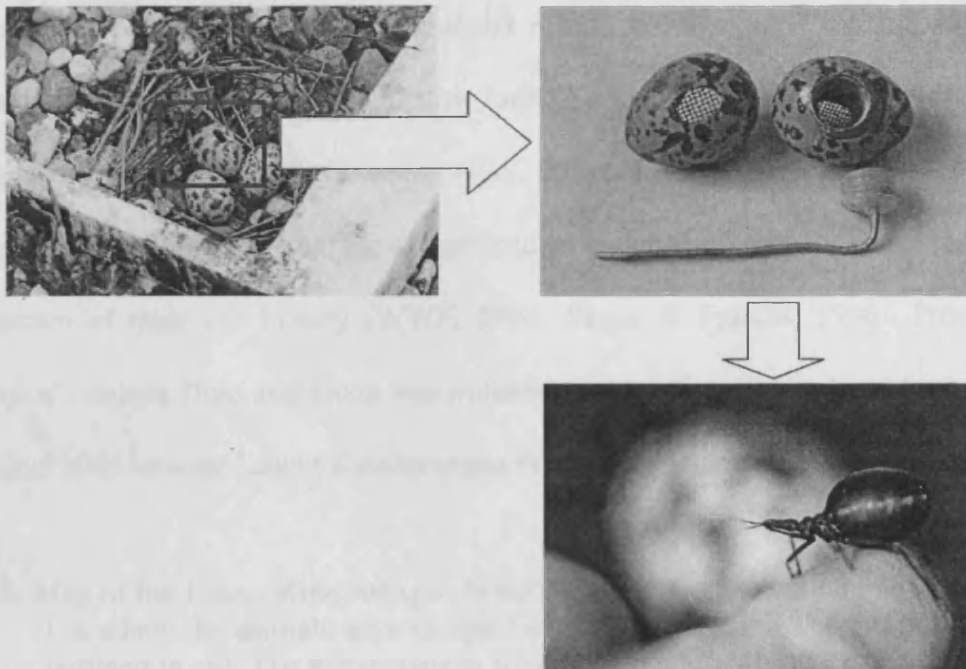
1.7 THE LOWER KINABATANGAN WILDLIFE SANCTUARY

The Kinabatangan River at 560 km is the longest river in Sabah and is also one of Borneo's few navigable rivers (WWF, 1998). The Kinabatangan floodplain is a mosaic of heterogeneous habitats consisting of mangrove, freshwater swamp, riverine, seasonally flooded forest and dry lowland dipterocarp forest (Azmi, 1998; WWF, 1998). Humans have been active here since at least the 7th century AD and the natural habitats have undergone considerable long term changes. However, since the 1950s, when logging licenses were first issued, the environment has significantly deteriorated. In addition, from the 1970s, agricultural has become more important to the Malaysian economy and 1980 saw the opening of the first oil-palm plantation. The subsequent growth of the oil-palm industry has resulted in marked fragmentation of the forested habitats (WWF, 1998). Nonetheless, the Kinabatangan floodplain still harbours many unique plants and animals, with two of the important inhabitants of Kinabatangan being the orang-utan and proboscis monkey (WWF, 1998).

Besides the orang-utan and proboscis monkey, LKWS is home to eight other species of primate; Bornean gibbon (*Hylobates muelleri*), long-tailed macaque, pig-tailed macaque (*Macaca nemestrina*), Hose's langur (*Presbytis hosei*), red langur (*Presbytis rubicunda*), silvered langur (*Semnopithecus cristata*), slow loris (*Nycticebus coucang*) and western tarsier (*Tarsius bancanus*) (WWF, 1998; Goossens *et al.*, 2003). The orang-utan, proboscis monkey, Bornean gibbon and Hose's langur are categorised as endangered, but for the latter too very little is known of their life history (WWF, 1998; Payne & Francis, 1998). Protection of Kinabatangan's unique flora and fauna was initiated by gazetting 10 blocks or Lots (Lots 1 – 10) in August 2005 into the Lower Kinabatangan Wildlife Sanctuary (LKWS) (see Fig. 1.8).

Figure 1.8. Map of the Lower Kinabatangan Wildlife Sanctuary showing the location of each Lot where the animals were sampled in the current study. The Lot boundaries are outlined in red. The Kinabatangan River is marked by the thick blue line with the thinner blue lines representing its tributaries. For detailed descriptions of the habitats see Ancrenaz *et al.* (2004).
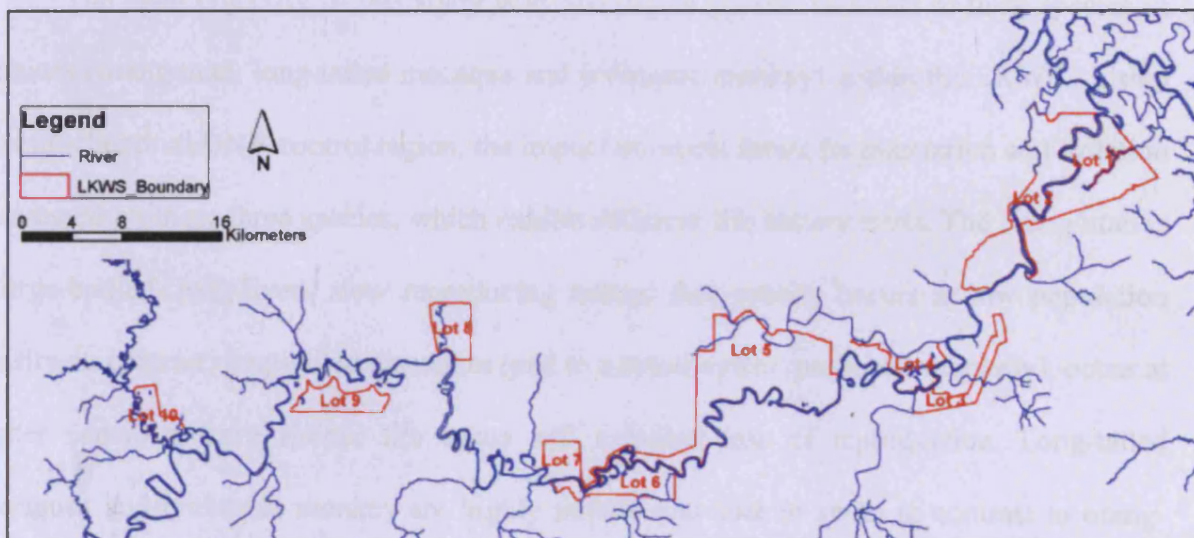
Amongst the ten species of primates in Kinabatangan, only the orang-utan and proboscis monkey have been widely studied. Ancrenaz *et al.* (2004) carried out the most recent census of the orang-utan within the LKWS and Goossens *et al.* (2005, 2006b) were the first to investigate the population genetics of LKWS orang-utans. The proboscis monkey has been studied in Kinabatangan since the 1970's. Kawabe & Mano (1972) investigated the

ecology and behaviour of this species in Kinabatangan. They noted troop sizes which ranged from 11 to 32 individuals, falling within the range obtained by Kern (1964) in Padas Bay, Brunei. More recent studies on feeding ecology and ranging behaviour of the Kinabatangan proboscis monkey were carried out from 1990 to 1991 (Boonratana, 1993; 2000). Recently, Murai (2004) described the social behaviour of proboscis monkey all-male groups found in Kinabatangan which he had monitored from 1999 to 2002. Only one study has attempted to estimate the population sizes of the long-tailed macaque and the proboscis monkey in Kinabatangan to date, when Goossens *et al.* (2003) estimated the population sizes of proboscis monkey based on their densities as at 3,430 individuals and long-tailed macaques at 3,170 individuals.

1.8 AIMS

The main objective of this study is to investigate genetic variation of three species of primates (orang-utan, long-tailed macaque and proboscis monkey) within the LKWS. Using the mitochondrial DNA control region, the impact of recent forest fragmentation and isolation is assessed on these three species, which exhibit different life history traits. The orang-utan is a large-bodied, long-lived, slow reproducing animal that usually occurs at low population density, in contrast, long-tailed macaques (and to a lesser extent, proboscis monkeys), occur at higher densities, have shorter life spans and a higher rate of reproduction. Long-tailed macaques and proboscis monkey are highly mobile and able to swim in contrast to orang-utans which are bulky and cannot swim thus restricting movement across the rivers. Deforestation also degrades the quality of the forest affected thus influencing the food availabilities for these primates (Johns, 1986; Bawa & Seidler, 1998; Chapman *et al.*, 2000). Each of the species selected has different diets, with orang-utans and long-tailed macaques being versatile feeders, whereas proboscis monkeys are specialist folivores, thus food availability is predicted to limit populations, particularly of proboscis monkeys (Wasserman

& Chapman, 2003). The combination of these factors (i.e. primate life histories traits, behaviour etc.) allows the following hypotheses to be tested.

Hypothesis I: Habitat fragmentation has resulted in a reduction of effective population size (N_e) and genetic drift in each of the primate populations. If N_e reduction has occurred, then species with long lifespan and slow reproduction rates (i.e. the orang-utan) will show a lower reduction of genetic diversity than species with a short lifespan and faster reproduction rates (i.e. macaque and proboscis monkey). However, this will also be influenced by population size, macaques and proboscis monkeys being three times more abundant than orang-utans (see Chapters 3, 4 & 5).

Hypothesis II: The Kinabatangan River is a barrier to primate gene flow. It is predicted that levels of genetic differentiation in primate populations on either side of the river will be greater than between populations on the same side of the river, the magnitude of which will correlate with the ability of each species to cross the river (see Chapters 3, 4 and 5). Specifically (i.e. across species), orang-utan populations on either side of the river will be least similar and proboscis monkey populations most similar. Long-tailed macaques are predicted to show intermediate levels of genetic variation due to their occasional swimming and social structure.

In Chapter 2, I describe the field and laboratory methods used during my study, including details of the genetic analyses used in subsequent chapters. In the subsequent Chapters (3, 4 and 5), I describe the population genetic structure and historical demography of each species beginning with long-tailed macaque and then describing the orang-utan and proboscis monkey. In the final Chapter (6), I assimilate all the data from the three different primate species and present a scenario on the events that occurred in Kinabatangan basin that might

have caused the present day levels of genetic diversity and population structure in the orang-

utan, long-tailed macaque and proboscis monkey.

CHAPTER 2: MATERIAL & METHODS

2.1 SAMPLING

Prior to fieldwork, we decided to utilize faeces as a non-invasive source of DNA from each primate species. As the primates (long-tailed macaques and proboscis monkeys) are known to use riverine trees for sleeping sites (van Schaik *et al.,* 1996) and all ten Lots of Kinabatangan border the river this enabled the animals to be observed and counted with ease using a boat survey. Effectively the river was used as a transect to census the primates during their inactive periods. Collection of the orang-utan faecal samples used in the current study (see Chapter 4) has previously been described by Goossens *et al.* (2005).

Sampling for long-tailed macaques and proboscis monkeys was carried out between June and September 2003. During the fieldwork, the whole length of the Lower Kinabatangan Wildlife Sanctuary (LKWS) was divided into five sections (see Fig. 2.1). For each section, census data and sampling was carried out for about two weeks (see Table 2.1).

Figure 2.1: Location of the Lower Kinabatangan Wildlife Sanctuary study sites in the Sabah state of Malaysia. The sanctuary consists of ten Lots, each marked in red. The Kinabatangan River and its tributaries are shown in blue.
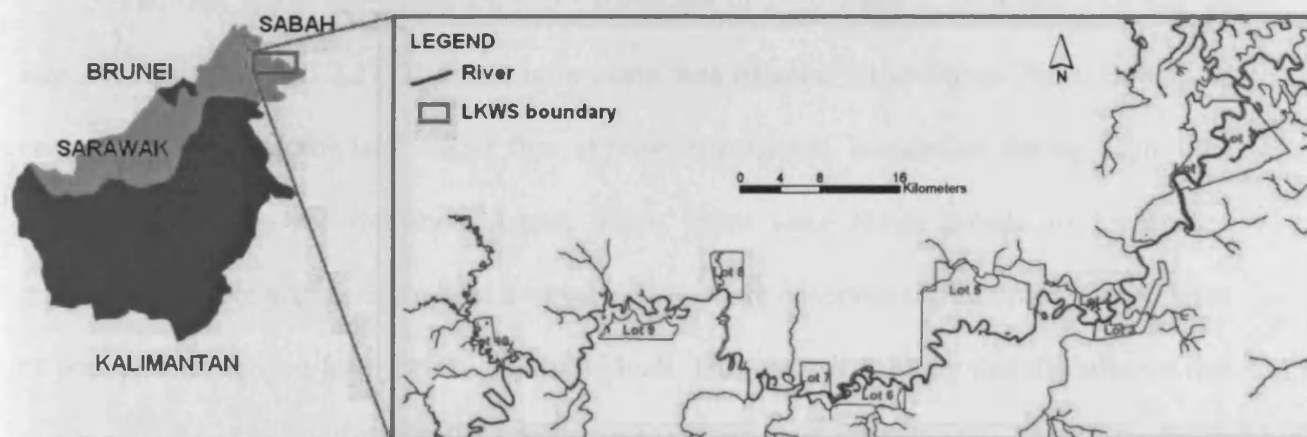
Table 2.1: Fieldwork timetable (dry season June to July and rainy season August to September)

No.	Fieldwork	Months	Areas
1.	Section 1	June 2003	Lots 1 and 2
2.	Section 2	July 2003	Lots 5, 6 and 7
3.	Section 3	August 2003	Lots 3 and 4
4.	Section 4	September 2003	Lots 8, 9 and 10

In total, the surveys covered ~330 km, equalling 660 km of riverbank including smaller tributaries. The surveys were carried out from 0600 to 0900 hours. The survey team consisted of the boat-man and one observer while the boat stayed in the centre of the river with a constant speed of 4 km/h. For each primate sighting, the following information was recorded: date, start and end times of observation, forest quality (i.e. disturbed or heavily disturbed as defined by Ancrenaz *et al.*, 2004), species, location (GPS coordinates), estimated group size and, where possible, group composition. Group size was estimated to the nearest 10 individuals (A <10, B = 11-20, C = 21-30 and D 31-40).

2.1.1 Section 1 (Lots 1 and 2)

The first leg of the fieldwork was carried out in Lots 1 and 2, 33.4 and 37.6 km² in size, respectively (Fig. 2.2). The first base camp was established at Danau Pitas. Both Lots consist of disturbed lowland forest that experiences diurnal inundation during each tide, exacerbated during the monsoon season. Here, there were seven groups of long-tailed macaques and five groups of proboscis monkeys that were observed and sampled. Group sizes of both species ranged from two to ten individuals. However, it is likely that disturbance due to a high level of river traffic (fishing, barge and tourist boats) reduced the number of primate sightings during the census.

Figure 2.2. Lots 1 and 2 of the Lower Kinabatangan Wildlife Sanctuary and the position of the base camps (yellow triangle). Lot boundaries are demarcated by the river or shown by the red dotted lines.
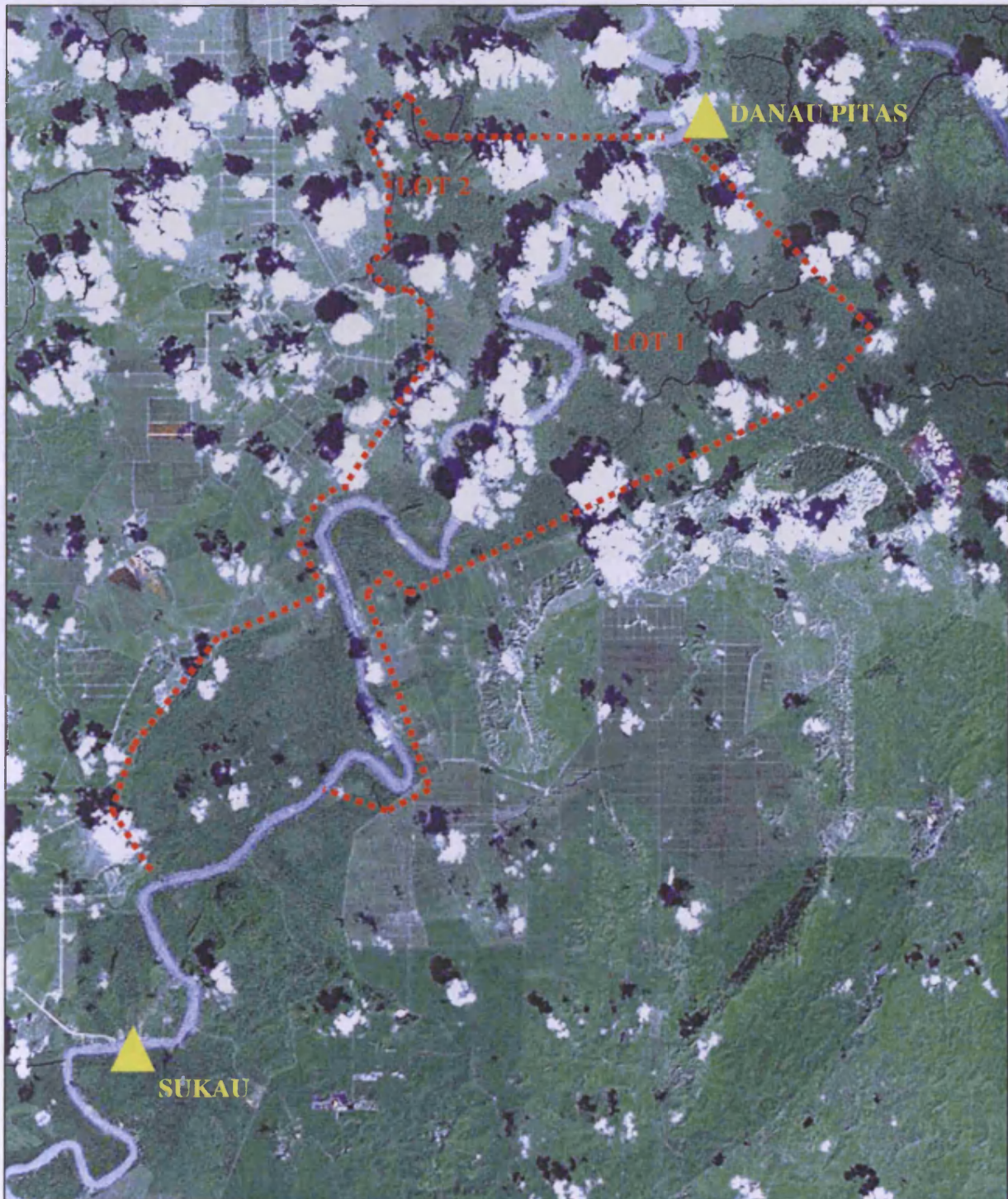
2.1.2 Section 2 (Lots 5, 6 and 7)

Figure 2.3. Lots 5, 6 and 7 of the Lower Kinabatangan Wildlife Sanctuary and the position of the base camp (yellow triangle). Lot boundaries are demarcated by the river or shown by the red dotted lines.
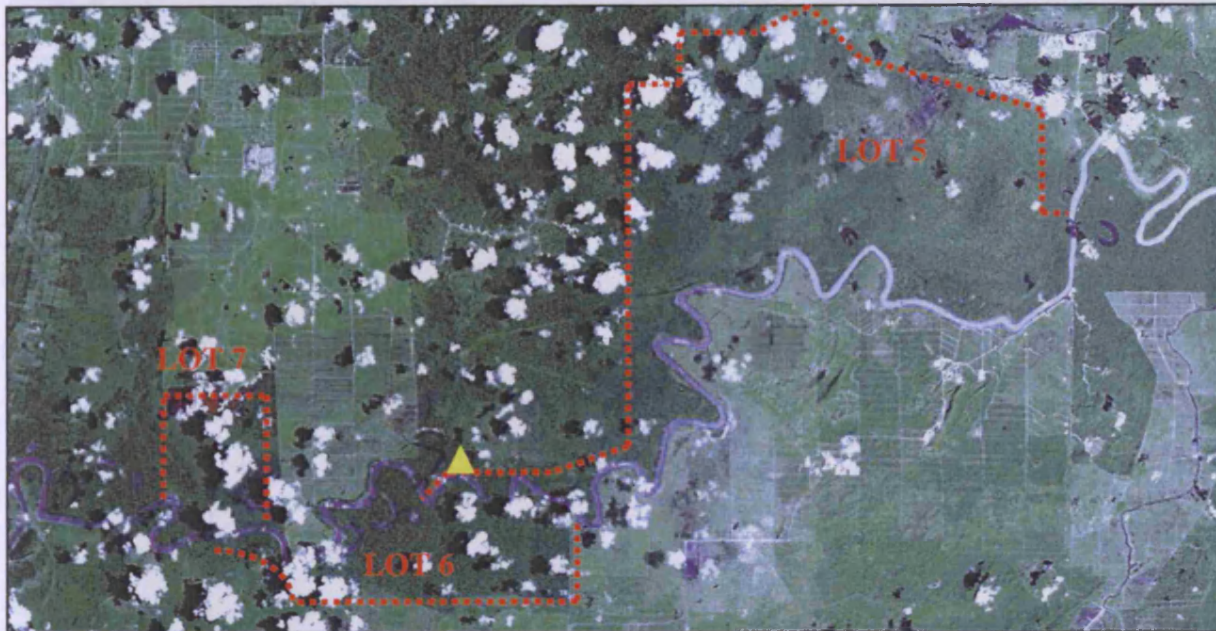
The second part of the field sampling was carried out in Lots 5, 6 and 7 (Fig. 2.3). Both Lots 5 and 6 are classified as disturbed forest, but Lot 7 is classified as heavily disturbed forest. As in Lots 1 and 2, this forest is also occasionally flooded during monsoon rains. There are many ox-bow lakes with large numbers of primates found near the lakes. Habitat disturbance is more pronounced than in Lots 1 and 2. Lot 5 covers the largest area within the sanctuary at 74.2 km^2. Lot 6 is more modest at 26.7 km^2 and Lot 7 at 10.3 km^2. An oil palm plantation separates Lots 5 and 7, Lot 6 is bordered by the Kinabatangan River in the north and oil palm plantation in the south. On the north-western border of Lot 5 lies the Gomantong Forest Reserve. In Lots 5 - 7, 103 long-tailed macaques and 43 proboscis monkey groups were recorded. The sizes of the groups ranged from two to 40 long-tailed macaque individuals and from two to 30 proboscis monkey individuals. Both species were commonly found sleeping in the same trees. In contrast to Lots 1 and 2, the impact of tourist disturbances is less apparent due to the increased distances from the nearest village.

2.1.3 Section 3 (Lots 3 and 4)

Figure 2.4. Lots 3 and 4 of the Lower Kinabatangan Wildlife Sanctuary and the position of the base camp (yellow triangle). Lot boundaries are demarcated by the river or shown by the red dotted lines.

The third sampling was carried out in Lots 3 and 4 (Fig. 2.4). Both are comparatively small in size at 22.1 km^2 and 18.8 km^2 consisting of disturbed forest with seasonal inundation. There were several ox-bow lakes showing different stages of plant colonization. There were 14 and 36 groups of long-tailed macaques and proboscis monkeys observed, respectively. Most of the groups observed were small ranging from two to ten individuals. Rarely were large groups of 20 or more individuals observed. On a number of occasions, the riverbanks were devoid of any primates due to the noise of tourist boats and, as the animals were sensitized to boat disturbances, the number of animals in these Lots was probably underestimated.

2.1.4 Section 4 (Lots 8, 9 and 10)

Figure 2.5. Lots 8, 9 and 10 of the Lower Kinabatangan Wildlife Sanctuary and the position of the base camps (yellow triangle). Lot boundaries are demarcated by the river or shown by the red dotted lines.
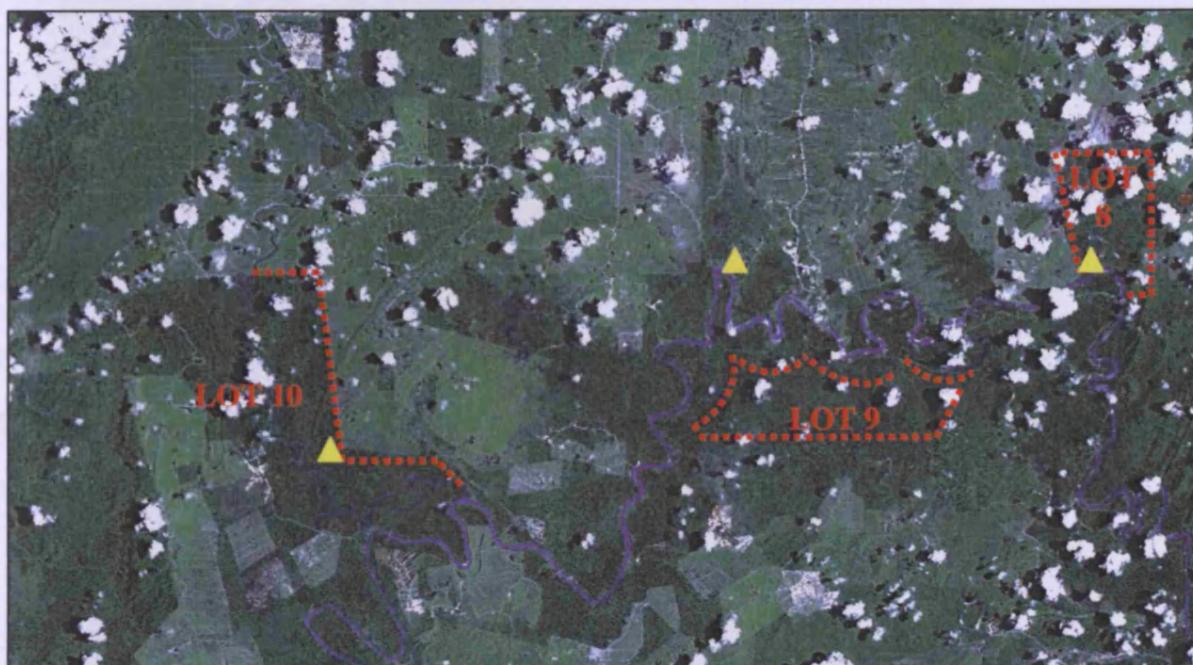
The fourth sampling period was the most difficult due to the start of the monsoon rains. Flooding of the forest greatly limited suitable sites for our base camps. Three base camps were set up due to the large distances between Lots. Both Lots 8 and 9 are heavily disturbed semi-inundated forest. Lot 10 consists of three fragments of mixed heavily disturbed and disturbed forest (Fig. 2.5). The Lot sizes ranged from 11.2 to 12.0 to 28.1 km^2 for Lots 9, 8, and 10, respectively. Only 12 groups of long-tailed macaques and 13 groups of proboscis monkeys were observed. All the group sizes were small ranging from two to 10 individuals, but this may have been caused by the reduced visibility in the rainy season. There was no disturbance from tourists, but occasionally crocodiles were detected basking on the river banks or lying in shallow water, which seemed to greatly reduce primate activity.

Immediately following each morning and evening census survey, the same sites were revisited and faecal samples were collected below trees previously occupied by sleeping or

resting primates. Samples were carefully collected to avoid potential cross contamination with human DNA and stored into 95% ethanol as suggested by Goossens *et al.* (2003). GPS coordinates of samples were noted for confirmation of sample location within each study area. Samples were then stored in a cold room before DNA extraction.

2.2   GENETIC STUDIES

DNA was extracted from all faecal samples using a QIAamp DNA Stool Mini Kit (QIAGEN, GMBH Cat. #51504) according to the manufacturer's protocol with the following modification: at the final step, the amount of AE elution buffer was reduced from 200 μl to 150 μl (see part 2.2.1 for further details). The extracted DNA samples were later stored at 4°C.

2.2.1   Faecal DNA extraction protocol

DNA was extracted from faecal samples using the QIAGEN QIAamp DNA Stool Mini Kit (QIAGEN, GMBH Cat. #51504). For each sample, five sets of Eppendorf tubes were prepared and labelled A-E (A and B were 1.8 ml tubes and C-E 1.5 ml tubes), in addition to one set of QIAamp spin columns and two collection tubes for each sample marked X and Y. In tube A, 1.4ml of Buffer ASL was added; in tube B, one Inhibitex tablet was added; and in tube D, 15 μl of Proteinase K was added.

The surface of the faecal samples were scraped with a clean sterile blade and this together with an inner portion of the pellet was added to tube A (150-200 mg) and mixed using a clean sterile spatula. The mixture was vortexed for 1 min and then heated at 70°C for five min. The mixture was then vortex for 15 s and centrifuged at 13,000 g for one min. The supernatant (1.2 ml) was transferred from Tubes A to B, and Tube A was discarded. Tube B was vortexed for one min, incubated at room temperature for one min, re-centrifuged for 3

min, before the supernatant was transferred to Tube C. Tube B was discarded. Tube C was centrifuged for 3 min and then 200 ml of supernatant transferred to Tube D. To this was added 200ml of buffer AL before vortexing for 15 s. After 10 min incubation at 70°C, 200 ml of ethanol was added to the lysate and vortexed for 15 min.

The lysate was transferred from Tube D into a new QIAamp spin column and Tube D was discarded. The QIAamp spin column was centrifuged at 13,000 g for one min. and then the spin column was transferred into a new collection tube (Tube X) and the tube containing the filtrate was discarded. In the same QIAamp spin column, 500μl AW1 buffer was added and centrifuged for one min. The spin column was then transferred into a new collection tube (Tube Y) and tube X was discarded. In the same QIAamp spin column, 500μl AW2 buffer was added and then centrifuged at 13K for three min. The same spin column was then transferred into a new eppendorf tube (Tube E). 150μl AE buffer was added into the QIAamp spin column and incubated for one min at room temperature. The column was centrifuged at 8,000 g for one min. Finally, the spin column was discarded and the resulting DNA extraction (Tube E) was stored at 4°C prior to use.

2.2.2   Hair extraction protocol

Human DNA could easily cause contamination when studying non-human primates causing false positive in PCR. To control for cross-species amplification, DNA from human hair was extracted from the author as a control for each PCR carried out in the study. The DNA was extracted from hair follicles using a buffer based protocol of Engström *et al.* (1998) and Vigilant (1999).

Hair samples were washed several times in ultrapure water. Using a clean, sterile blade, hair shafts were removed and the remaining hair and root were transferred into a 1.5ml

screw capped eppendorf tube containing 20µl 10X PCR buffer, 1µl Proteinase K (50µg/µl) and 79µl ultrapure water. The mixture was agitated at 37°C overnight.

. The next day, the mixture was boiled for 10 min and centrifuged at 13K for one min. The resulting DNA extraction was stored at 4°C prior to use. For subsequent PCRs, there was no need to add extra buffer.

2.2.3   Standard PCR protocol

PCR reactions were performed in a final volume of 20μl, containing 2μl DNA extract, 1.5μl 4 mg/ml BSA, 2μl 10X PCR Buffer, 1.5μl 25 mM MgCl$_2$, 1μl 10mM of dNTP mix, 0.2μl of 50 pmol of each primer, 0.2μl of AmpliTaq Gold^{TM} and 11.4μl of ultrapure water. Amplification was carried out in a Perkin Elmer 9700 thermocycler following an initial denaturation for 12 min at 94°C followed by 40 cycles of 94°C denaturation for 40 s, annealing at 58°C for 30 s, one min of extension at 72°C and final extension of 10 min at 72°C. The annealing temperature ranged between species, 58°C, 60°C and 62°C for long-tailed macaque, orang-utan and proboscis monkey, respectively.

Amplification of the mtDNA control region was achieved using the general primers (L16517 and 12SAR-3', Fumagalli et al., 1996; Palumbi, 1996) and species-specific primers (see Table 2.2). For this study, three sets of species-specific primers were developed that successfully amplified only the target species. These primers were tested against other target species (primate vs. primate) and humans (human vs. primate), but only amplified the target species.

Table 2.2 Species-specific PCR primers developed and used for this study.

No.	Species	Primer code	Primer Sequence
1.	Long-tailed macaque	*Mf-5'*	5'-GCA ACT ACT TTC TGC ACT-3'
2.	Long-tailed macaque	*Mf-3'*	5'-GAA CAA GGG ATT CCT AAG-3'
3.	Orang-utan	*Pp-5'*	5'-GCA CTT AAC TTC ACC ATC-3'
4.	Orang-utan	*Pp-3'*	5'-AAA CAA GGG ACC ACT AAC-3'
5.	Proboscis monkey	*Nl-5'*	5'-CGT AAA CCA GAA ACG GAT-3'
6.	Proboscis monkey	*Nl-3'*	5'-TAA TGG GAA TAT CCG TGC-3'

2.2.4 DNA sequencing

PCR products were purified using Exonuclease I/Shrimp Alkaline Phosphatase (ExoSap) (USB Corp, USA) (Hanke & Wink, 1994). An 0.5μl aliquot of ExoSap mastermix (Exonuclease I (10 units/μl) and Shrimp Alkaline Phosphatase (1 unit/μl) in a ratio of 1:1) was added to each 5μl PCR product. Purification of PCR products was carried out following activation of the enzymes at 37°C for 60 min and deactivation at 80°C for 15 min.

Sequencing PCRs were performed in a final volume of 8μl, containing 2μl of purified PCR product, 2.5μl Better Buffer (Web Scientific), 0.5μl BigDye Terminator Ver.1 (Applied Biosystems), 1μl of 1.6 pmol of primer and 2μl of water. Sequencing PCR was performed separately for forward and reverse primers. Sequencing PCR was carried out following an initial denaturation for 3 min at 96°C followed by 25 cycles of 96°C denaturation for 15 s, annealing at 50°C for 10 s and 2 min of extension at 60°C. Sequencing PCR was performed using ABI Big Dye Terminator vs. 1 (Applied Biosystems). Each PCR product was sequenced in both directions.

PCR products were precipitated by adding 90μl of 63% isopropanol to each PCR tube. Tubes were vortexed for 20 s, left to stand for 15 min, then centrifuged for 30 min at 13,000 g. The supernatant was discarded and 150μl of 70% isopropanol was added to each PCR tube. The PCR strips were then centrifuge for one min at 500 g and dried at 52°C for 2 min. Sequencing was performed in an ABI3100 automated sequencer.

2.3 DATA ANALYSIS

2.3.1 Sequence alignment and editing

Sequences obtained were checked by eye and contigs created from each forward and reverse sequence in SEQUENCHER 3.1.2 (Genecode Corp.). Each contig was checked again for any base discrepancies and corrected by eye. The contig was then converted into a consensus sequence and all the consensus sequences were aligned into one file consisting of all individual samples for each species. This file was then exported in NEXUS format.

2.3.2 Genetic Diversity

For this study, two basic measures of genetic diversity were used to describe the data, nucleotide and haplotype diversity, both calculated in ARLEQUIN version 3 (Excoffier *et al.*, 2005). Nucleotide diversity (π) is a measure of DNA sequence polymorphism within a population, and is defined as the average number of nucleotide differences per site between two sequences (Nei, 1987). Haplotype diversity (h) is defined as the probability that two randomly chosen haplotypes are different (Nei, 1987), where n is the number of gene copies in the sample, k is the number of haplotypes, and Pi is the sample frequency of the i-th haplotype. It is equivalent to the expected heterozygosity for diploid data.

Haplotype and nucleotide diversities are typically interpreted according to Grant & Bowen (1998) who analysed mtDNA of marine fish as a basis of assessing demographic

histories (Table 2.3). Low values found for both indices indicate a recent bottleneck, whereas high values indicate a large stable population with a long evolutionary history or secondary contact of differentiated lineages. High haplotype and low nucleotide diversities suggest that there has been rapid growth from a small ancestral population but in too short a time to accumulate mutations. Low haplotype and high haplotype diversities may be the result of transient bottlenecks in large ancestral populations since short crashes eliminate many haplotypes but do not affect nucleotide diversity severely.

Table 2.3. Grant & Bowen' (1998) interpretation of differences between haplotype and nucleotide diversities.

		Haplotype diversity	
		Low	**High**
Nucleotide diversity	**Low**	Recent population bottleneck. Founder effect with single or few lineages.	Bottleneck followed by rapid population growth and mutation accumulation.
	High	Divergence between geographically subdivided populations.	Large stable population with long evolutionary history. Secondary contact between differentiated lineages

2.3.3 Analysis of molecular variance (AMOVA)

Hierarchical population genetic structure was investigated using Analysis of Molecular Variance (AMOVA) as implemented in ARLEQUIN version 3 (Excoffier *et al.*, 1992, 2005). This statistical method was initially defined by Cockerham (1969, 1973) and extended by others (Weir & Cockerham, 1984; Long, 1986). The AMOVA approach is based on analysis of variance of gene frequencies which can take into account the number of mutations between molecular haplotypes (Excoffier *et al.*, 2005). For this study, populations were defined according to their geographic locations and to investigate the possibility that the Kinabatangan River impedes gene-flow, primate populations on the north and south riversides were compared.

The fixation index (Φ) ranges from 0 (indicating no differentiation between the overall population and its subpopulations) to a theoretical maximum of 1. Occasionally AMOVA produces small negative values of Φ_{ST} which indicates that haplotypes drawn at random from one population have a higher probability of being identical to haplotypes of another population rather than the population of origin (Excoffier *et al.*, 1992). Fixation indexes (Φ) can be determined for differentiated hierarchical levels of a population structure, for example, Φ_{CT} describes the regional apportionment of genetic variation with respect to all haplotypes, Φ_{SC} describes the apportionment of variation within the population of a given region, and Φ_{ST} characterizes the variation between haplotypes in a single population relative to all haplotypes (Excoffier *et al.*, 1992).

2.3.4 Networks

In this study, gene genealogies were investigated using networks rather than bifurcating evolutionary trees, which assume no reticulation (horizontal gene-flow - an assumption violated in many population studies). Intraspecific genealogies can be affected by low divergence among individuals, persistence of ancestral haplotypes, reticulation and large sample sizes (Posada & Crandall, 2001).

Two different methods were used to create networks: (i) minimum spanning networks (MSN) implemented in ARLEQUIN version 3 (Excoffier *et al.*, 2005) and (ii) median joining network (MJN) using NETWORK version 4.1.1.1. (Bandelt *et al.*, 1999). The minimum spanning network includes all equally parsimonious minimum spanning trees (MST) into a single network. MST construction is based on a matrix of pairwise distances among haplotypes. If there is more than one solution present, alternative links are created in the network (Excoffier & Smouse, 1994). In the median joining approach, all MSTs are first combined within a single network (MSN) following an algorithm analogous to Excoffier &

Smouse (1994). Using a parsimony criterion, median vectors (which represent missing intermediates) are added to the network (Bandelt *et al.*, 1999; Posada & Crandall, 2001).

2.3.5 Mismatch distributions

Population demography was analysed using mismatch distributions. A mismatch distribution is the distribution of the observed number of differences between pairs of haplotypes. Samples drawn from a population at demographic equilibrium that exhibit highly stochastic gene tree topologies usually have a multimodal distribution. In contrast, populations having passed through a recent demographic bottleneck and expansion (Slatkin & Hudson, 1991; Rogers & Harpending, 1992) or though a range of expansions with high levels of migration between neighbouring demes (Ray *et al.* 2003; Excoffier 2004) tend to be unimodal. Mismatch distribution can be calculated using three different models: pure demographic expansion, sudden expansion and spatial expansion model. The first two models are very similar, and therefore in the current study just the pure demographic expansion and spatial expansion models were calculated in ARLEQUIN 3 (Excoffier *et al.*, 2005).

A demic spatial expansion generally occurs if the range of a population is initially restricted to a very small area, and then the range of the population increases over time and space. In the demic spatial expansion model, the shape of the gene genealogies and the overall pattern of diversity within demes are influenced by the age of the expansion and the level of gene-flow, Nm (N=size of deme; m= proportion of migrants) between neighbouring demes (Ray *et al.*, 2003). Low gene-flow (<1 migrant per generation) produces a substantial proportion of coalescent events early in the genealogy producing gene genealogies that are star-shaped and multimodal mismatch distributions. For large Nm values, most coalescent events that occur around the time of the onset of the spatial expansion produce a mixture of both very short and long branch lengths gene genealogies resulting in unimodal mismatch

distributions. Tests of selective neutrality (Tajima's D or Fu's F_S) will show significant negative values after a spatial expansion only in demes with high Nm values

2.3.6 Selective neutrality

Statistics based on mismatch distribution are not always robust at detecting expansion and critically depend on sample size, therefore, a range of tests were employed to detect traces of past population growth or stability based on DNA sequences (Ramos-Onsins & Rozas, 2002). Tajima's D (Tajima, 1989), Fu's Fs (Fu, 1997), Fu & Li, D* and F* (Fu & Li, 1993) tests of neutrality were also used to assess demographic history. All estimates are based on the infinite-site model (ISM) without recombination, appropriate for mtDNA. Fu & Li's (1993) D* and F* statistics together with Fu's (1997) F_S can be used to distinguished the effects of background selection from population growth or range expansion. If F_S is significant and F* and D* are not, this indicates population growth or range expansion, whereas, the reverse indicates selection (Fu, 1997). All selective neutrality tests were performed in ARLEQUIN 3 (Excoffier et al., 2005) with the exception of Fu & Li's (1993) D* and F*, which were calculated in DNASP 4.10.3 (Rozas et al., 2003).

2.3.7 Coalescent simulation

Coalescent-based methods were used to test for evidence of population expansion using FLUCTUATE v.1.4. (Kuhner et al., 1998). This approach uses a maximum likelihood method to simultaneously estimate theta (θ_{Kuhner}; the scaled mutation rate theta (Θ) = $N\mu$ where N is the size of each subpopulation and μ is the mutation rate, assuming an infinite allele mutation model) and population growth rate (g). FLUCTUATE assumes that the loci sampled are not affected by selection or recombination. If these assumptions are violated the results may be erroneous. To achieve convergence for the estimates, the simulation was executed five times. Large positive values of theta and g indicate population growth.

CHAPTER 3

MITOCHONDRIAL GENETIC STRUCTURE OF THE LONG-TAILED MACAQUE (*Macaca fascicularis*) IN THE KINABATANGAN FLOODPLAIN, SABAH, MALAYSIA

Jalil, M.F.[1,2], Goossens, B.[1,2], Bruford, M.W.[1], Sinyor, J.[3] and Cable, J.[1]

[1] School of Biosciences, Cardiff University, Cardiff CF10 3TL, UK.

[2] Institut Biologi Tropika & Pemuliharaan, Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah, Malaysia.

[3] Kinabatangan Orangutan Conservation Project, P.O.Box 3109, 90734 Sandakan, Sabah, Malaysia.

Abstract

To investigate the population genetic and hence demographic structure of the long-tailed macaque (*Macaca fascicularis*) in the Lower Kinabatangan Wildlife Sanctuary (Sabah), sequence variation in the mitochondrial DNA control region was analyzed from 89 faecal samples taken from 88 sites north (n = 31) and south (n = 57) of the Kinabatangan River. As predicted, this widely distributed, invasive species showed both high haplotype and nucleotide diversity, indicating a historically large and stable population, an inference also supported by a multimodal mismatch distribution. Surprisingly and despite the fact that long-tailed macaques are known to be good swimmers and therefore able to cross rivers, analysis of molecular variance indicated restricted gene-flow between populations on the north and south side of the Kinabatangan River. However, 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.

3.1   INTRODUCTION

Riverine biodiversity is threatened around the world by habitat degradation which includes pollution, deforestation, creation of dams and over-harvesting (see Dudgeon, 2000). Riverine species also face severe competition with humans because riverine wetlands are also

productive agricultural lands. This is particularly important in Asia where rapidly expanding human settlements inexorably encroach upon riverine habitats (Dudgeon, 2000). Many river basins in South East Asia are crowded with growing human settlements (e.g. the Mekong and Chow Phraya Basins). River pollution has reduced many populations of freshwater flora (i.e. *Cyrpticoryne*) and fauna (i.e. *Pangasianodon gigas*, see Na-Nakorn *et al.*, 2006), and a combination of damming and water-flow regulation have disrupted migration patterns of freshwater fauna (e.g. the Chinese sturgeon, Dudgeon, 2000). One main impact of riverine pollution is the displacement of many wild species with domesticated forms. Many large bodied species that are unable to adapt to changing habitats (i.e. Yangtze river dolphin, Indian rhino and wild water buffalo) are almost extinct (Dudgeon, 2000; Yang *et al.*, 2006) due to habitat loss. Associated habitat fragmentation has also resulted in the restriction of many species to small habitat patches separated by a matrix of inhospitable environments (such as plantations). All such changes potentially have an adverse impact on genetic diversity, including reduction in local population size, reduced migration (i.e. Chinese Paddlefish, Dudgeon, 2000) and inbreeding (Frankham *et al.*, 2002; Frankham, 2005). Short term effects of habitat fragmentation and isolation are influenced by effective population size and genetic drift, which are controlled by factors such as availability of mates (sex ratio) and food (productivity), predation pressure and parasite loads. Long term consequences may include genetic loss, deleterious mutation accumulation and fixation of certain haplotypes.

Asia is unique in having many terrestrial mammals associated with riverine wetlands (Dudgeon, 2000), some of which are now classified as vulnerable (e.g. fishing cats) or endangered (e.g. orang-utan, proboscis monkey, IUCN, 2006) due to habitat loss and degradation. In addition to the charismatic megafauna that can be found in riverine wetlands (such as Malayan Tapir and Asian elephant), some common and commensal species are also present, including the long-tailed macaque (*Macaca fascicularis,* see van Schaik *et al.*, 1996;

Dudgeon, 2000). This species occurs from Burma and Indochina in the north, to Bali in the east, and Sumatra and Nicobar in the extreme west. They are extremely common throughout their native range (Abegg & Thierry, 2002; Umapathy *et al.*, 2003). In fact, the long-tailed macaque has 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 (Lowe *et al.*, 2000; Long, 2003). Long-tailed macaques inhabit 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). Their only natural predators are the clouded leopard and crocodile, but these predators occur in low densities with no obvious impact on long-tail macaque populations. Long-tailed macaques required large home ranges depending on the quality of the forest, 25-50 ha in primary forest and up to 200 ha in secondary or degraded forest (Wolfheim, 1983; de Ruiter & Geffen, 1998).

Using mitochondrial DNA (mtDNA) sequences, Melnick and Hoelzer (1996) and Cowlishaw and Dunbar (2000) showed that primates with female philopatry (such as macaques) have greater local genetic homogeneity and greater among–population genetic differentiation than species which conform more closely to panmixis. The utility of mtDNA is widely recognised for inferring genetic relationships among and within populations, especially as it is easily isolated and occurs in multiple copies per.cell. The control region of the mtDNA is particularly useful due to its rapid rate of sequence evolution in humans (Vigilant *et al.*, 1989, 1991), other primates (Marmi *et al.*, 2004) and other vertebrates such as birds (Wennick *et al.*, 1994; Delport *et al.*, 2002). Traditionally, the left domain of the control region has been favoured due to its rapid rate of evolution and lack of conserved blocks (Fumagalli *et al.*, 1996). Many recent papers have used the left domain to study phylogeography of primates across their geographic ranges showing high levels of

polymorphism (i.e. gorillas, Clifford *et al.*, 2004; orang-utans, Warren *et al.*, 2001; barbary macaques, Modolo *et al.*, 2005). However, to date there have been no empirical studies that have directly compared polymorphism in the left and right domain of the control region in non-human primates and very few examining polymorphism within and among social groups and populations on a small spatial scale. This study both compares levels of polymorphism in the left and right domain of the mtDNA control region, prior to investigating the fine-scale population genetic structure of the long-tail macaque. Specifically, the genetic diversity within and between populations of long-tailed macaques in the fragmented forest of Lower Kinabatangan is assessed in the context of their highly structured social systems along the Kinabatangan River.

3.2 MATERIAL AND METHODS

DNA samples

Samples were collected along the Kinabatangan River and Pangi Forest Reserve (Pangi) within the Lower Kinabatangan Wildlife Sanctuary, Sabah, Malaysia (LKWS) in 2003 (Fig. 3.1 and Appendix 1). Long-tailed macaques were surveyed on each Lot of the LKWS by boat at a constant speed following methods described by Goossens *et al.* (2003). All faecal samples were collected under trees following earlier sightings of long-tail macaques, therefore individual samples could not always be related back to defined social groups. Faecal samples (n = 88 + one from Labuk Bay outside of LKWS) were collected in sterile falcon tubes and preserved in 95% ethanol at 4°C. Total DNA was extracted using a QIAamp DNA Stool Mini Kit following the manufacturer's protocol (QIAGEN GMBH Cat. #51504). However, at the final step, the amount of AE elution buffer was reduced from 200μl to 70μl.

Figure 3.1. Distribution of long-tailed macaque faecal samples (purple squares) collected along the Kinabatangan River. Red line indicates the Lower Kinabatangan Wildlife Sanctuary (LKWS) boundaries for each Lot and blue lines indicate the Kinabatangan River and its tributaries.
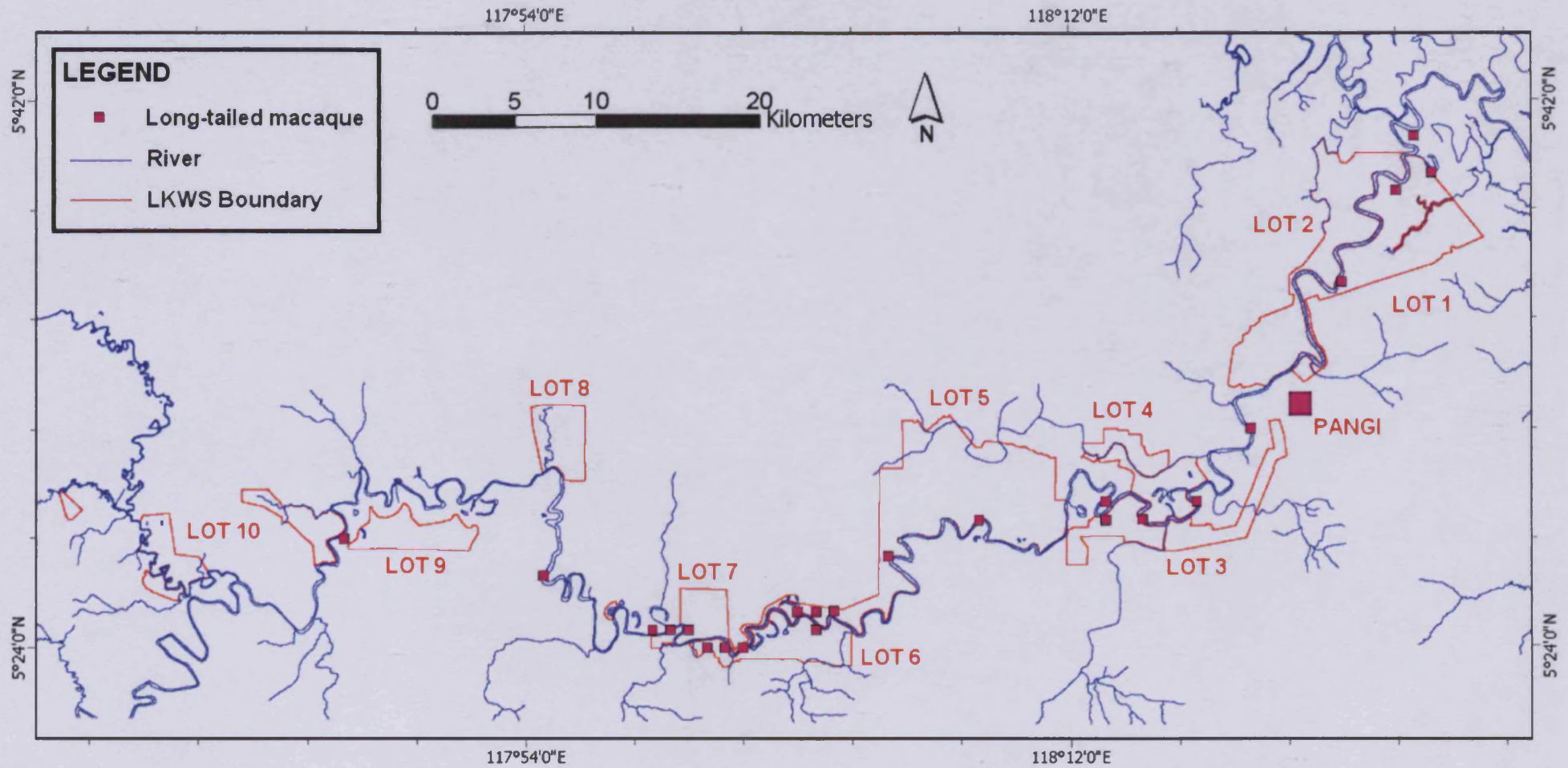
The left (350 bp) and right domains (545 bp) of the mitochondrial DNA control region were amplified using primers Mf-5'/Mf-3' (see Chapter 2) and L16517/12SAR-3', respectively (Fumagalli et al., 1996; Palumbi, 1996) (Fig. 3.2). PCR reactions were performed in a final volume of 20μl, containing 2μl DNA extract, 1.5μl 4 mg/ml BSA, 2μl 10X PCR Buffer, 1.5μl 25mM MgCl$_2$, 1μl 10mM of dNTP mix, 0.2μl of 50pmol of each primers, 0.2μl of AmpliTaq Gold^{TM} (Applied Biosystems) and 11.4μl of ultrapure water. Amplifications were performed in a Perkin Elmer 9700 thermocycler following an initial denaturation for 12 min at 94°C followed by 40 cycles of 94°C for 40 s, 50°C for 30 s and one min of extension at 72°C, and a final extension of 10 min at 72°C.

PCR products were purified using Exonuclease I/Shrimp alkaline Phosphatase (ExoSap) (Hanke & Wink, 1994). ExoSap mastermix were prepared by adding Exonuclease I (10 units/μl) (USB Corp. USA) and Shrimp Alkaline Phosphatase (1 unit/μl) (USB Corp. USA) in a ratio of 1:1. 0.5μl of the mastermix was added into each 5μl PCR product aliquots (in a new PCR strips). Purification of the PCR products was carried out following an activation of the enzymes at 37°C for 60 min and deactivation at 80°C for 15 min.

Sequencing PCR was performed in a final volume of 8μl, containing 2μl purified PCR product, 2.5μl Better Buffer (Web Scientific ltd. UK), 0.5μl BigDye Terminator Ver.1 (Applied Biosystems, Europe), 1μl of 1.6 pmol of primer and 2μl of water. Sequencing was performed using both forward and reverse primers. Sequencing PCR was carried out following an initial denaturation for 3 min at 96°C followed by 25 cycles of 96°C denaturation for 15 s, annealing at 50°C for 10 s and 2 min of extension at 60°C. PCR products were precipitated by adding 90 μl of 63% isopropanol to each PCR tube, vortexing for 20 s, standing for 15 min, and then centrifuging for 30 min at 13,000 g. The supernatant was discarded and 150 μl of 70% isopropanol was added to each PCR tube. The PCR strips were

Figure 3.2: Schematic diagram of the organization of the mitochondrial DNA control region in mammals (from Fumagalli *et al.*, 1996).

Left Domain   Central Conserved Domain   Right Domain

Mf-5'   L16517   CSB1 CSB2 CSB3

Pro-tRNA   Phe-tRNA

Mf-3'   12SAR-3'
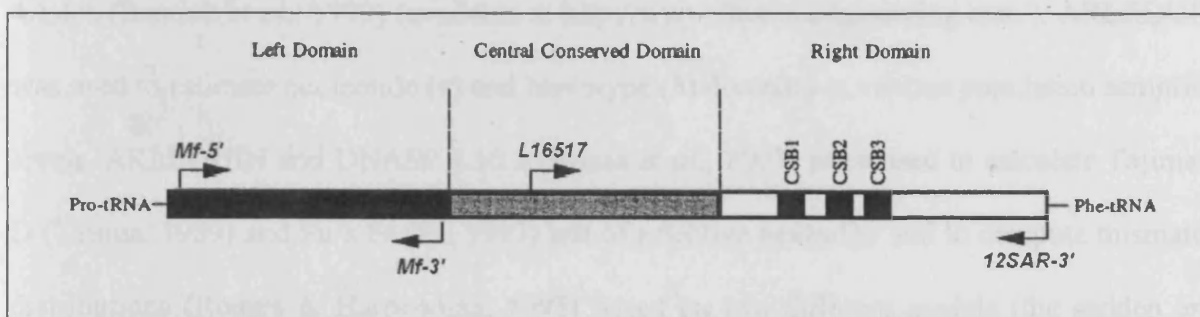
then spun for 1 min at 500 g and dried at 52°C for 2 min. Each PCR product was sequenced in

both directions. Sequencing was performed in an ABI3100 semi automated sequencer.

Genetic structure of control region haplotypes

MtDNA control region sequences were aligned using the program SEQUENCHER

3.1.2 (GeneCodes Corp.) with correction by eye. Minimum spanning and median joining

networks were estimated using ARLEQUIN 3 (Excoffier *et al.*, 2005) and NETWORK

4.1.1.1 (Bandelt *et al.*, 1999) (available at http://www.fluxus-engineering.com/). ARLEQUIN

was used to estimate nucleotide (π) and haplotype (*h*) diversity at various population sampling

levels. ARLEQUIN and DNASP 4.10.3 (Rozas *et al.*, 2003) were used to calculate Tajima's

D (Tajima, 1989) and Fu's Fs (Fu, 1997) test of selective neutrality and to compute mismatch

distributions (Rogers & Harpending, 1992) based on two different models (the sudden and

spatial expansion models) for all LKWS long-tailed macaque samples and separately for each

riverside as implemented in ARLEQUIN. A hypothesis of exponential growth using a

Markov Chain Monte Carlo (MCMC) approach was also tested as implemented in

FLUCTUATE v. 1.4. (Kuhner *et al.*, 1998). ARLEQUIN was also used to estimate pairwise

population differentiation (Φ_{ST} values) by Analysis of Molecular Variance (AMOVA)

(Excoffier *et al.*, 1992). The significance of Φ_{ST} was assessed through a permutation test with

1,000 repetitions.

3.3   RESULTS

I. Left Domain of mtDNA Control Region

From 350 bp of the left domain (LD) of the CR sequenced from 89 samples, there were 43

polymorphic sites with 42 transitions and 1 transversion (Table 3.5). Twenty two haplotypes

were identified, 21 in LKWS and one in Labuk Bay (Tables 3.1 & 3.2). Of the 21 haplotypes

in LKWS, 11 were identified in multiple sites and ten were found only on the southern side of

Table 3.1. Variable nucleotide positions in the alignment of mtDNA sequences of the long-tailed macaque control region. Dots indicate positions where the bases are identical to haplotype LD01.

	14	32	73	106	116	122	123	147	151	162	168	176	177	181	185	186	187	197	202	203	204	205	206	208	209	210	217	219	220	224	228	230	233	238	239	242	248	254	255	256	283	285	310
LD01	A	C	A	T	C	C	C	T	T	C	A	G	A	A	T	C	T	C	G	C	A	T	G	C	T	C	A	T	T	C	C	T	C	A	A	T	G	G	C	T	A	T	A
LD02	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	T	.	.	C	.
LD03	.	T	.	C	.	T	.	C	C	.	G	.	.	.	.	.	.	.	A	T	.	C	.	T	C	.	.	.	C	T	.	.	.	G	.	C	A	.	T	C	.	C	.
LD04	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
LD05	.	T	.	C	.	T	.	C	C	.	G	.	.	.	.	.	.	.	A	T	.	C	.	T	C	.	.	.	.	T	.	.	.	G	.	C	A	.	T	C	G	C	.
LD06	.	T	.	C	.	T	.	C	C	.	G	.	.	.	.	.	.	.	A	T	.	C	.	T	C	.	.	.	.	T	.	.	.	G	.	C	A	.	T	C	.	C	.
LD07	.	.	.	C	.	T	T	.	C	.	.	A	.	.	C	T	C	A	A	T	.	.	.	.	C	T	.	C	.	T	.	C	T	.	.	.	A	.	T	C	.	C	.
LD08	.	.	.	C	.	T	.	C	C	.	.	A	.	.	.	.	.	.	A	T	.	.	.	.	C	.	.	.	.	T	.	.	T	.	.	.	A	.	T	C	G	C	.
LD09	.	T	.	C	.	T	.	C	C	.	.	A	.	.	.	.	.	.	A	T	.	C	A	T	C	.	.	.	.	T	.	.	.	G	.	C	A	.	T	C	.	C	.
LD10	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
LD11	.	T	.	C	.	T	.	.	C	.	G	.	.	.	.	.	.	.	A	T	G	C	.	T	C	.	.	.	.	T	T	C	T	G	.	C	A	A	T	C	G	C	G
LD12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
LD13	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	A	.	.	.	C	.
LD14	.	.	.	C	T	T	.	C	C	.	.	.	.	G	.	.	T	.	A	T	.	C	.	T	C	.	.	.	.	T	T	C	.	.	G	.	A	.	T	C	G	C	.
LD15	.	T	.	C	.	T	.	C	C	.	.	.	.	.	.	.	.	.	A	.	.	C	.	T	C	.	.	.	.	T	.	C	T	G	.	C	A	.	T	C	.	C	.
LD16	.	T	.	C	.	T	.	.	C	.	G	.	.	.	.	.	.	.	A	T	G	C	.	T	C	.	.	.	.	T	T	C	T	G	.	C	A	A	T	C	G	C	.
LD17	.	.	.	.	.	T	.	.	.	T	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	C	G
LD18	.	T	.	C	.	T	.	C	C	.	.	.	.	.	.	.	.	.	A	.	.	C	.	T	C	.	.	.	.	T	.	.	T	.	.	C	A	.	T	C	.	C	.
LD19	.	T	.	C	.	T	.	C	C	.	G	A	.	.	.	.	.	.	A	T	.	C	.	T	C	.	G	C	.	T	.	.	T	.	.	C	A	.	T	C	.	C	.
LD20	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
LD21	.	T	.	C	.	T	.	C	C	.	G	.	.	.	.	.	.	.	A	T	.	C	.	T	C	.	.	.	.	T	.	C	T	G	.	C	A	.	T	C	.	C	.
LABUK	.	T	.	C	.	T	.	C	C	.	G	.	.	.	.	.	.	.	A	T	.	C	.	.	C	.	.	.	.	T	.	.	T	.	.	C	A	.	T	C	.	C	.

Chapter 3

Table 3.2. Haplotypes of left domain (LD) of mtDNA control region observed for long-tailed macaques and their frequency in each Lot and forest reserve.

HAPLOTYPE	NORTHERN			SOUTHERN					LABUK
	LOT 4	LOT 5	LOT 7	LOT 1	LOT 3	LOT 6	LOT 9	PANGI	
LD01	-	-	-	1	-	-	1	-	-
LD02	-	-	-	-	-	2	-	-	-
LD03	-	2	1	-	-	2	-	-	-
LD04	-	-	-	-	-	1	-	-	-
LD05	-	-	-	-	-	-	-	4	-
LD06	1	-	-	1	-	-	-	2	-
LD07	-	-	-	-	-	-	-	1	-
LD08	-	-	-	-	1	-	-	-	-
LD09	-	-	-	-	1	-	-	.·	-
LD10	-	-	-	-	10	1	-	-	-
LD11	-	-	-	-	-	1	-	-	-
LD12	-	-	-	2	-	-	-	7	-
LD13	-	-	-	-	-	-	2	-	-
LD14	1	2	-	-	-	2	-	3	-
LD15	-	1	1	-	1	-	-	-	-
LD16	-	9	3	-	-	2	-	1	-
LD17	-	1	-	-	-	1	-	-	-
LD18	-	3	1	-	-	1	-	-	-
LD19	1	-	4	-	1	-	-	-	-
LD20	-	-	-	2	-	-	-	-	-
LD21	-	-	-	-	-	-	3	-	-
LABUK	-	-	-	-	-	-	-	-	1
TOTAL	3	18	10	6	14	13	6	18	1

the river (Lot 1 – LD20; Lot 3 – LD08, LD09; Lot 6 – LD02, LD04, LD11; Lot 9 – LD13, LD21 and Pangi – LD05, LD07).

Both minimum spanning and median joining networks revealed groupings of closely related haplotypes connected by long mutational steps. Both networks also suggested partial separation of populations between the Northern and Southern riversides with the Labuk haplotype showing closest affinities to the 'Northern' sequences (Fig. 3.3A & B). Individual haplotypes from Pangi were observed in both groups (Fig. 3.3). Most of the haplotypes occurring in more than one Lot were found in adjoining Lots or forest reserves (i.e. LD16 was found in Lots 5, 6 and 7, and Pangi).

Haplotype diversity (h) ranged from 0.809 ± 0.053 to 0.924 ± 0.019 to 0.926 ± 0.012 for Northern, Southern groupings and all LKWS, respectively. Nucleotide diversity (π) varied from $0.022 + 0.012$ to $0.034 + 0.017$ to $0.035 + 0.018$ for Northern, Southern groupings and all LKWS, respectively (Table 3.5). Both indexes revealed higher values in the Southern compared to the Northern grouping.

Tajima's test of selective neutrality, D, revealed values ranging from -0.131 to 0.936 to 1.423 for Northern, Southern grouping and all LKWS, respectively. Fu's Fs ranged from 0.385 to 2.676 to 4.365 for the Southern grouping, all LKWS and Northern grouping, respectively. However, none of the population expansion estimates were significant at the 95% level (Table 3.5). Mismatch distributions provided multimodal patterns for all LKWS samples, and for the Northern and Southern groupings analysed separately (Fig. 3.4). Coalescent analysis yielded a low value of Θ at 0.026 ± 0.002 and a small growth parameter value of 1.52 ± 23.93. AMOVA estimated that 72.1% of genetic variability was

Figure 3.3. Minimum spanning network (A - MSN) and median joining network (B – MJN) of the Kinabatangan long-tailed macaque mtDNA control region. Each circle represents a different haplotype and the diameter indicates haplotype frequency. The smallest circle represents a singleton. Mutational steps are represented by black bars on lines connecting haplotypes for MSN. In MJN, black dots are median vectors presumed unsampled or missing intermediates and numbers indicate the location of sites that have undergone mutations.
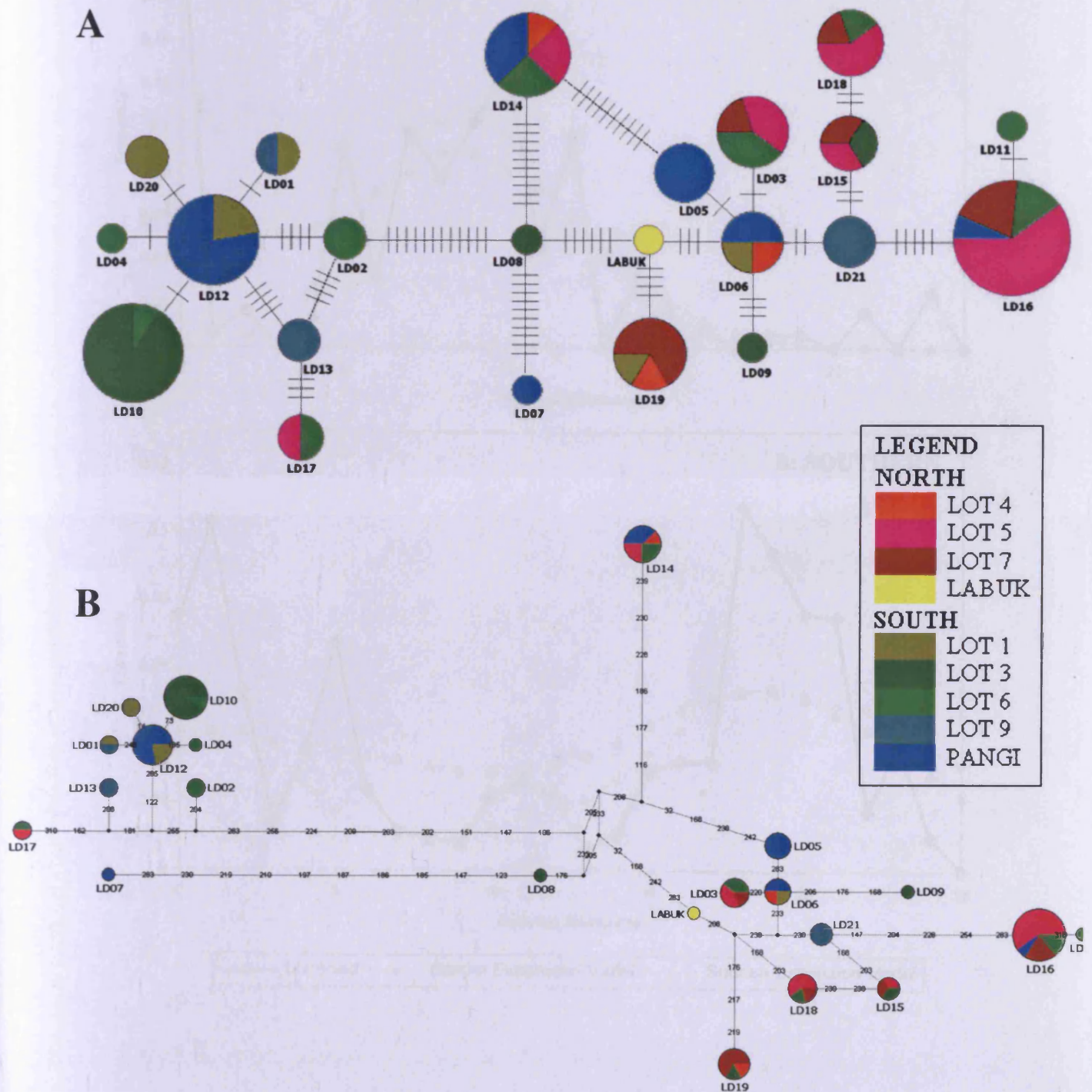
Figure 3.4. Mismatch distribution for long-tailed macaque left domain mtDNA control region sequences (LD). Observed (solid lines) and expected (dotted lines) showing the frequencies of pairwise differences within the (A) northern, and (B) southern riversides.
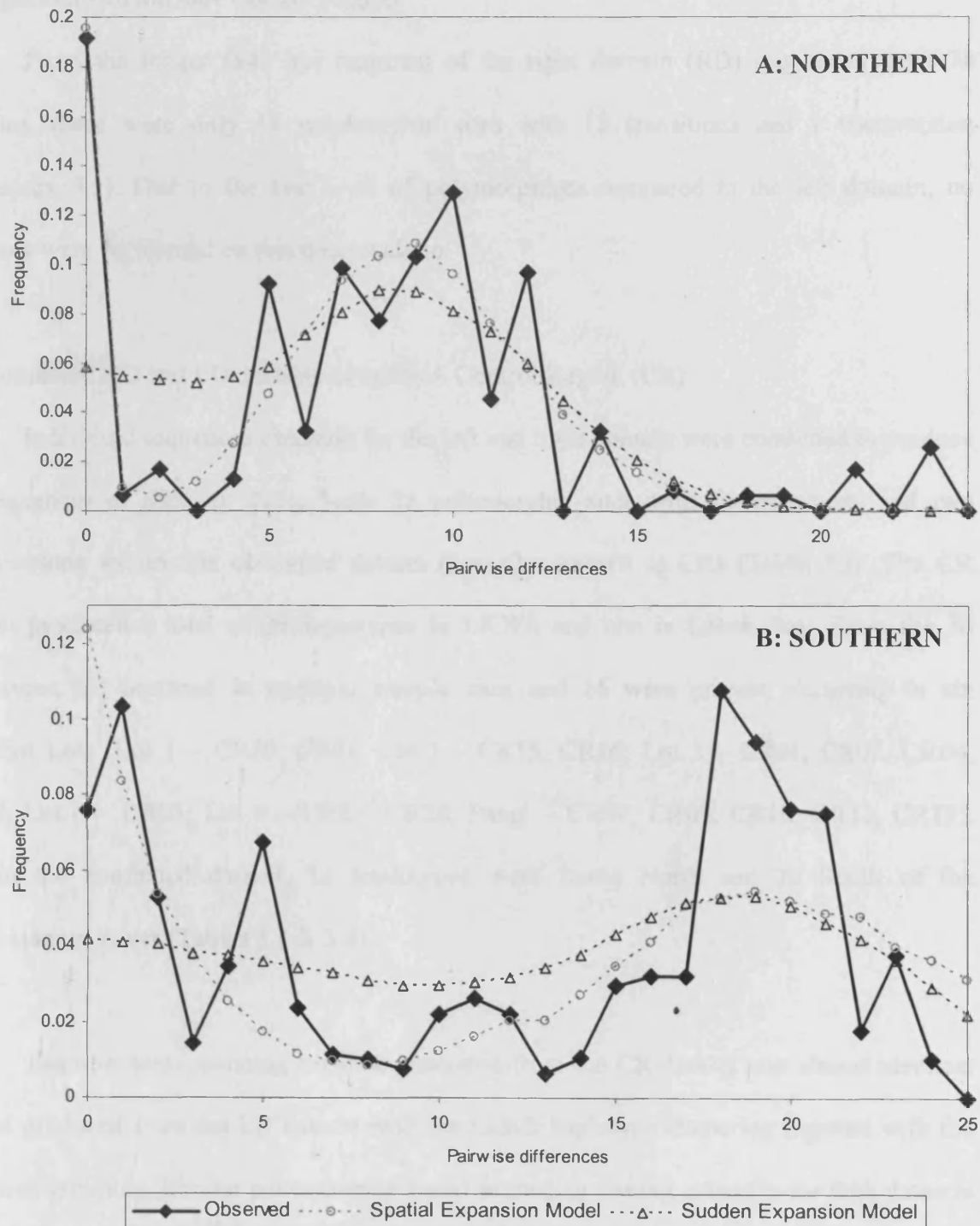
attributable to variance within populations and 27.9% was accounted for by differences among riversides. The riverside Φst value of 0.279 was highly significant (P=0.000).

II. Right Domain mtDNA Control Region

From the longer (545 bp) fragment of the right domain (RD) sequenced from 74 samples, there were only 13 polymorphic sites with 12 transitions and 1 transversion (Appendix 3.1). Due to the low level of polymorphism compared to the left domain, no analyses were performed on this dataset alone.

III. Combined RD and LD datasets of mtDNA Control Region (CR)

Individual sequences obtained for the left and right domain were combined to produce 74 sequences of 895 bp. There were 55 polymorphic sites with 53 transitions and two transversions within this combined dataset (hereafter known as CR) (Table 3.3). The CR dataset produced a total of 26 haplotypes in LKWS and one in Labuk Bay. From the 26 haplotypes, 10 occurred in multiple sample sites and 16 were private, occurring in six different Lots (Lot 1 – CR20, CR21; Lot 3 – CR15, CR16; Lot 5 – CR01, CR02, CR04, CR08; Lot 6 – CR05; Lot 9 – CR03, CR26; Pangi – CR07, CR09, CR10, CR12, CR13). Within the combined dataset, 12 haplotypes were found North and 20 South of the Kinabatangan River (Tables 3.3 & 3.4).

The minimum spanning network generated from the CR dataset was almost identical to that produced from the LD dataset with the Labuk haplotype clustering together with the Northern grouping. Similar patterns were found in median joining networks for both datasets (Figs. 3.5A & B). Haplotype diversity values obtained for the CR dataset were lower in the North (0.890 ± 0.035) compared to the South (0.901 ± 0.032) and the entire LKWS (0.938 ±

Table 3.3. Variable nucleotide positions in the alignment of mtDNA sequences of the long-tailed macaque control region. Dots indicate positions where the bases are identical to haplotype CR01. The vertical line marks the border between left and right domains of mtDNA control region.

```
         1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 | 3 3 4 5 5 6 6 6 6 6 8 8 8
       1 3 7 0 1 2 4 5 6 6 7 7 8 8 8 8 9 0 0 0 0 0 0 0 1 1 1 2 2 2 3 3 3 3 4 5 5 5 8 8 1 | 6 9 4 1 4 0 0 1 5 8 0 6 6
       4 2 3 6 6 2 3 7 1 2 8 6 7 1 5 6 7 7 2 3 4 5 6 8 9 0 7 9 0 4 8 0 3 8 9 2 4 5 6 3 5 0 | 2 6 3 1 1 1 5 2 2 3 2 1 8
CR01   A T A C C T C T C C G G A A T C T C A T G C G T C C A T T T T C T G A C A T C G C A | G T C T T G T A A C T A A
CR02   . C . . T . . C . . A . G . . T . . . . A . . . . . . . . . . . . . . . C A G T G . | . C . C . . C . . . C . .
CR03   . . . . . . . C . . . . . . . . . . . . A . . . . . . . . . . C . . . . . G . . A . . | . C . C . . . . . . . . .
CR04   . . . . . . . C . . A . . . . . . . . C A . . . . . . . . . . C . . . . . G . . A . . | . C . C . . . . . . C . .
CR05   . C . T . . . T . A . . . . . . . . . G C . T . C T . . . . C C T C A . T G . T A . . | A . . C C A . G G . C G G
CR06   . . . . . . . C . . . . . . . . . . . . A . . . . . . . C . C T C . . . G . . A . . | . . . C . . . . . . . C .
CR07   . . . . . . . C . . . . . . . . . . . . A . . . . . . . . . C T C . . . G . . . . . | . . . . . . . . . . . C .
CR08   . . . . . . . C . . A . . . . . . . . C A . . . . . . . . . C T . A . . G . . A . . | . . T . . . . . . . G C .
CR09   . . . . T . . C . . . . . . . . . . . . A . . . . . . . . . C T C . . . G . . . . . | A C . C C A . G G . C G G
CR10   . C . . . . . C . . A . G . . T . . . . A . . . . . . . . . . . . . . . C A G T G . | A C . C C A . G G . C G G
CR11   . . . . . . . C . . . . . . . . . . . . A . . . . . . . . . C T C . . . G . . A . . | . . . . . . . . . . . C .
CR12   . C . . . T . . . A A . . C T C A . . A T . C . T . C . . C . . A . T G . . A . . | A C . C C A . G G . C G G
CR13   . C . . T . . C . . A . G . . T . . . . A . . . . . . . . . . . . . . . . C A G T . . | A C . C C A . G G . C G G
CR14   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . | . C . C . . C . . . C . .
CR15   . C . . . . . C . . A A . . . . . . . . A T . C . . . . . . C T . A . T G . . . . . | A C . C C A . G G . C G G
CR16   . . . T . . . C . . A A . . . . . . . . A . A . . . . . . . C T C . . . G . . A . . | . C . . . . . . . . . C .
CR17   . C G . . C . . T . A . . . . . . . . G C A T . C T . . . . C C T C A . T G C T A T . | A C . C C A . G G . C G G
CR18   . . . . . . . C . . A . . . . . . . . . C A . . . . . . . . C T . A . . G . . A . . | . . T . . . . . . . . C .
CR19   . C . T . C . . T . A . . . . . . . . G C A T . C T . . . . C C T C A . T G C T A T . | A C . C C A . G G . C G G
CR20   . C . T . C . . T . A . . . . . . . . G C A T . C T . . . . C C T C A . T G C T A T . | A C . C C A . G G G C G G
CR21   G C . T . C . . T . A . . . . . . . . G C A T . C T . . . . C C T C A . T G C T A T . | A C . C C A . G G . C G G
CR22   . . . . . . . C . . A . . . . * . . . . C A . . . . . . . . . C . . . . . G . . A . . | . . . . . . . . . . . . .
CR23   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | . C . C . . . . . . C . .
CR24   . C . T . . . . T T A . . G . . . . . G C A T . C T . . . . C C T . C A . T G C T A . G | A . . C C A . G G . C G G
CR25   . . . . . . . C . . . A . . . . . . . . A . . . . . . . G C . . C T . A . . G . . A . . | . . T . . . . . . . . C .
CR26   . . . . . . . C . . . . . . . . . . . . A . . . . . . . . . . C . . . . . G . . A . . | . . . . . . . . . . . . .
LABUK  . . . . . . . C . . . . . . . . . . . . A . . C . . . . . . C T . A . . G . . A . . | A . . . . . . . . . . C G .
```

Chapter 3

Table 3.4. Haplotypes of the mtDNA control region (CR) for long-tailed macaques and their frequency in each Lot and forest reserve.

HAPLOTYPE	NORTHERN			SOUTHERN					LABUK
	LOT 4	LOT 5	LOT 7	LOT 1	LOT 3	LOT 6	LOT 9	PANGI	
CR01	-	1	-	-	-	-	-	-	-
CR02	-	1	-	-	-	-	-	-	-
CR03	-	-	-	-	-	-	1	-	-
CR04	-	1	-	-	-	-	-	-	-
CR05	-	-	-	-	-	2	-	-	-
CR06	-	2	1	-	-	2	-	-	-
CR07	-	-	-	-	-	-	-	2	-
CR08	-	1	-	-	-	-	-	-	-
CR09	-	-	-	-	-	-	-	1	-
CR10	-	-	-	-	-	-	-	2	-
CR11	1	-	-	-	-	-	-	2	-
CR12	-	-	-	-	-	-	-	1	-
CR13	-	-	-	-	-	-	-	1	-
CR14	-	3	-	-	-	-	-	1	-
CR15	-	-	-	-	1	-	-	-	-
CR16	-	-	-	-	1	-	-	-	-
CR17	-	-	-	-	10	1	-	-	-
CR18	-	2	-	-	-	1	-	-	-
CR19	-	-	-	1	-	-	-	7	-
CR20	-	-	-	1	-	-	-	-	-
CR21	-	-	-	1	-	-	-	-	-
CR22	1	-	1	-	-	-	-	-	-
CR23	-	5	3	-	-	1	-	-	-
CR24	-	1	-	-	-	1	-	-	-
CR25	1	-	4	-	1	-	-	-	-
CR26	-	-	-	-	-	-	2	-	-
LABUK	-	-	-	-	-	-	-	-	1
TOTAL	3	17	9	3	13	8	3	17	1

Chapter 3

0.013). Similarly, nucleotide diversity in the North (0.011 ± 0.006) was lower than that in the

South (0.018 ± 0.009) and entire LKWS (0.019 ± 0.010) (Table 3.5).

The Tajima's D value obtained for the Northern grouping was lower (-0.618) when

compared to the Southern grouping (0.828) or all of the LKWS (1.693). The values of Fu's Fs

were all positive, ranging from 0.833 in the Southern grouping to 1.442 in all LKWS to 1.604

in the Northern grouping, but none were significant (Table 3.5). The mismatch distribution for

all LKWS samples, the Northern and Southern subpopulations showed a multimodal pattern

(Fig. 3.6). Coalescent analysis yield a low value of Θ at 0.020 ± 0.002 and the growth

parameters also yielded a very low value of 49.15 ± 42.96 (Table 3.5). Comparing long-tailed

macaque populations on either side of the river, AMOVA analysis estimated variability

within populations of 62% and between riverside variability of 38%. Like the LD dataset, the

combined dataset produced Φst value of 0.380 which was again highly significant (P=0.000).

3.4   DISCUSSION

This is the first study of non-human primates to directly compare variation in the left

and right domain of the mtDNA control region in the same individuals. As in human studies

(e.g. Salas *et al.*, 2000), the left domain was much more variable with a 3 - 4 fold difference

in the number of polymorphic sites in 350 bp compared to 545 bp of the right domain. In fact

there appeared to be little value at all in additional sequencing of the right domain as

statistical analysis of the combined data sets produced virtually identical results to those

obtained from the left domain alone. Thus, the left domain seems the most productive and

cost effective sequence to target when evaluating primate control region genetic variation.

This study revealed high levels of genetic diversity in long-tailed macaques in the

Lower Kinabatangan, with 21 and 13 haplotypes detected from the left and right domain,

Figure 3.5. Minimum spanning network (A-MSN) and median joining network (B-MJN) of mtDNA control region haplotypes of the long-tailed macaque. The diameter of the circle indicates number of sequences. The smallest circle represents a singleton. Black bars in MSN and numbers in MJN indicates the mutational steps (MSN) and the sites that undergone mutations. Black dots (in MJN) represent missing samples.
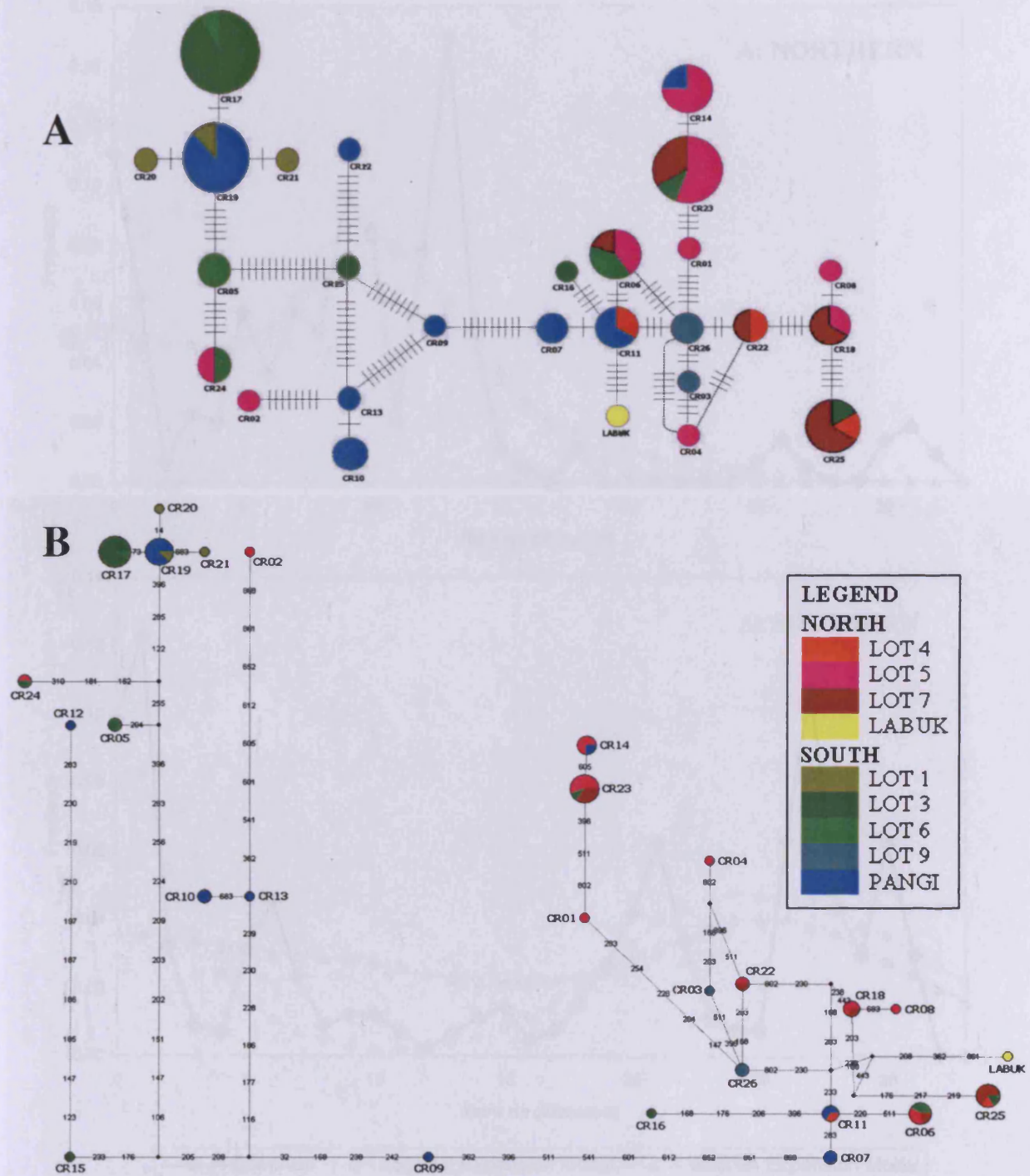
Figure 3.6. Mismatch distribution for long-tailed macaque mtDNA control region sequences (CR). Observed (solid lines) and expected (dotted lines) showing the frequencies of pairwise differences within the (A) northern, and (B) southern riversides.
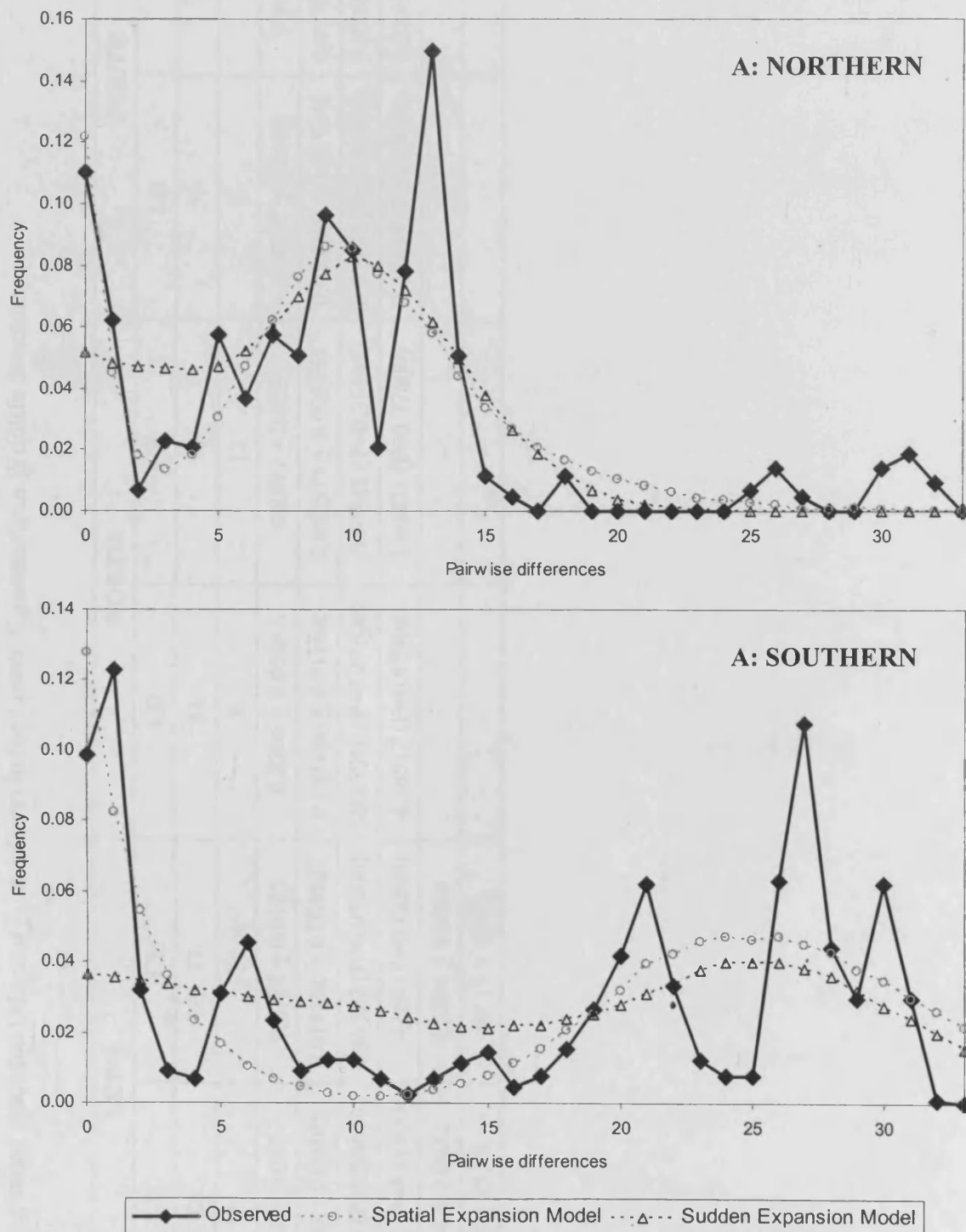
Table3.5. Summary on number of sequences and haplotypes, values of nucleotide (π) and haplotype (h) diversity, neutrality tests (Tajima's D and Fu's Fs), population parameters of Θ and growth parameter of g for mitochondrial DNA control region datasets (RD, LD & CR datasets) for long-tailed macaque (*Macaca fascicularis*) in the Lower Kinabatangan Wildlife Sanctuary.

	LKWS		NORTH		SOUTH	
	LD	**CR**	**LD**	**CR**	**LD**	**CR**
Seq.	88	73	31	30	57	43
Hap.	21	26	8	12	8	12
h	0.926 ± 0.012	0.9384 ± 0.0125	0.8206 ± 0.0506	0.8897 ± 0.053	0.9242 ± 0.0188	0.9014 ± 0.0315
π	0.03520 ± 0.00105	0.019108 ± 0.009527	0.021866 ± 0.011645	0.010575 ± 0.005555	0.034060 ± 0.017345	0.017523 ± 0.008856
D	1.42336 (P=0.94600)	1.69302 (P=0.97000)	-0.13050 (P=0.47600)	-0.61841 (P=0.31500)	0.93572 (P=0.91668)	0.82812 (P=0.84700)
Fs	2.67646 (P=0.81900)	1.44169 (P=0.72300)	4.36517 (P=0.93900)	1.60421 (P=0.77800)	0.38459 (P=0.60600)	0.38459 (P=0.60600)
Θ	0.0264 ± 0.0022	0.0204 ± 0.0019	-	-	-	-
g	1.52 ± 23.93	49.15 ± 42.96	-	-	-	-

respectively. In contrast, Perwitasari-Farajallah *et al.* (1999) studying the same species using PCR-RFLP of the control region, identified only five haplotypes from West Java. Even though direct comparison is not valid, our small geographic coverage produced a higher haplotype diversity illustrating the increase in sensitivity of PCR sequencing over PCR-RFLP. From 392 bp of the left domain of the mtDNA control region, Marmi *et al.* (2004) detected only nine haplotypes from 50 individuals of the Japanese macaque, *Macaca fuscata,* from mainland Japan. Using the right domain, Hayaishi & Kawamoto (2006) studied the same species, from a small island (Yakushima), and found only six haplotypes from 488 faecal samples collected from the whole island. However, although Hayaishi & Kawamoto (2006) genotyped almost 10 times as many individuals as Marmi *et al.* (2004) the results are not comparable as different domains were targeted. From the Barbary macaque (*Macaca sylvanus*), Modolo *et al.* (2005) identified 24 haplotypes from the left domain of the control region from 212 samples across its entire geographic range. This low level of diversity probably reflects population bottlenecks due to contraction of suitable habitats (Modolo *et al.,* 2005). With an estimated population of ~10,000 Barbary macaques remaining in the wild, the remaining populations clearly need proper management to maintain genetic diversity (Modolo *et al.,* 2005).

Neutrality tests, mismatch distribution and coalescent analysis all indicated that the Lower Kinabatangan long-tailed macaque population has remained stable for a long period of time and has not undergone a demographic expansion in the recent evolutionary past. This long-term stability could be due to their adaptability to different habitats and being opportunistic, especially within human settlements in the Kinabatangan basin, although this would not explain their post-glacial demographic profile, since humans only arrived in the region relatively recently. Nonetheless, wild primates with access to human waste or other such food resources are much heavier with significantly higher birth (hence productivity)

rates, especially as in well fed populations females mature earlier (Wheatley *et al.,* 1996; Cowlishaw & Dunbar, 2000).

The social system of long-tailed macaques consists of multi-male and multi-female groups dominated by an alpha male and alpha female. de Ruiter & Geffen (1998) found that within Sumatran long-tailed macaques, a stable social group comprised eight to 60 individuals. In our Kinabatangan study, we observed group sizes ranging from two to 40 individuals, but the most common group size was 10 to 20 (see Chapter 2). This species is known to prefer open, degraded habitats (Supriatna *et al.*, 1996). The differences in group size between de Ruiter & Geffen's (1998) and the current study could potentially be attributed to the more degraded habitat in LKWS or a lack of resources to sustain such a large group size like in Sumatra, although this is clearly speculation.

Perhaps surprisingly, evidence of genetic differentiation was detected between long tail macaque populations on each side of the Kinabatangan River. However, as these primates are known to be good swimmers, the differences between riversides may actually reflect differences in social structure between populations. Like other macaques species, female long-tailed macaques are highly philopatric (Melnick, 1987; Hoelzer *et al.*, 1994; de Ruiter & Geffen, 1998) and social groups tend to consist of closely related females. Female philopatry is common in cercopithecoid primates usually associated with strong matrilineal bonds among sisters/females (Difiore & Randall, 1994; Cowlishaw & Dunbar, 2000). New social groups are formed, usually through group-fission (de Ruiter, 1994; de Ruiter & Geffen, 1998) which results in closely related females staying together and unrelated females forming a new 'daughter' groups (Hoelzer *et al.*, 1994; de Ruiter & Geffen, 1998; Cowlishaw & Dunbar, 2000). These new groups usually occupy available habitat nearby the source group as demonstrated by toque macaques in Sri Lanka (Hoelzer *et al.*, 1994) and long-tailed

macaques in Sumatra (de Ruiter, 1994) and Sabah (current study). 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 (Hoelzer *et al.*, 1994).

For long-tailed macaques in Sumatra, de Ruiter & Geffen (1998) found that the river did not significantly disrupt gene-flow, instead relatedness values were similar for adults on one and both sides of the river, which was interpreted as being due to female philopatry and the formation of new groups by group fission along matrilineal lines. This result is supported by the current study which showed clustering of closely related haplotypes within and between adjacent Lots. Due to strong genetic drift operating at the social level, the linear habitat along the river in Sumatra may enhance group differentiation by a 'stepping stone' configuration among social groups (Scheffrahn *et al.*, 1996). Amazonian primate communities on facing riverbanks are more similar across narrow, slow flowing rivers, but less similar across faster flowing rivers (Ayres & Clutton-Brock, 1992). However, on flood plains the increased occurrence of the formation of oxbow lakes on large, slow flowing rivers, may lead to increased opportunities for primates to move between riverbanks. Another factor which might influence movement of individuals across rivers is the formation of arboreal bridges (tree canopies) over narrower channels (Cowlishaw & Dunbar, 2000). This situation has been observed in Chinese and Indian Rhesus macaques (Smith & McDonough, 2005). Despite the hazardous terrains of the Himalayan glaciers, it is considered unlikely that the Brahmaputra River Valley barrier (Melnick *et al.* 1993) restricted gene flow of the Rhesus macaques, as this species is known to inhabit regions in the Himalayas as high as 3,500 m (Smith & McDonough, 2005). The most likely explanation for differentiation between Indian and Chinese rhesus macaques is climatic. Towards the end of the Pleistocene, climatic

changes, including the desiccation of the tropical forest, caused near extinction of these

macaques and a bottleneck that persisted until quite recently (Smith & McDonough, 2005).

Unlike de Ruiter's (1994) work in Sumatra, populations in the current study were

sampled essentially at random with little information on social group origin of samples.

Genetic analysis revealed a mixture of closely related with some very distant related lineages

without any apparent structure except that many of the lineages were distributed between

adjacent Lots in LKWS but some lineages are unique to one Lot only (i.e. LD02 in Lot 6 and

CR03 in Lot 9). Here, the high mobility of macaques might contribute this kind of movement

but an alternative explanation is that groups were forced into the Kinabatangan from other

areas (adjacent forest) during the deforestation that occurred in the region during the 1960's

and 1970's and have subsequently interbred with the local long-tailed macaques hence the

highly differentiated lineages. Alternatively, several invasions of Kinabatangan by long-tailed

macaques might have occurred from the surrounding forest.

Considering the stability of the LKWS macaque population, it is interesting to ask the

question how can a relatively small-bodied monkey apparently out-compete the other nine

primate species in the same area? If we look at the LKWS primate fauna, two of them are

nocturnal prosimians (*Tarsius bacanus* and *Nycticebus coucang*) which have different dietary

requirement and are too small to compete with long-tailed macaques. Secondly the langurs,

Presbytis rubicunda, *P. hosei* and *Trachypithecus cristata*, which live in smaller social

groups, high up in the canopy, usually only compete for resources within the canopy itself as

they seldom move down to the ground in the way that macaques are known to. Thirdly, the

proboscis monkey (*Nasalis larvatus*) with its highly specialized diet does not compete with

long-tailed macaques; in fact on many occasions, both species forage in the same area

(Kawabe & Mano, 1972). The Bornean gibbon (*Hylobates muelleri*) and the orang-utan

---

(*Pongo pygmaeus*) might provide some competition for food resources but being either monogamous or solitary species occurring at relatively low densities, they will easily be out-competed by the numerous long-tailed macaques. The only other species that could provide competition to the long-tailed macaque is the pig-tailed macaque, however as both species show different habitat preferences (mainly terrestrial foragers in hill forest for pig-tailed and mainly arboreal foragers in riverine forest for long-tailed macaques), potential competition is reduced. Thus, the long-tailed macaque seems to flourish more than other primates in Kinabatangan as the forest is rich with fruiting trees and macaques are the only primates to obtain food from humans by crop raiding.

The current study had demonstrated that the left domain of the control region mtDNA is a suitable marker for intraspecific phylogeography study of long-tailed macaques and other primates, and that there is no additional benefit for phylogeographic studies in also sequencing the right domain. The study also revealed that Kinabatangan long-tailed macaques have a stable population, high level of genetic diversity and that differentiation between riversides is the merely the product of their social systems. The success of long-tailed macaques in Kinabatangan could be attributed to their adaptability, reflected by the fact that this genus is the most successful of all non-human primates (Abegg & Thierry, 2002).

CHAPTER 4

RIVERS INFLUENCE THE POPULATION GENETIC STRUCTURE OF BORNEAN ORANG-UTANS (*PONGO PYGMAEUS*)

Jalil, M.F.[1,2], Cable, J.[1], Sinyor, J.[3], Bruford, M.W.[1] and Goossens, B.[1,2]

[1] School of Biosciences, Cardiff University, Cardiff CF10 3TL, UK.

[2] Institut Biologi Tropika & Pemuliharaan, Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah, Malaysia.

[3] Kinabatangan Orangutan Conservation Project, P.O.Box 3109, 90734 Sandakan, Sabah, Malaysia.

Abstract

We examined mitochondrial DNA control region sequences of Kinabatangan orang-utans to test the hypothesis that the genetic structure of the Bornean orang-utan is influenced by major rivers. Our samples from the Lower Kinabatangan Wildlife Sanctuary represent the northern-most population of orang-utans (*Pongo pygmaeus*) on Borneo and they are separated into two subpopulations by the Kinabatangan River, the longest river in Sabah. Orang-utan samples on either side of the river were significantly differentiated by a high Φ_{ST} value of 0.404 (P=0.000). Our data also revealed an east-west gradient of genetic diversity and evidence for a population expansion along the river, possibly reflecting a post-glacial colonization of the Kinabatangan. We analysed our data together with previously published sequences of Bornean orang-utans and discuss the general importance of rivers as barriers to gene flow in this long-lived, solitary arboreal ape.

4.1   Introduction

Phylogeographic studies examine the biogeographic patterns of molecular sequences in a species and the data are used to infer population processes from an evolutionary (historical) perspective (Avise *et al.*, 1987). Biogeographic barriers, such as rivers (Telfer *et al.*, 2003; Eriksson *et al.*, 2004) and mountain ranges (Hewitt, 2000), may disrupt gene flow

thus shaping genetic structure. Comparative phylogeography analyses geographic patterns of genetic variation across species in order to find general patterns of evolutionary history within biogeographic regions and to reveal the evolutionary processes behind these patterns (Bermingham & Avise, 1986; Eizirik *et al.*, 1998). Nevertheless for widespread species, genetic divergence may occur simply as a function of distance (isolation by distance) due to limited gene flow (Avise *et al.*, 1987). At a regional level, comparative phylogeography can be used for conservation purposes to identify areas of high biodiversity and hence high conservation value (Moritz & Faith, 1998).

The orang-utan (*Pongo pygmaeus*), the largest arboreal ape, is found on Sumatra and Borneo in South-East Asia. Traditionally, the orang-utan was classified into two subspecies, *P. pygmaeus pygmaeus* in Borneo and *P. p. abelii* in Sumatra. However, recent molecular data has led to the re-classification of the orang-utan into two distinct species, *P. pygmaeus* and *P. abelii* (see Xu & Arnason, 1996, but see Muir *et al.*, 1998; Muir, 2000; Zhang *et al.*, 2001). Furthermore, based on mitochondrial control region DNA data, Warren *et al.* (2001) suggested there are four distinct evolutionary groups within the Bornean orang-utan, corresponding to populations living in (i) Sabah, (ii) Sarawak and northwest Kalimantan, (iii) southwest and central Kalimantan, and (iv) east Kalimantan. Warren *et al.* (2001) also suggested that these subpopulations should be treated as separate units for conservation with their genetic integrity maintained.

Genetic studies on orang-utans to date have been limited, with most work focusing on the taxonomic status of the Bornean and the Sumatran sub-species (Xu & Arnason, 1996; Zhi *et al.*, 1996; Muir *et al.*, 2000; Zhang *et al.*, 2001), rather than population genetic structure (Warren *et al.*, 2000, 2001; Goossens *et al.*, 2005). Furthermore, most of these studies have utilised invasive samples (see Kanthaswamy *et al.*, 2006) from zoo and rehabilitation centres

with very limited non-invasive sampling from wild animals (Kanthaswamy & Smith, 2002). Only recently have completely non-invasive samples (i.e. hair and faeces) been utilized to study wild orang-utans, as part of the antecedent study to this work, based in the Lower Kinabatangan floodplain (Goossens *et al.*, 2005, 2006a, b).

In the current study, the structure and historical demography of the fragmented orang-utan populations previously analysed by Goossens *et al.* (2005, 2006b) were re-examined using mitochondrial control region sequences. As mitochondria are maternally inherited in a single copy, any reduced migration rates in females (as has been inferred for Sumatran orang-utans; see Utami *et al.*, 2002) are predicted to lead to an increased structuring of mitochondrial lineages. Such demographic effects are also predicted to result in a decrease in intra-population mitochondrial variability and to increase mitochondrial differentiation between subpopulations. Recently, however, Goossens *et al.* (2006b) showed males and females to be equally philopatric in the Kinabatangan using microsatellite markers, casting doubt on whether mitochondrial DNA would show a different structure to nuclear DNA. Here, we analysed samples from the northern and southern side of the river to investigate phylogeographic structure and gene flow along an almost linear habitat. In addition, sequences from Warren *et al.* (2001) with known localities in Borneo were used to analyse the phylogenetic relationships between Lower Kinabatangan Wildlife Sanctuary samples and localities elsewhere in Borneo.

4.2   Material and methods

DNA Samples

Faecal samples (n = 73) used in the current study were collected by BG and members of the Kinabatangan Orang-utan Conservation Project (KOCP) in 2001 from all 10 Lots in the Lower Kinabatangan Wildlife Sanctuary (LKWS), Sabah (Fig. 4.1). Details of sampling and

Figure 4.1. Sampling localities (yellow squares) for orang-utan samples in Lower Kinabatangan Wildlife Sanctuary (LKWS). The red line indicates the Lower Kinabatangan Wildlife Sanctuary (LKWS) boundaries for each Lots and blue lines indicate the Kinabatangan River and its tributaries.

Table 4.1. List of orang-utan samples from Warren *et al.* (2001) analysed in the current study.

Samples	Assession No.	Locality
DS29	AJ391099	Leboyan, Danau Sentarum, Northwest Kalimantan
DSLE1	AJ391100	Leboyan, Danau Sentarum, Northwest Kalimantan
DSME1	AJ391101	Meliau, Danau Sentarum, Northwest Kalimantan
DSME2	AJ391102	Meliau, Danau Sentarum, Northwest Kalimantan
DSRA	AJ391103	Radai, Danau Sentarum, Northwest Kalimantan
GP31	AJ391105	Gunung Palung, Southwest Kalimantan
GPJA	AJ391106	Gunung Palung, Southwest Kalimantan
GPMA	AJ391107	Gunung Palung, Southwest Kalimantan
GPUN	AJ391108	Gunung Palung, Southwest Kalimantan
KA1	AJ391109	Sangatta, East Kalimantan
KPC	AJ391110	Sangatta, East Kalimantan
SB372	AJ391116	Sepilok, Sabah
SB57	AJ391117	Sukau, Kinabatangan, Sabah
SB60	AJ391118	Kinabatangan, Sabah
SB70	AJ391119	Lahad Datu, Sabah
SB71	AJ391120	Sandakan, Sabah
SE8	AJ391121	Semongok, Sarawak
SEAH	AJ391122	Semongok, Sarawak
SEBU	AJ391123	Semongok, Sarawak
SEOA	AJ391124	Semongok, Sarawak
SEUA	AJ391125	Semongok, Sarawak
TNK36	AJ391133	Kutai National Park, East Kalimantan
TNK37	AJ391134	Kutai National Park, East Kalimantan
TNK39	AJ391136	Kutai National Park, East Kalimantan
TNK41	AJ391137	Kutai National Park, East Kalimantan
TP14	AJ391138	Tanjung Putting, Central Kalimantan
TP15	AJ391139	Tanjung Putting, Central Kalimantan
TP24	AJ391140	Tanjung Putting, Southwest Kalimantan
TP6	AJ391141	Tanjung Harapan, Tanjung Putting, Central Kalimantan

verify amplification. PCR products were cleaned with an Exonuclease I (10 units/μl) (USB

Corp. USA) and Shrimp Alkaline Phosphatase (1 unit/μl) (USB Corp. USA) mix at a ratio of

10μl PCR product to 1μl mix (Exonuclease I: Shrimp Alkaline Phosphatase – 1:1) (Hanke &

Wink, 1994). The product was incubated at 37°C for 1 h and at 80°C for 15 min to deactivate

the enzymes. Sequencing PCR was performed using ABI Big Dye Terminator vs. 1 (Applied

Biosystems). Each PCR product was sequenced in both directions. PCR products were

precipitated by adding 90μl of 63% isopropanol to each PCR tube. The product and

isopropanol were vortexed for 20 s, left to stand for 15 min, then centrifuged for 30 min at

13,000 g. The supernatant was discarded and 150μl of 70% isopropanol was added to each

PCR tube. The PCR strips were then centrifuged for one min at 500 g and the pellets dried at

52°C for 2 min. Sequencing was performed in an ABI3100 automated sequencer. The mtDNA

control region sequences were aligned using the program SEQUENCHER 3.1.2 (GeneCodes

Corp.) with correction by eye. A BLAST (Basic Local Alignment Search Tool) nucleotide

search was performed on each sequence.

Phylogenetics, mtDNA diversity and population structure

Sequences were collapsed into unique haplotypes using DAMBE 4.2.13 (Xia & Xie,

2001). Intra-specific gene genealogies were inferred using two network construction methods,

minimum spanning and median joining using ARLEQUIN vs. 3 (Excoffier *et al.*, 2005) and

NETWORK vs. 4.1.1.1, respectively (available at http://www.fluxus-engineering.com/)

(Bandelt *et al.*, 1999). Intra-specific nucleotide level (π) and haplotype diversities (h) were

estimated using ARLEQUIN (Tajima, 1983; Excoffier *et al.*, 2005). Analysis of molecular

variance (AMOVA) was used to analyze how genetic variability was partitioned within and

between riversides using Φ-statistics in ARLEQUIN. This approach incorporates information

on the absolute number of differences among haplotypes and haplotype frequencies. The

significance of variance, designated Φ-statistics (F statistic analogue), was tested by 1000

random permutations. In order to test the sequences for deviation from the expectations based on neutral theory, Tajima's D (Tajima, 1989), Fu's F_s (Fu, 1997), Fu and Li's D* (Fu & Li, 1993), and Fu and Li's F* (Fu & Li, 1993) were calculated using ARLEQUIN and DNASP 4.10.3 (Rozas *et al.*, 2003). Past demography was also assessed by mismatch distribution (distribution of pairwise sequence differences; Rogers & Harpending, 1992) in ARLEQUIN based on a spatial and sudden expansion models for orang-utan samples at three different levels; for each riverside, for Sabah and for Borneo. A coalescent-based simulation method to test for evidence of population expansion was also carried out, implemented in FLUCTUATE v.1.4 (Kuhner *et al.*, 1998). The model was run five times to ensure convergence of the estimates.

4.3   RESULTS

I. *Lower Kinabatangan Wildlife Sanctuary (LKWS)*

Genetic diversity

From all 73 DNA samples, 323 bp of the left hypervariable domain of the control region was successfully amplified (Table 4.2). From the 323 nucleotides, 314 were invariant and nine were variable with eight transitions and one transversion.

Amongst the 13 haplotypes identified (Table 4.2), OU11 and OU12 were dominant and found in 28 (38.3%) and 25 (34.2%) samples, respectively. OU11 was identified in eight out of nine Lots (not in Lot 1) and OU12 in six Lots (absent from Lots 2, 4 and 5). Only haplotypes OU10 (North = 4; South = 1), OU11 (North = 21; South = 7) and OU12 (North = 5; South = 20) were found on both sides of the river. Haplotypes OU01, OU02, OU07, OU08 and OU13 were present exclusively on the Northern side, whereas OU03, OU04, OU05, OU06 and OU09 were only on the Southern bank. Besides haplotypes OU11 and OU12, only

Table 4.2. Condensed matrix displaying variables sites of the 323 bp alignment of the mtDNA control region for 13 haplotypes found in Lower Kinabatangan orang-utan. Haplotype codes and nucleotide position are displayed on the left, and haplotype frequencies for each Lot (L) are given on the right.

	VARIABLES SITES									NORTH					SOUTH			
	99	175	209	226	243	255	261	393	308	L2	L4	L5	L8	L10	L1	L3	L6	L9
OU01	C	A	C	C	A	G	T	C	T			1						
OU02	.	.	.	.	.	.	.	A	C	1								
OU03	.	G	.	T	.	.	C	.	C								1	
OU04	T	.	.	.	G	A	.	.	C						2			
OU05	T	.	.	.	.	A	.	.	C						1		1	1
OU06	.	.	.	.	.	A	.	A	C						1			
OU07	.	G	T	.	.	.	.	.	C			1						
OU08	.	G	T	.	.	A	.	.	.		1							
OU09	.	.	.	T	.	A	.	.	C							3		
OU10	.	G	.	.	.	.	.	.	.	4								1
OU11	.	G	T	.	.	.	.	.	.	3	10	5	2	1		3	3	1
OU12	.	.	.	.	.	A	.	.	C				1	4	8	4	5	3
OU13	T	.	T	.	.	.	.	.	.			1						
TOTAL										8	11	8	3	5	12	10	10	6

OU05 and OU10 were recorded in more than one Lot (OU05 = Lot 1, 6 and 9; OU10 = Lot 2 and 9) (see Table 4.3).

A minimum spanning network connecting the 13 haplotypes in a linear form revealed a partial separation into two groups, largely corresponding to populations on the northern and southern sides of the river (Fig. 4.2A). Haplotypes OU12 and OU11, distinguished by 10 substitutions, were the most common haplotypes on each riverside. There were four haplotypes between OU11 and OU12, two haplotypes radiating from OU12 and three haplotypes from OU11. The haplotypes are similarly grouped in a median joining network (Fig. 4.2B), which also highlights five potentially unsampled or missing haplotypes.

Based on AMOVA, the genetic variation in LKWS orang-utans was mostly attributable to among rather than within populations (Φ_{ST} = 0.404; P=0.000) suggesting that the river is a significant barrier to gene flow. Within the Northern riverside, the Φ_{ST} value was higher (Φ_{ST} = 0.388, P=0.000) than for the Southern riverside (Φ_{ST} = 0.067, P=0.123) indicating genetic structure on the Northern but not on the Southern riverside. Nucleotide diversity (π) ranged from 0.005 $\pm$ 0.003 on the Northern riverside to 0.008 $\pm$ 0.005 in the whole of LKWS. Overall haplotype diversity (h) within LKWS was 0.734 $\pm$ 0.035. Comparing the North and South populations, the latter had the highest nucleotide (π, 0.006 $\pm$ 0.004) and haplotype diversity (h, 0.690 $\pm$ 0.071) (Table 4.3). The overall high haplotype and low nucleotide diversity in the LKWS indicates a population bottleneck event followed by rapid population growth and accumulation of mutations.

Historical demography

Neutrality tests were performed to detect for additional evidence of population expansion. Negative values indicate the presence of either some form of selection (unlikely at

Table 4.3. Number of sequences and haplotypes, nucleotide diversity (π), haplotype diversity (h), test of selective neutrality (Tajima's D, Fu's F_S, Fu & Li's D* & Fu & Li's F*) and population parameters of theta (Θ) and growth parameter (g) of orangutan mtDNA control region sequences for LKWS, Sabah and Borneo.

Samples	ALL LKWS	LKWS North	LKWS South	SABAH	BORNEO
Sequences	73	35	38	78	102
Haplotypes	13	8	8	17	40
h	0.736 ± 0.035	0.620 ± 0.086	0.690 ± 0.071	0.768 ± 0.034	0.865 ± 0.025
π	0.007 ± 0.005	0.005 ± 0.003	0.006 ± 0.004	0.011 ± 0.007	0.026 ± 0.014
D	0.789 (P=0.837)	0.162 (P=0.591)	-0.157 (P=0.471)	-0.798 (P=0.236)	-1.582 (P=0.030)
Fs	-2.586 (P=0.170)	-1.920 (P=0.120)	-0.811 (P=0.371)	-5.147 (P=0.023)	-18.897 (P=0.000)
D*	-	-	-	-	-3.140 (P<0.05)
F*	-	-	-	-	-3.046 (P<0.02)
Θ	0.041 ± 0.003	0.008 ± 0.001	0.006 ± 0.001	0.053 ± 0.004	0.177 ± 0.019
g	305.44 ± 36.70	170.68 ± 173.35	119.42 ± 130.52	526.12 ± 56.03	147.54 ± 16.96

the non-coding control region) or population expansion. Based on Table 4.3, all datasets (LKWS, North and South) had negative values for Fu's F_S, indicating population expansion, although these values were not significant at the 95% level. In contrast, Tajima's D results were much lower (<1.0) and only one dataset (South) showed a negative value, but again this was not significant (Table 4.3).

The bimodal mismatch distributions for LKWS, the Northern and Southern populations (Fig. 4.3) indicate two population expansions or the presence of two or more mixed populations that have subsequently expanded. The latter interpretation is clearly supported by the network in which the haplotypes OU11 and OU12 were found on both sides of the river. A coalescent approach to detect population expansion using FLUCTUATE indicated positive estimates for the growth parameters for all sample groups (LKWS, Northern and Southern), (Table 4.3). Estimates of theta ($\Theta \pm$ SD) were 0.041 ± 0.003 (LKWS), 0.008 ± 0.001 (North) and 0.006 ± 0.001 (South). The growth parameters (g $\pm$ SD) were 305.44 ± 36.70 (LKWS), 170.68 ± 173.35 (North) and 119.41 ± 130.52 (South) which indicates population growth. Both positive theta values and high values of growth parameters support the results of neutrality test and mismatch distribution that the Kinabatangan orang-utan population is expanding.

II. *Sabah*

Genetic diversity

The LKWS samples were combined with five Sabah samples from Warren *et al.* (2001) and reanalysed. From 78 nucleotides, there were a total of 17 haplotypes with 18 polymorphic sites. There were 17 transitions and one transversion in the dataset with a haplotype diversity *h* of 0.768 ± 0.034 and nucleotide diversity π of 0.011 ± 0.007. Out of five Warren samples, only one, SB70 (AJ391119, Lahad Datu, Sabah) matched the LKWS

Figure 4.2. Minimum spanning (A - MSN) and median joining networks (B - MJN) of the Kinabatangan orang-utan sequences. Each circle represents a haplotype and the diameter scales to haplotype frequency. The smallest circles represent singletons. Mutational steps are represented by black bars on lines connecting haplotypes (MSN). In the MJN, black dots are median vectors presumed unsampled or missing intermediates and numbers indicate the locations of site that have undergone mutations.
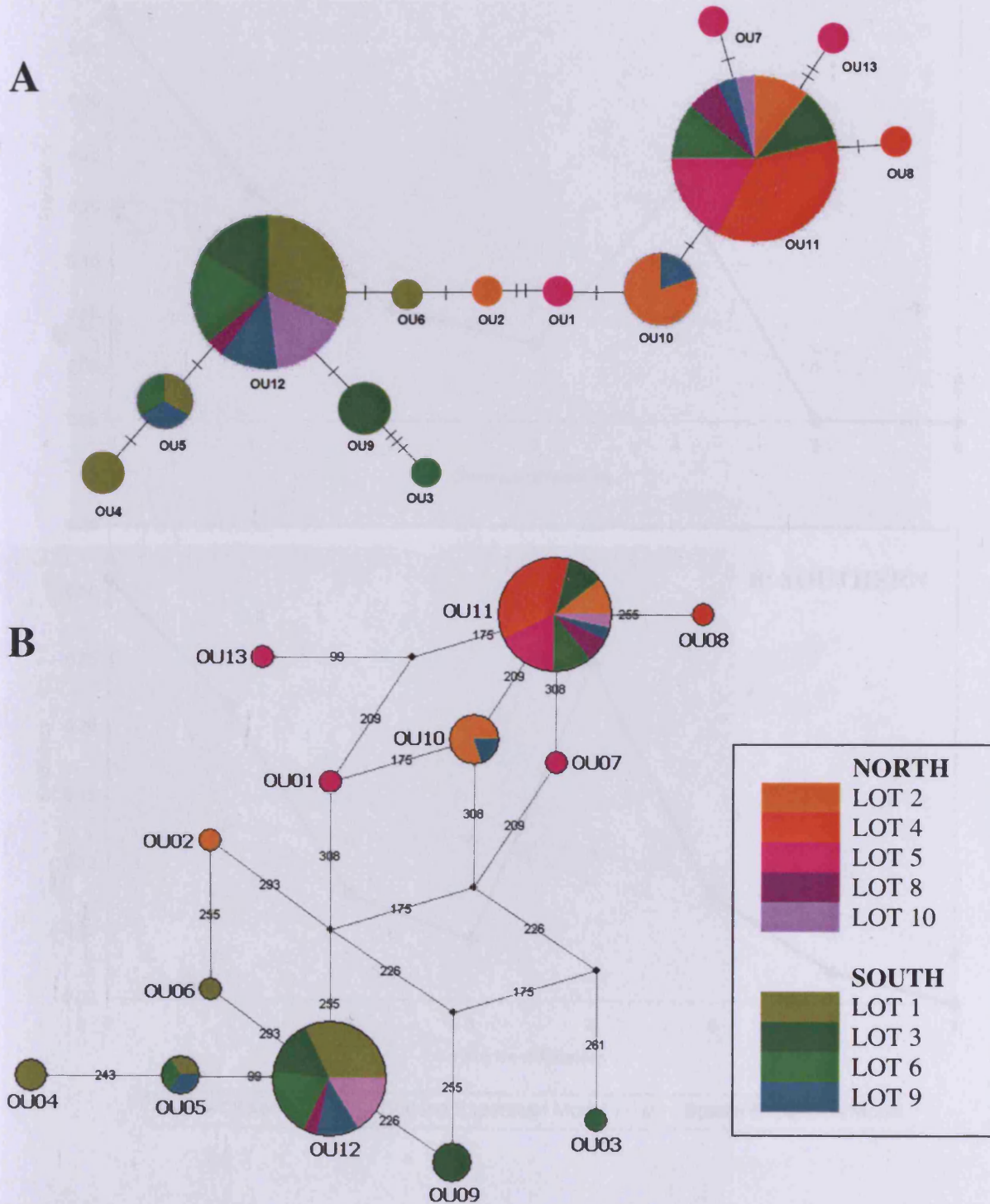
Figure 4.3. Mismatch distribution for each of the analysed LKWS orang-utan samples. Observed (solid lines) and expected (dotted lines) showing the frequencies of pairwise differences within the (A) northern, and (B) southern riversides.
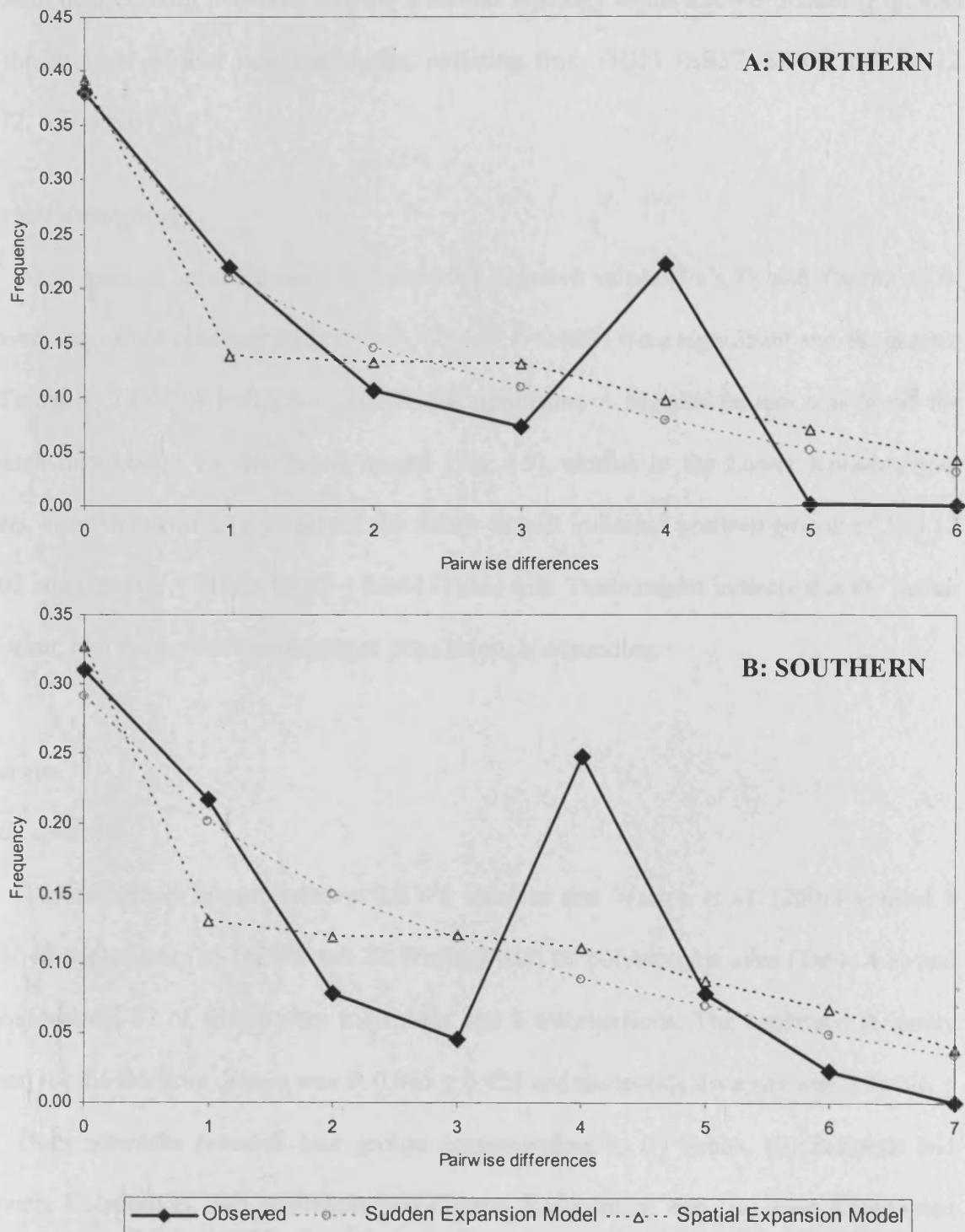
haplotypes, OU08. Four other haplotypes grouped within the dominant Northern haplotypes, OU11 (SB57-AJ391117) or OU12 (SB372-AJ391116; SB71-AJ391120; SB60-AJ39118). For the Sabah dataset, both networks showed a similar topology to the LKWS dataset (Fig. 4.4) with the addition of four new haplotypes, radiating from OU11 (SB57, SB70) and OU12 (SB372, SB60, SB71).

Historical demography

Both tests of selective neutrality revealed negative values (Fu's F_S and Tajima's D). However, the values obtained from Fu's F_S (-5.147 P=0.023) were significant and far greater than Tajima's D (-0.798 P=0.236) which is not significant. A bimodal pattern was found for mismatch distribution for the Sabah dataset (Fig 4.5), similar to the Lower Kinabatangan datasets. Growth estimate (g $\pm$ SD) for the Sabah dataset indicated positive growth of 526.12 $\pm$ 56.03 and theta ($\Theta \pm$ SD) of 0.053 $\pm$ 0.004 (Table 4.3). These results indicate that the Sabah population, like the Lower Kinabatangan population, is expanding.

III. *Borneo*

Genetic diversity

The combined dataset between LKWS samples and Warren *et al.* (2001) yielded a total of 40 haplotypes (13 LKWS and 27 Warren) with 62 polymorphic sites (Table 4.3) and 65 substitutions, 57 of which were transitions and 8 transversions. The haplotype diversity obtained for the Bornean dataset was *h*: 0.865 $\pm$ 0.025 and nucleotide diversity was π: 0.026 $\pm$ 0.014. Both networks revealed four groups corresponding to (i) Sabah, (ii) Sarawak and Northwest Kalimantan, (iii) southwest and Central Kalimantan, and (iv) East Kalimantan (Fig. 4.6). However, median joining network present the result better than minimum spanning network because the four grouping were well separated from each other. The closest to the Sabah samples were the East Kalimantan group, which were separated by four mutations (Fig. 4.7).

Figure 4.4. Minimum spanning network (A: MSN) and median joining network (B: MJN) for Sabah orang-utan (LKWS and Warren datasets).
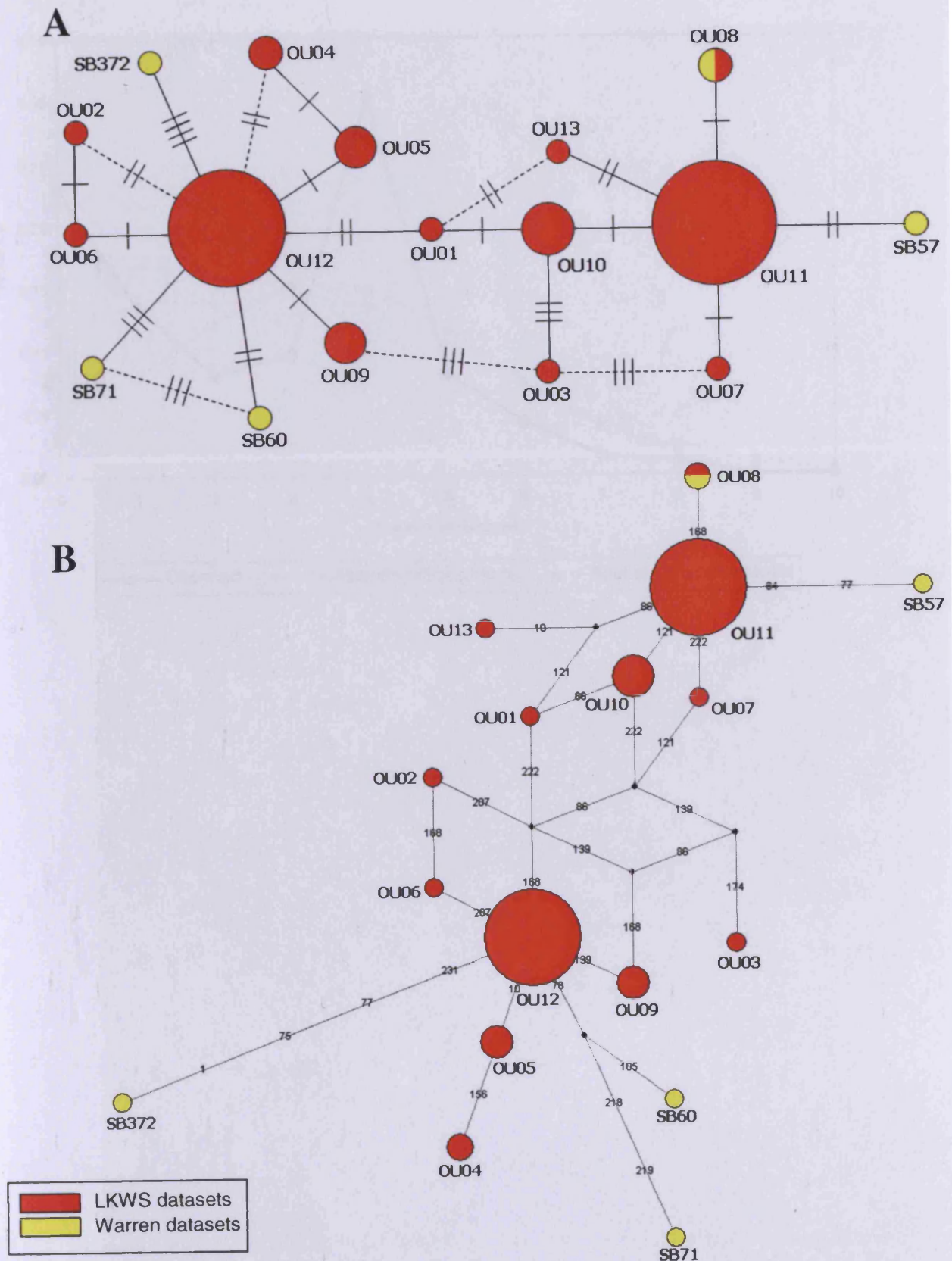
Figure 4.5. Mismatch distribution for Sabah orang-utan (LKWS and five Warren samples). Solid line represents the observed and dashed lines represent the expected for each model.
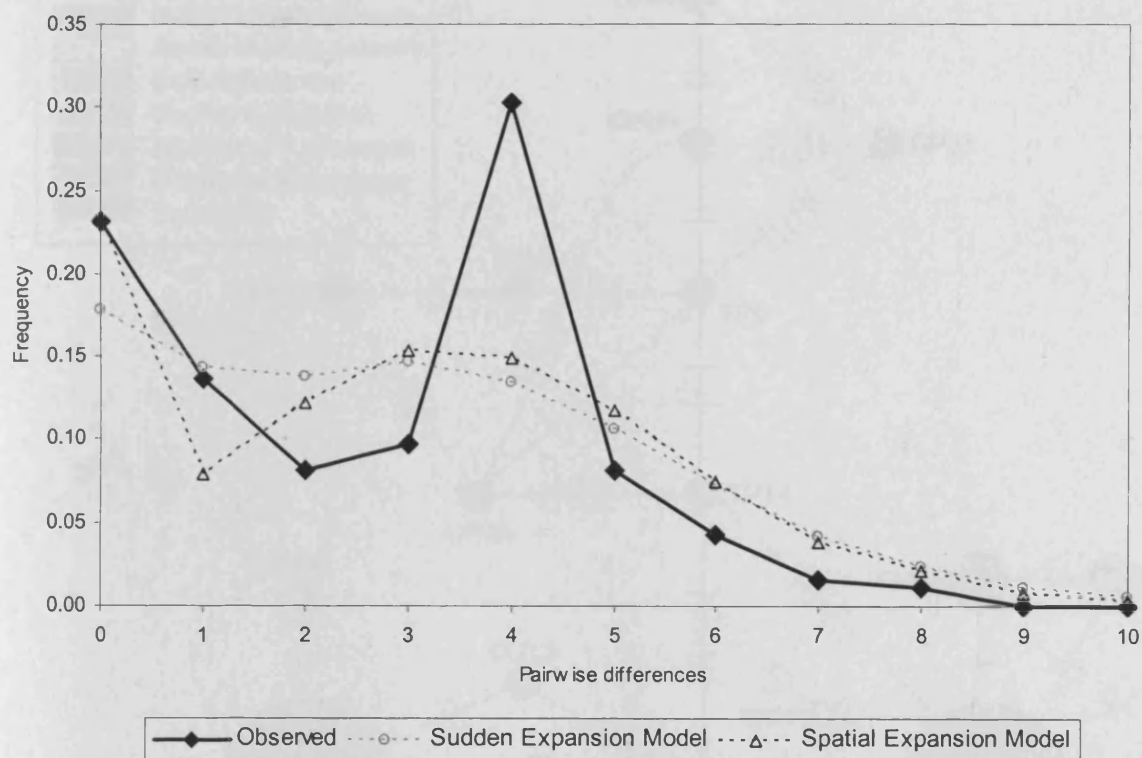
Figure 4.6. Minimum Spanning Network of Bornean orang-utan samples. Each circle represents a different haplotype. The diameter of the circle represents the frequency of each haplotype with the smallest indicating a singleton. Each bar represents a mutational step and dotted lines indicate an alternative connection.
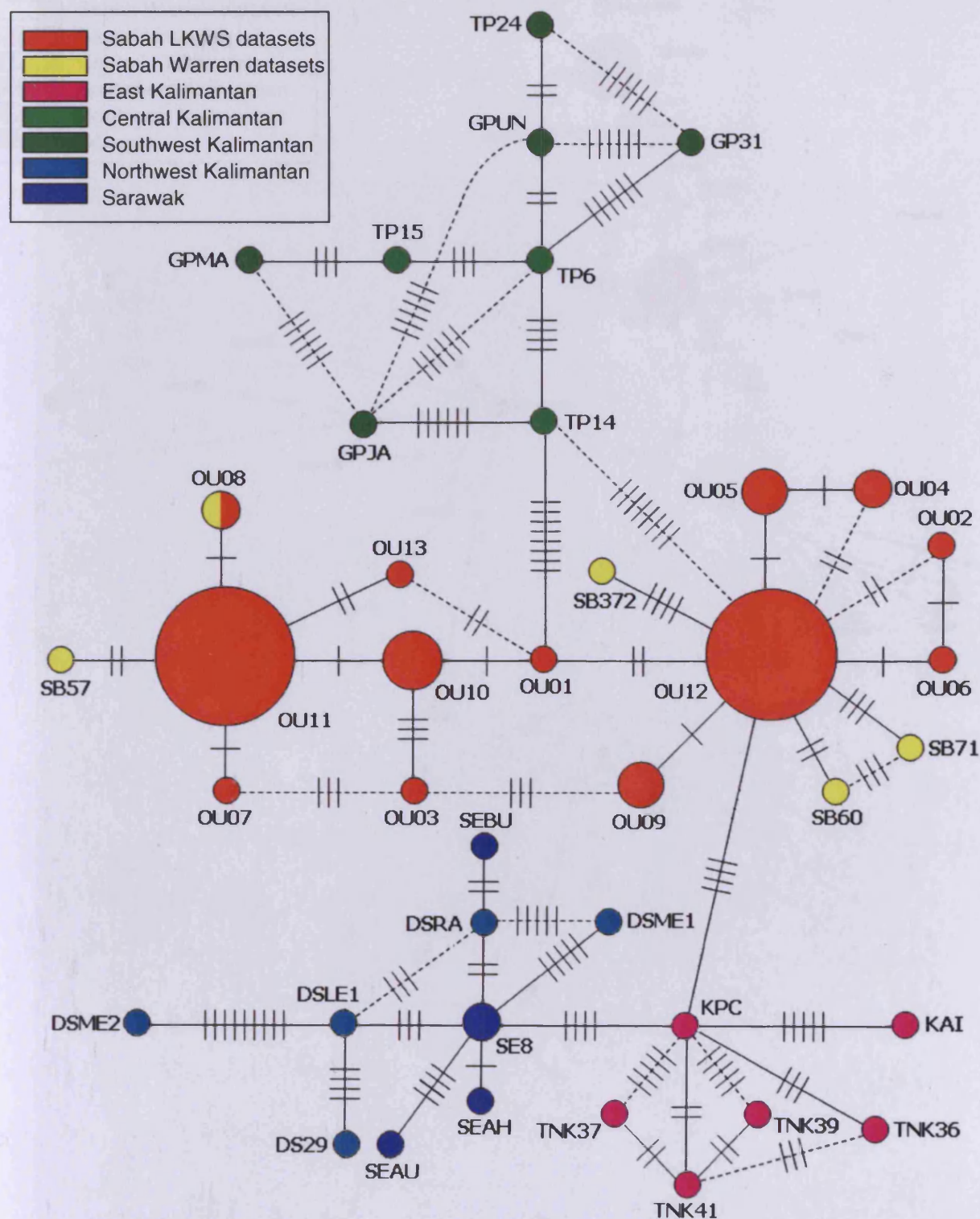
Figure 4.7. Median Joining Network of Bornean orang-utan samples. Black dots are median vectors presumed unsampled or missing intermediates. The number on the line indicates the site which has undergone mutations.

Legend:
- Sabah LKWS datasets
- Sabah Warren datasets
- East Kalimantan
- Central Kalimantan
- Southwest Kalimantan
- Northwest Kalimantan
- Sarawak
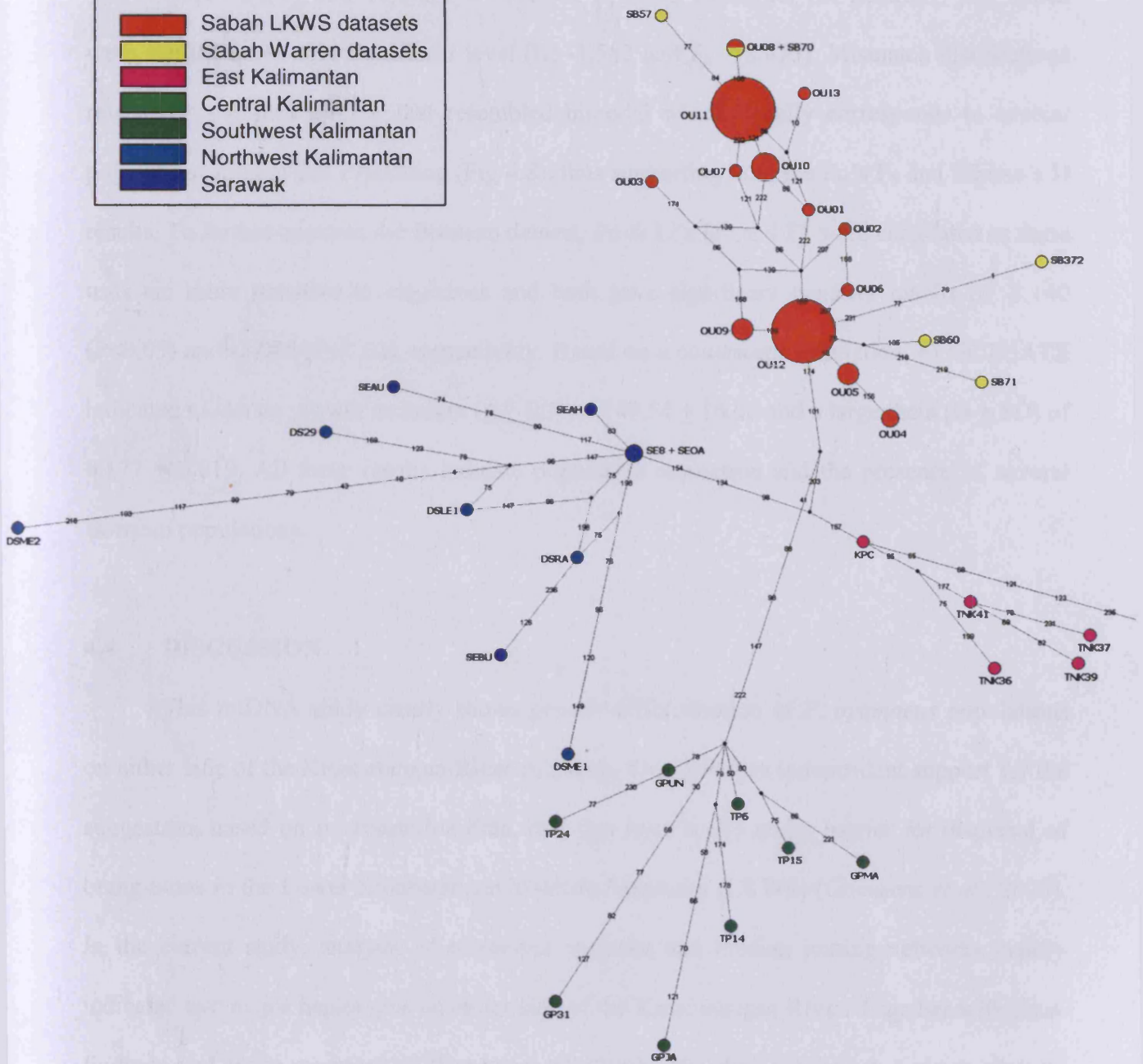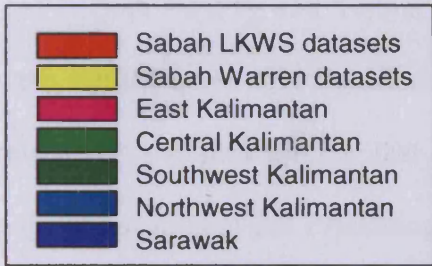
Historical demography

Both Fu's F_S and Tajima's D revealed negative values for the neutrality test which were significant at 95% significant level (D: -1.582 and F_S -18.973). Mismatch distributions revealed a complex pattern that resembled bimodal which usually corresponds to several populations mixing and expanding (Fig 4.8), thus supporting both the Fu's F_S and Tajima's D results. To further examine the Bornean dataset, Fu & Li's D* and F* were calculated as these tests are more sensitive to singletons and both gave significant negative results of -3.140 (P<0.05) and -3.046 (P<0.02), respectively. Based on a coalescent simulation, FLUCTUATE indicated moderate growth estimates (g ± SD) of 147.54 ± 16.96 and a large theta (Θ ± SD) of 0.177 ± 0.019. All these results indicate population expansion and the presence of several Bornean populations.

4.4   DISCUSSION

This mtDNA study clearly shows genetic differentiation of *P. pygmaeus* populations on either side of the Kinabatangan River in Sabah. This provides independent support for the suggestion, based on microsatellite data, that this river is the major barrier for dispersal of orang-utans in the Lower Kinabatangan Wildlife Sanctuary (LKWS) (Goossens *et al.*, 2005). In the current study, analysis of molecular variance and median joining networks clearly indicated two major haplotypes on either side of the Kinabatangan River. Together with these findings and those reported by Warren *et al.* (2001), the data indicate that rivers play an important role in shaping the genetic structure of Bornean orang-utans. The low nucleotide and high haplotype diversity exhibited by LKWS orang-utans suggests population expansion by a few founder lineages. The two haplotypes (OU11 and OU12) exhibiting the highest frequencies are most likely to represent the co-founders of the current population. This interpretation was further supported when Sabah samples from Warren *et al.* (2001) were included in the analysis as these also clustered with OU11 and OU12. The almost star-shaped

Figure 4.8. Mismatch distribution for Bornean orang-utan (LKWS and Warren). The solid line represents the observed and dashed lines represent the expected for each models.
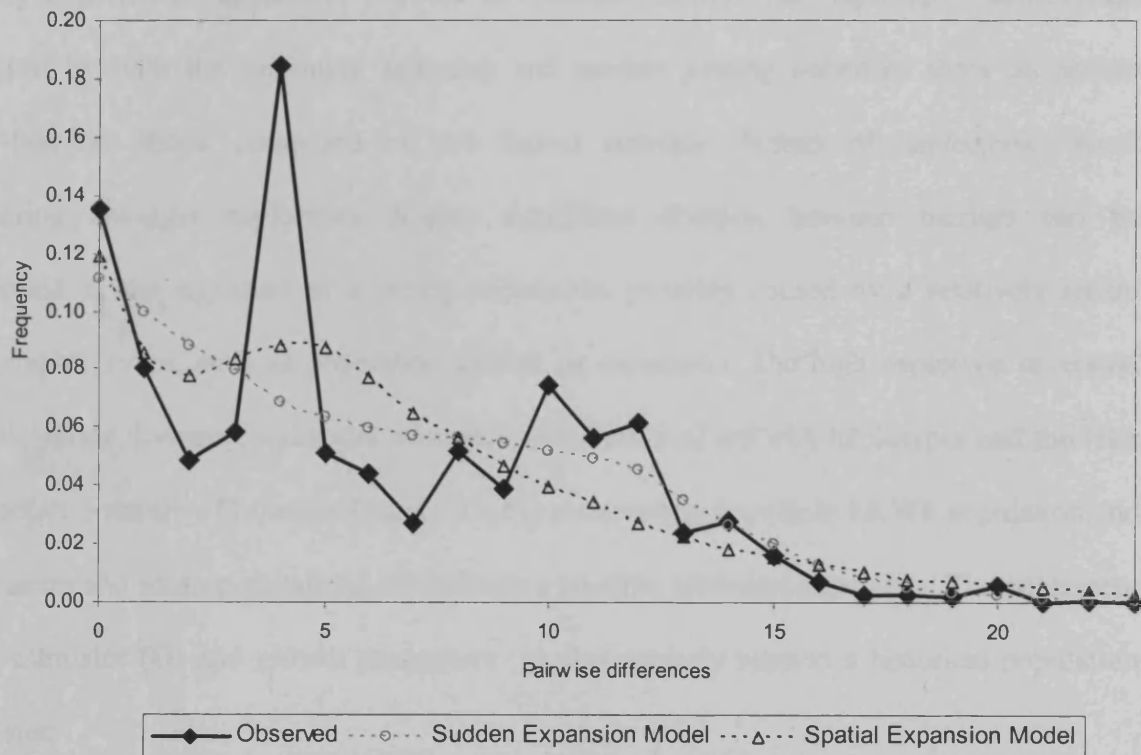
phylogeny with OU11 and OU12 surrounded with rare haplotypes differentiated by single mutational steps implies the potential ancestral lineages or founders (OU11 and OU12) recently experienced expansion (Posada & Crandall, 2001). The haplotypic relationships portrayed by both the minimum spanning and median joining networks show an almost dumb-bell-like shape composed of two linked star-like clusters of haplotypes. Weak structuring amongst haplotypes despite significant division between barriers can be interpreted as the signature of a young population, possibly caused by a relatively recent demographic event, such as population growth or expansion. The high haplotype diversity, low nucleotide diversity, unimodal mismatch distribution of mtDNA haplotypes and the tests of selective neutrality (Tajima's D and Fu's F_S) observed in the whole LKWS population and in the north and south populations, all indicate a possible historical expansion. The coalescent based estimator (Θ) and growth parameters (g) also strongly support a historical population expansion.

The present population of orang-utans in the LKWS is small, about 1000 individuals (Ancrenaz et al., 2004), and we have sequenced 73 individuals, about 7% of the total population. However, out of these 7%, we identified 13 haplotypes of which 12 were new to Sabah. Warren et al. (2001) reported only five haplotypes from their Sabah study and only one of these matched our haplotypes. Amongst the few studies on orang-utans, our work identifies the largest number of haplotypes from one study site (27,000 ha). We also found evidence that the population of orang-utans in Borneo is not static but expanding after a series of bottlenecks. Zhi et al. (1996) found only 9 haplotypes using 16s RNA and they also reported complete isolation between Sumatran and Bornean orang-utans. However, they also reported that there is very little differentiation within the Bornean orang-utan and, contrary to our data, found no evidence for past bottlenecks or founder effects (Zhi et al., 1996). Muir et al. (2000) identified three orang-utan lineages, two in Sumatra and one in Borneo. The latter

consisted of just four cytochrome *b* and five NADH haplotypes. In contrast to Zhi *et al.*'s (1996) finding of complete dichotomy between Sumatran and Bornean orang-utans, Muir *et al.* (2000) found one Sumatran individual with a haplotype that was identical to the most widespread Bornean haplotype. However, both of these studies were based on limited geographic coverage. Warren *et al.* (2001) had access to all subspecies from Borneo and Sumatra, and using the control region, identified 37 haplotypes from 41 Borneo individuals that clustered into four distinct subpopulations. These were distinct from four haplotypes found from five Sumatran individuals (Warren *et al.* 2001). The current study is the only comprehensive population genetic study using mtDNA of orang-utans in a single, localised area. Despite the restricted area (27,000 ha), we show that orang-utans in the Kinabatangan retain high levels of mtDNA genetic diversity. This complements the findings of Goossens *et al.* (2005, 2006a) who identified high levels of nuclear genetic diversity in the same orang-utan populations.

During the early and middle Pleistocene, orang-utans were widely distributed throughout mainland South East Asia (Harrison *et al.*, 2006). When the Sunda Shelf was exposed, orang-utans colonised Sumatra, Borneo and Java. Even though a land bridge existed between mainland South East Asia and its islands, the movement of orang-utans might have been disrupted by a drier arid landscape of seasonal woodlands and grasslands that bordered the eastern edge of the Malay Peninsula and continued onto the lowland areas between Sumatra and Borneo, through southern Kalimantan and eastern Java and the Lesser Sunda Islands (Bird *et al.*, 2005; Harrison *et al.*, 2006). This almost completely isolated the orang-utan populations within Borneo, probably contributing to their morphological and molecular separation from Sumatran orang-utans, where there is no current evidence of population subdivision (Harrison *et al.*, 2006). Brandon-Jones (1998) described a severe glacial drought around 190,000 years ago (YA) and a second less severe drought 80,000 YA. Contraction and

expansion of rainforest distribution has been the prime mediator of extant primate distribution in Borneo (Brandon-Jones, 1998). The first severe glaciation might have caused a population decline and fragmentation of orang-utan populations due to contraction of the forest habitats. Orang-utans were previously considered highly dependent on primary forest (Delgado & van Schaik, 2000) and being solitary animals, the effects of fragmentation and drought during glaciation might have resulted in local extinctions. Recently, Ancrenaz *et al.* (2004) suggested that orang-utan populations could survive in small degraded forest based on the finding that within a restricted area (Kinabatangan) these primates can utilize a wide range of food items, from fruits and leaves to insects and bird eggs. This might explain why the contraction of rainforest during the last glaciation did not wipe out all orang-utan populations but some might have survived in rainforest remnants (Brandon-Jones, 1998). However, the climatic change brought by glaciation might have further isolated the remnants of these populations thus severing potential gene flow which might have caused allopatric speciation (Slatkin, 1987).

Based on fossil records, the Kinabatangan area is a relatively recent habitat (Noad, 2001), which was probably colonized by orang-utans from Mount Kinabalu (see Appendix I), a known glacial refugia for many species, including orchids (Barkman & Simpson, 2001), termites (Gathorne-Hardy *et al.*, 2002) and oaks (Cannon & Manos, 2003). Mount Kinabalu lies to the west of the Kinabatangan headwaters, therefore assuming orang-utans colonized from this refugium, we predicted greater genetic diversity of orang-utans in the west compared to the east. Assuming the Kinabatangan populations were founded by very few lineages, then a small number of haplotypes would dominate (Lawler *et al.*, 1995) as detected in the current study. Haplotypes OU11 and OU12 were the most common haplotypes, each dominating one river side, and each being more common in the east compared to the west (see Figure 4.9). This process of colonization is likely to have been slow considering that orang-

utans have a long life span, very low fecundity and occur at low densities (Degado & van Schaik, 2000). Recently, Goossens *et al.* (2006b) showed that the orang-utan in Kinabatangan were highly philopatric, this further supports the hypothesis of only a few founders colonizing Kinabatangan.

Figure 4.9. The possible sequential event of lineage sorting in orang-utan populations on the Northern and Southern sides of the Kinabatangan River.
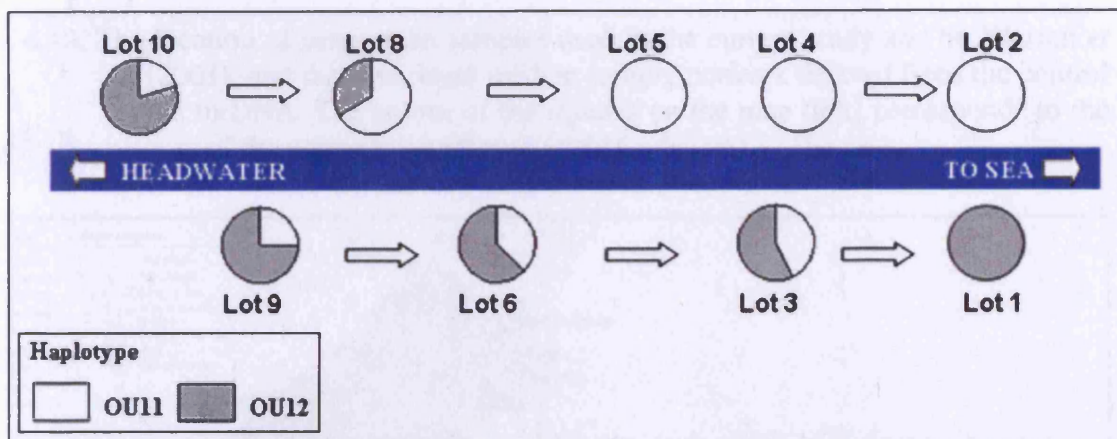
The Bornean orang-utan samples cluster into four monophyletic groups on the median joining network. Mapping the locality of samples used for genetic analyses reveals the possible separation effects of five major rivers, the Rajang, Kapuas, Barito, Mahakam and Kayan (Fig. 4.10). Warren *et al.* (2001) hinted that geographic barriers might be responsible for isolating the four separate subpopulations of Bornean orang-utans. As described above for Kinabatangan, the other three haplogroups (east Kalimantan, Sarawak & northwest Kalimantan, central and southwest Kalimantan) might have originated in a similar fashion during the last glaciation. The mountain range on which Mount Kinabalu is located, *i.e.* the Crocker Range, is actually considered to consist of several separate glacial refugia (Tanaka *et al.*, 2001). In addition to Mount Kinabalu in the north, there is considered one refugium in the east and another in the west (Gathorne-Hardy *et al.*, 2002) from which the different orang-utan haplogroups might have originated. Such divergence within a species has also been shown in another species of great ape, the bonobo (*Pan paniscus*), which occurs within the

Democratic Republic of Congo (Eriksson *et al.*, 2004). Bonobos show high genetic diversity within the Congo across their entire distribution range and can be loosely divided into two major clades (Eriksson *et al.*, 2004). The greatest differentiation was observed between samples in the east and the north, and between the north and northeast, populations which are divided by Congo and Lomami rivers, respectively. Eriksson *et al.* (2004) concluded that this pattern was indicative of the rivers hindering gene flow within this species.

Figure. 4.10. The location of orang-utan samples used in the current study and by Warren *et al.* (2001), and the associated median joining network derived from the control region mtDNA. The colour of the squares on the map (left) corresponds to the colours of the network haplotypes (right).
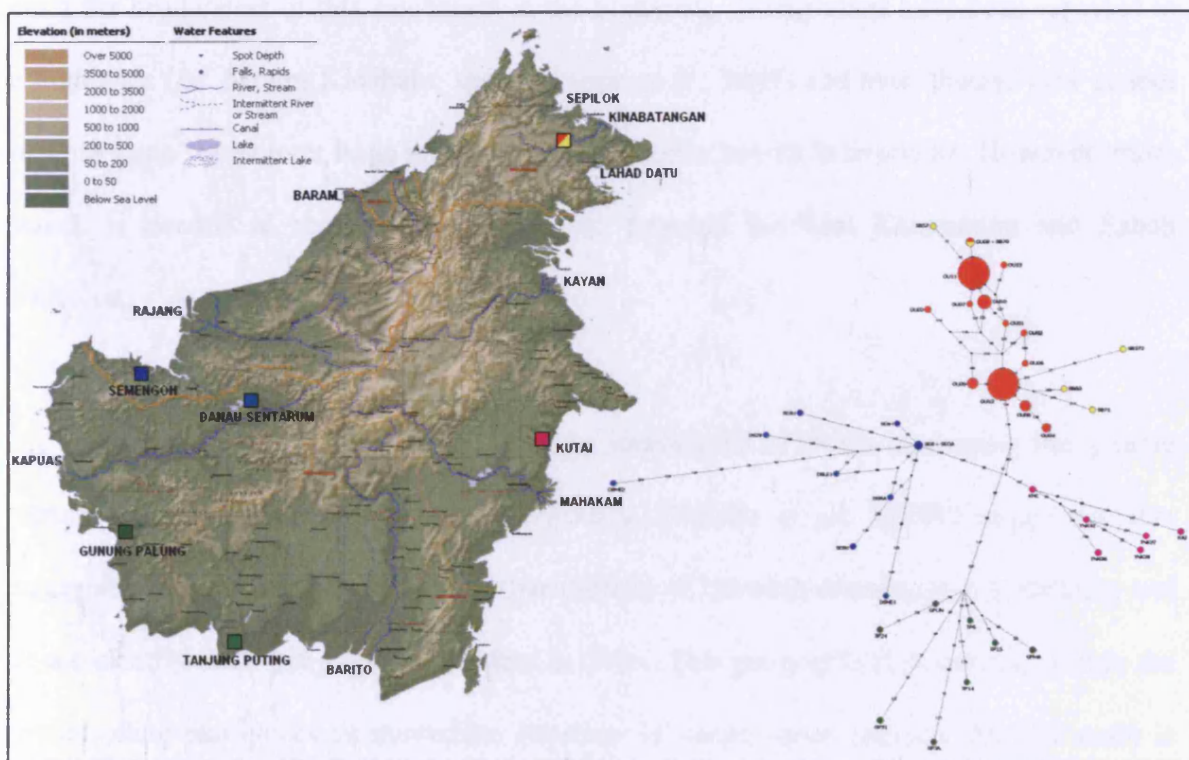
Groves (1986, 2001), Groves *et al.* (1992) and Uchida (1998) suggested that Bornean orang-utans could be separated into three subspecies based on cranial and dental (post canine) morphologies, (i) *P. p. morio* found in Sabah and East Kalimantan, (ii) *P. p. pygmaeus* found in Sarawak and northwest Kalimantan, and (iii) *P. p. wurmbii* found in central and southwest Kalimantan. Only recently has this hypothesis of three subspecies been accepted (Goossens *et al.*, 2005; Harrison *et al.*, 2006), and in fact Warren *et al.* (2001) and the current study suggest

that there may even be four subspecies. However, the current data can not be explained by Groves' (1986, 2001) morphological subspecies because our Sabah and East Kalimantan clades do not form reciprocally monophyletic groups, unlike the Sarawak and northeast Kalimantan (*P. p. pygmaeus*) and central and southwest Kalimantan (*P. p. wurmbii*) which do correspond to subspecies groupings. Instead, the divergence between the Sabah and East Kalimantan could be attributed to the Kayan River that separates the population. However, this explanation might appear less likely if more samples are obtained south of Kinabatangan and north of Kayan River. Unlike the other three great rivers of Kalimantan (Kapuas, Barito and Mahakam), the Kayan River is relatively short and there is a possibility of gene flow around the headwaters of this catchment in the highlands. Orang-utans have been reported in the highlands (*i.e.* Mount Kinabalu; see Ancrenaz *et al.*, 2005) and even though they cannot swim, they can travel over large areas providing suitable habitat is available. However, more research is needed to resolve the relationship between the East Kalimantan and Sabah subspecies.

The current study has demonstrated the importance of rivers in shaping the genetic structure of orang-utan populations. Previously, Warren *et al.* (2001) suggested that geographical barriers are responsible for partitioning of Bornean orang-utan populations, and here we identify these geographical barriers as rivers. This geographical structuring within the Bornean orang-utan poses an immediate question for conservation projects. Further study is needed to explore the extent of geographical structuring in Bornean orang-utans to detect distinct populations for conservation purposes. Populations that are highly divergent must be protected to safeguard the genetic diversity of the dwindling Bornean orang-utan, especially in Kalimantan. Further research with a larger coverage of samples and different molecular markers (*i.e.* Y-chromosomes; microsatellites see Kanthaswamy & Smith, 2002) could provide alternative views on population genetics of this species. In the current study and of

Goossens *et al.* (2005, 2006a,b), which utilized the same DNA extracts, we have demonstrated unequivocally that non-invasive samples yield high enough DNA quality, without the need for invasive samples (see Kanthaswamy *et al.*, 2006), for genetic studies of endangered species, such as the orang-utan.

CHAPTER 5

GENETIC STRUCTURE AND DISPERSAL IN A PROBOSCIS MONKEY (*NASALIS LARVATUS*) POPULATION IN SABAH, MALAYSIA.

Jalil, M.F. [1,2], Goossens, B. [1,2], Cable, J. [1], Sinyor, J. [3] and Bruford, M.W. [1]

[1] Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3TL, UK.

[2] Institut Biologi Tropika & Pemuliharaan, Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah, Malaysia.

[3] Kinabatangan Orangutan Conservation Project, P.O.Box 3109, 90734 Sandakan, Sabah, Malaysia.

Abstract

To investigate the population structure and demographic history of the proboscis monkey in northern Borneo, we examined variation in the mitochondrial DNA control region from 133 faecal samples obtained from the Lower Kinabatangan Wildlife Sanctuary and Labuk Bay, Sabah. We found very high haplotype diversity and moderate to high nucleotide diversity in this population, and identified three distinct mitochondrial lineages. However, there was no support for geographic partitioning between these lineages and no indication of any form of restricted gene flow across the Kinabatangan River, in accordance with the observed strong swimming behaviour of this species and in direct contrast to other primates inhabiting the Kinabtangan, for example the Bornean orang-utan. However, evidence for demographic expansion of more than one population probably reflects post-glacial refugial colonization of the Kinabatangan.

5.1. INTRODUCTION

The Sunda Shelf experienced several major episodes of climatic change in the late Pliocene and early Pleistocene which had a dramatic impact on the flora and fauna of the region (van de Bergh *et al.*, 2001; Bird *et al.*, 2005). Several authors have hypothesized that an 'arid corridor' was present across the continent of Sundaland during the last glaciation

(Voris, 2000; Bird *et al.*, 2005). This provided a potential route for fauna (e.g. orang-utans; Harrison *et al.*, 2006) from continental Asia to disperse to the islands (such as Java; van de Bergh *et al.*, 2001), but also served as a barrier for forest dependent species, such as some primates (Brandon-Jones, 1996, 1998). Meijaard (2003) suggested that during the last glacial maximum (LGM) several areas remained forest-covered which included West Sumatra, North-west Borneo, the Malacca Straits and the region around Palawan Island. This refugium hypothesis is supported by the disjunct and restricted distributions of several primate taxa in both Sumatra and Borneo (Brandon-Jones, 1996, 1998). Recently, Tanaka *et al.* (2001) suggested that the mountain range on which Mount Kinabalu is located, *i.e.* Crocker Range may comprise several separate glacial refugia (see Appendix I). In addition to Mount Kinabalu in the north, a second refuge may lie to the east and a third in the west (Gathorne-Hardy *et al.*, 2002).

Within this context, the genus *Nasalis* is a severely range restricted, monotypic genus, now found only in Borneo with no evidence that its distribution ever extended beyond Sundaland (Brandon-Jones, 1996, 1998; Harcourt & Schwartz, 2001). Sterner *et al.* (2006) suggested that ancestral Asian colobines split from African colobines around 10.8 million years ago (MYA) and began to diversify around 6.7 MYA. This corresponds to Harrison *et al.*'s (2006) suggestion that the ancestor for both the proboscis monkey (*Nasalis larvatus*) and its closest relative, the Simakobu monkey (*Simias concolor*) arrived on the Sunda Islands during the early Pliocene (Sterner *et al.*, 2006). During the late Pliocene, several cold phases led to fragmentation of tropical and subtropical rainforest creating refugia separated by open vegetation (Meijaard & Groves, 2006). As both species are forest dependent (Brandon-Jones, 1996) this contraction of suitable habitat during the LGM is likely to have reduced their population sizes, thus establishing their endemism by the late Pliocene (Harisson *et al.*, 2006). Subsequently, early Pleistocene glaciation and climate change may have formed an arid

barrier between Sumatra and Borneo preventing these primates from recolonising their previous habitat (Brandon-Jones, 1996; Bird *et al.*, 2005).

Traditionally, proboscis monkeys have been 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). These primates are highly folivorous, but can survive on high or low quality food resources at different times of the year allowing them to exploit seasonal rainforest where food is more abundant than open woodlands (Bennett & Davis, 1994; Brandon-Jones, 1996; Agoramoorthy & Hsu, 2005a). Thus, the proboscis monkey, like the other odd-nose genera *Simias* and *Pygathrix*, has evolved an adaptive foraging strategy that enables them to thrive in seasonal forest (Bennett & Davies, 1994; Sterner *et al.*, 2006).

The proboscis monkey, known locally as the bangkatan or Dutchman monkey, occurs 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 occurs quite frequently and is usually more common in subadult females than adult females (Murai *et al.*, 2006). Boonratana (2000) reported that the home range size of the proboscis monkey in Sukau, Kinabatangan, is approximately 220 ha which is much smaller than in Samunsam, Sarawak (900 ha) but larger than in Tanjung Putting, Kalimantan (137 ha). The home range size of the proboscis monkey is influenced by the availability of food items. The large home-range in Samunsam probably compensates for lack of suitable resources thus the animals forage over larger distances (Bennet & Davies, 1994). The poor diversity of trees in Samunsam (~10 genera), compared to

Kinabatangan and Tanjung Putting, probably accounts for the four fold increase in size of home range.

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 (Fleagle, 1998; Oates *et al.*, 1994). According to the Aquatic Ape Hypothesis (AAH) (Morgan, 1999), ancestral humans may have evolved aquatic adaptation to survive living in wet/semi-aquatic environments. As proboscis monkeys are most at home in the water, the ancestral proboscis monkey might have lived in similar semi-aquatic environments, however the AAH remains an extremely controversial idea.

Relatively few genetic studies have been carried out on colobines (i.e. *Trachypithecus* spp., Rosenblum *et al.*, 1997; *Rhinopithecus roxellana,* Li *et al.*, 2003) and there have been no previous molecular studies on *N. larvatus*. This is the first study of population genetics and demography of wild proboscis monkeys in the fragmented forest of the Lower Kinabatangan in Sabah, Malaysia. It is predicted that this endemic Bornean colobine will display relatively high genetic diversity and no population structure within Kinabatangan due to their social structure and swimming ability.

5.2.   MATERIAL AND METHODS

Sampling

Faecal samples (n = 132) were collected along the Kinabatangan River within the Lower Kinabatangan Wildlife Sanctuary (LKWS), Sabah, Malaysia (Fig 5.1) in 2003. An additional sample was collected in Labuk Bay, north of the Kinabatangan basin in the same

Figure 5.1. Distribution of proboscis monkey faecal samples (green squares) collected within the Lower Kinabatangan Wildlife Sanctuary, Sabah.
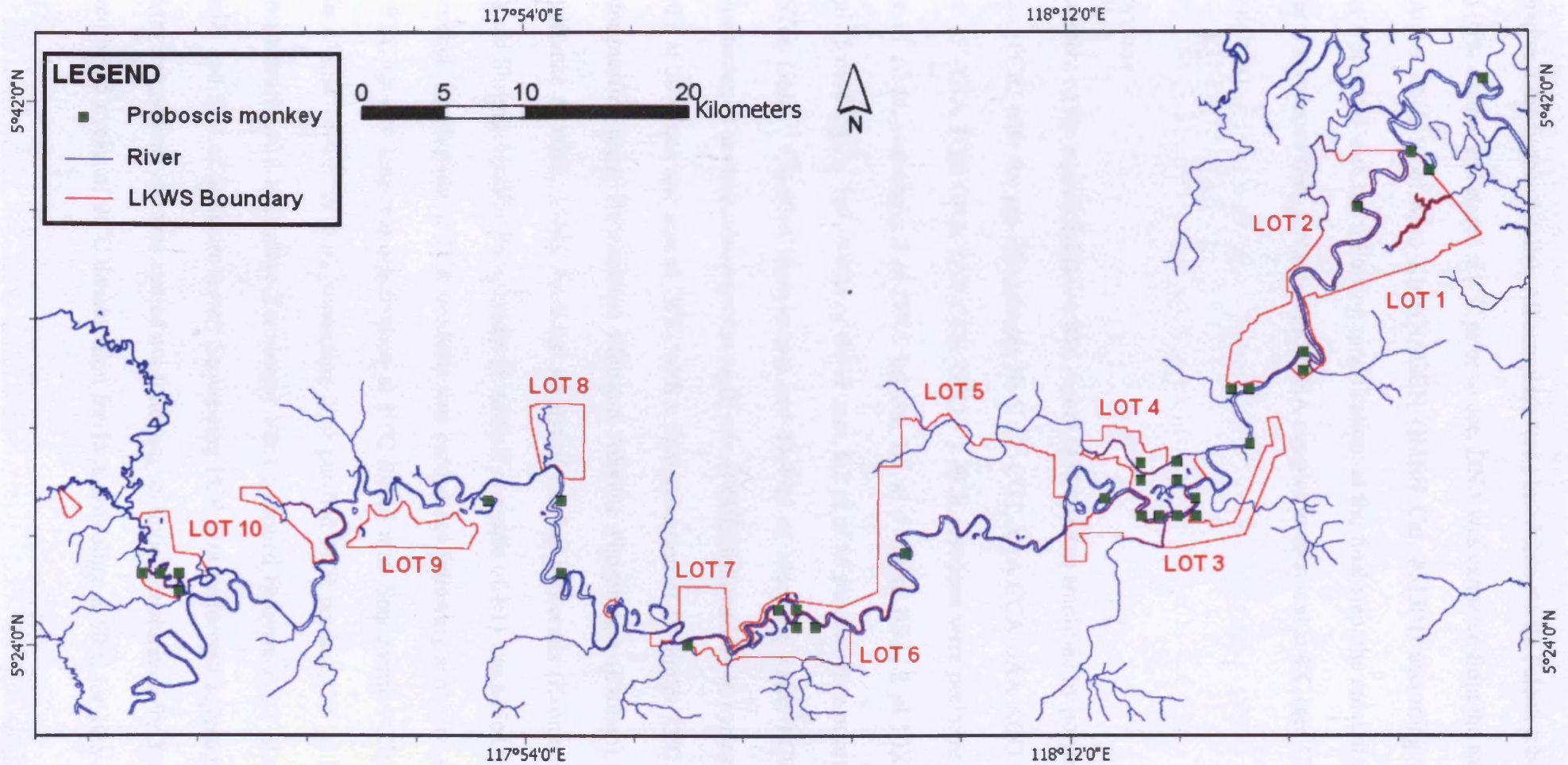
year (see Chapter 2 for further details). All samples were placed into sterile Falcon tubes and preserved in 95% ethanol and stored at 4°C prior to use. DNA was extracted from the samples using a QIAamp DNA Stool Mini Kit (QIAGEN GMBH Cat. #51504) according to the manufacturer's protocol with the following modification: at the final step, the amount of AE elution buffer was reduced from 200μl to 70μl. DNA samples were stored at 4°C (see Chapter 2 for further details).

DNA amplification

A fragment of the mitochondrial control region (434 bp) was amplified by polymerase chain reaction (PCR) with the specific primers Nl-5' (5'-CGT AAA CCA GAA ACG GAT-3') and Nl-3' (5'-TAA TGG GAA TAT CCG TGC-3'). PCR reactions were performed in a final volume of 20 μl, containing 2 μl DNA extract, 1.5 μl 4 mg/ml BSA, 2 μl 10X PCR Buffer, 1.5 μl 25 mM MgCl$_2$, 1μl 10 mM of dNTP mix, 0.2 μl of 50 pmol of each primer, 0.2 μl of AmpliTaq Gold^{TM} (Applied Biosystems) and 11.4 μl of ultrapure water. PCR was carried out consisting of an initial denaturation of 12 min at 94°C, followed by 40 cycles of 30 s at 94°C, 30 s at 50°C and one min at 72°C, with a final extension of 10 min at 72°C. PCR products were purified using Exonuclease I/Shrimp Alkaline Phosphatase (ExoSap) (USB Corp, USA) (Hanke & Wink, 1994). An 0.5μl aliquot of ExoSap mastermix (Exonuclease I (10 units/μl) and Shrimp Alkaline Phosphatase (1 unit/μl) in a ratio of 1:1) was added to each 5 μl PCR product. Purification of PCR products was carried out following activation of the enzymes at 37°C for one hour and deactivation at 80°C for 15 min. Sequencing PCRs were performed in a final volume of 8 μl, containing 2 μl purified PCR product, 2.5 μl Better Buffer (Web Scientific), 0.5 μl BigDye Terminator Ver.1 (Applied Biosystems), 1 μl of 1.6 pmol of primer and 2 μl of ultrapure water. Sequencing PCR was performed separately for forward and reverse primers and was carried out following an initial denaturation for 3 min at 96°C followed by 25 cycles of 96°C denaturation for 15 s, annealing at 50°C for 10 s and 2

min of extension at 60°C. Sequencing PCR was performed using the ABI Big Dye Terminator vs. 1 (Applied Biosystems). Each PCR product was sequenced in both directions. PCR products were precipitated by adding 90 μl of 63% isopropanol to each PCR tube. The product and isopropanol were vortexed for 20 s, left to stand for 15 min, then centrifuged for 30 min at 13,000 g. The supernatant was discarded and 150 μl of 70% isopropanol was added to each PCR tube. The PCR strips were then centrifuged for one minute at 500 g and dried at 52°C for 2 min. Sequencing was performed in an ABI3100 automated sequencer.

Mitochondrial DNA analysis

MtDNA control region sequences were aligned using the program SEQUENCHER 3.1.2 (GeneCodes Corp.) with correction by eye. The program DAMBE 4.2.13 (Xia & Xie, 2001) was used to search the total data alignment for unique haplotypes. These haplotypes were assembled into a separate data set and used for all subsequent analyses. Genetic diversity was measured for the entire sample based on haplotype diversity (h) and nucleotide diversity (π). Haplotype diversity was defined as the probability that two randomly chosen sequences from a sample were different and nucleotide diversity was the average number of nucleotide differences per site between two randomly chosen sequences (Nei, 1987). These diversity indices were computed with the software ARLEQUIN 3 (Excoffier *et al.*, 2005). Patterns of mtDNA genetic structure between riversides of LKWS were examined using an analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992) as implemented in ARLEQUIN. The significance of the genetic structure within the whole population (Φ_{ST}) was assessed by 1000 permutation tests implemented in ARLEQUIN. To investigate the relationship between haplotypes, minimum spanning and median joining networks were constructed using ARLEQUIN and NETWORK 4.1.1.1 (Bandelt *et al.*, 1999; available at http://www.fluxus-engineering.com/), respectively. Tests for selective neutrality, Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) which can independently assess the demographic

trajectories of populations (specifically, signatures of population expansion) were performed using ARLEQUIN. A mismatch distribution was constructed to further test for demographic expansion using ARLEQUIN. A coalescent-based method to test for evidence of population expansion was also carried out using FLUCTUATE v. 1.4 (Kuhner *et al.*, 1998). The programme used a maximum likelihood approach based on coalescent theory to simultaneously estimate theta (Θ) and population growth rate (g). The programme was run 5 times to ensure convergence of the estimates.

5.3.   RESULTS

In the 434 bp fragment of the mtDNA control region from the proboscis monkey, alignment of 133 sequences revealed 45 variable nucleotide positions (Table 5.1). Of the 44 LKWS haplotypes identified, 17 were found at multiple sites and 27 were identified only in individual Lots. The site with most haplotypes recorded was Lot 4 and PM33 was the most commonly found haplotype in LKWS. Lot 4 also contained the highest number of private haplotypes (9) followed by Lot 10 (5) and Lot 5 (4), whereas only two private haplotypes were recorded in Lots 1, 3, 6 and 8. No transversions were observed amongst the proboscis monkey sequences (Table 5.2). The sample from Labuk Bay contained a unique haplotype. No insertions or deletion were observed in any of the samples.

Nucleotide diversity (π) was 0.016 $\pm$ 0.008 across all LKWS Lots, but higher for samples from the northern compared to southern side of the river (Table 5.3). Haplotype diversity (*h*) was also highest on the northern side of the river (Table 5.3). The minimum spanning and median joining networks indicate three distinct groups of proboscis monkey mtDNA sequences in the LKWS. However, the groups are not clearly separated according to riverside (Figs. 5.2 & 5.3). This observation was supported by AMOVA, which indicated an extremely high proportion of variation residing within populations (99.8%), but only 0.2%

Table 5.1. Variable nucleotide positions in the alignment of mtDNA sequences of the proboscis monkey control region. Dots indicate positions where the bases are identical to haplotype PM01.

```
         1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4
         2 9 5 7 7 0 0 1 3 4 4 4 6 7 7 7 8 8 8 9 9 0 1 1 1 2 3 4 4 4 5 5 5 5 6 6 7 8 8 9 0 1 2 3 3
         6 7 6 0 2 0 5 4 1 2 5 8 9 3 8 9 2 6 8 3 6 3 1 3 4 6 1 4 5 9 3 4 5 9 5 7 6 7 9 9 5 0 9 1 3
PM01  G G G C G T T A T C A C G T A G G C G C T C C T C A A C C C G C C T T A T G T G C T A C A
PM02  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A . . . . .
PM03  . A . . . . . C T G . A . G A A . . . . . . . . G . . . . . T . . . . C A . . . . . . . .
PM04  . . . . . . . C T G . A . G . A . . . . . . T G . . T . A . T . . . . . . . . . . . . . .
PM05  . . . . . . . C T G . A . G . A . . . . . . T G . . . . T . . . . . A . . . . . . . . . .
PM06  . . . . . . . C T G . A . G . A . . . . . . T . . . . . T . . . . . A . . . C . . . .
PM07  . . . . . . . C T G . A . G . A . . . . . . T G G . . . A . T . . . . . . . . . . . . . .
PM08  . . . . . . . C T G . A . G . A . . . . . . . G . . . . T . . . . . A . A . . . . .
PM09  A . . . . C C G C T . . A C . . . . . . C . . . T G . . . . T . . . . . . . . . . . . . .
PM10  . . A . . . . C T G . A . G . A . . . . . . . G . . . . T . . . . . A . A . . . . .
PM11  . . . . . . . C T . . A . G . A . . . . . C T G . . . . T . . . . . A C . . . . . .
PM12  . . . . . . . C T G . A . G . A . . . . T . . G . . . . T . . . . . A . . . . . . . .
PM13  . . . . . . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . . . . . . . . . .
PM14  A . . . . . C G C T . . A . . . . . A . C . . . T G . . . . T . . . . . . . . . . . . .
PM15  A . . . . C G C T . . A . . . . . . C . . . T G . . . . T . . . . . . . . T . . . .
PM16  . . . . . . . . . . . . . . . . T . . . . . . G . . . . . . . . . . . . . . . . . . . .
PM17  . . . . . . . C T G . A . . . A . . . . . . T G . . . . . T . . . . . . . . . . . . .
PM18  . . . . . . . C T . . . . . . T . . . . . . T G . T . . . T . . . . . . . . . . . . .
PM19  . . . . . . . C T G . A . G . A . . . . . . . G . . . . A . T . . . . . . . . . . . . .
PM20  . . . . . . . . T . . . . . . . . . T . . . . G . . . . . . . . . . . . . . . . . . . .
PM21  A . . . . C G C T . . A . . . . . . C . . . T G . . . . T . . . . . A . . T . . . .
PM22  . . . . . . . C T G . A . . . A . . . . . . T G . . . . A . T . . G . . . . . . . . . .
PM23  . . . . . . . C T G . A . G . A . . . . . . T G . . . . . . . . . . . . . . . . G . .
PM24  . . . . A . . . C T . . . . . . . T . . . . T G . . . . . . . . . . . . . . . . . . T .
PM25  . . . . . . . C T G . A . G . . . . . . . . . G . . . . T . . . . . A . . . . . . . .
PM26  . . . . . . . C T G . A . . . A . . T . . . . T G . . . . . T . . . . . . . . . . . . G
PM27  . . . T . . . C T G . A . G . A . . . . . . T G . . . . T . . . . . . . . . . . . . . .
PM28  . . . . . . . C T G . A . . . A . . . . . . T G . . . . T T . . . . . . . . . . . . .
PM29  . . . . . . . C . G . A . G . A . . . . . . T G . . . . T . . . . . . . . . . . . . . .
PM30  A . . . . C G C T . . A . G . . . . . C . . . T G . . . . T . . . . . . . . . . . . .
PM31  . . . . . . . C T G . A . G . A . . . . . . T G . . . A . T . . . . . . . . . . . . .
PM32  A . . . . C G C T . . A . . . . . . C . . . T A . . . . T . . . . . A . . . . . .
PM33  . . . . . . . C T G . A . . . A . . . . . . T G . . . A . T . . . . . . . . . . . . .
PM34  A . . . . C G C T . . A . . . . . . C . . . T G . . . T . T . . . . . . . . . . . . .
PM35  . . . . . . . C T . . . . . . T . . . . . . T G . T . . . T . . . . . . A . . . . . .
PM36  . . . . . . . C T G . A . G . A . . . . . . T G . . . . T . . C . . . . . . . . . . .
PM37  . . . . . . . . . . . . . . . . . . . . . . . G . . . . . T . . . . . . . . . . . . .
PM38  . . . . . . . C T G . A . G . A . . . . . . . G . . . . T . . . . . A . . . . . . . .
PM39  A . . . . C G C T . . A . . . . . . C . . . T G . . . . T . . . . . . . . . . . . .
PM40  A . . . . C G C . . . A . . . . . . C . . . T G . . . . T * . . . . . . . . . . . . .
PM41  A . . . . C G C T . . A . . . . . . C . . . T G . . . . T . C . . . A . . . . . . .
PM42  . . . . . . . C T G . A . G . A . . . . . . T G . . . . T . . . . . . . . . . . . .
PM43  . . . . . . . C T . . . . . . . . . . . . . T G . . . . T . . . . . . . . . . . . .
PM44  . . . . . . . . . T . . . . . . . . . . . . T G . . . . T . . . . . . . . . . . . .
LABUK A . . . A . C G C T . . A . G . . . . . C . . . T G . . . . . T . . . . . . . . . . . . .
```

Table 5.2. Frequency of each mitochondrial DNA control region haplotype of proboscis monkey within the Lower Kinabatangan Wildlife Sanctuary. (L=Lot; BD=Balad Dami)

HAPLOTYPE	L 1	L 2	L 3	L 4	L 5	L 6	L 8	L 9	L 10	BD
PM01	-	-	-	1	-	-	-	-	-	-
PM02	-	-	-	1	-	-	-	-	-	-
PM03	-	-	-	1	-	-	-	-	-	-
PM04	-	-	-	-	-	1	-	-	-	-
PM05	-	-	-	-	-	-	-	-	-	2
PM06	-	-	-	-	-	-	-	-	1	-
PM07	-	-	-	1	-	-	-	-	-	-
PM08	-	-	-	3	-	-	-	-	-	-
PM09	-	-	-	-	-	-	1	-	-	-
PM10	-	-	-	1	-	-	-	-	-	-
PM11	-	-	-	-	-	-	-	-	1	-
PM12	-	-	1	-	-	-	-	-	-	-
PM13	1	-	3	2	-	-	-	-	-	-
PM14	-	-	-	-	-	-	-	-	2	-
PM15	-	-	2	-	-	-	-	-	-	-
PM16	-	-	-	1	-	-	-	-	-	-
PM17	-	-	-	1	-	-	-	-	-	-
PM18	-	-	1	-	1	-	-	-	-	-
PM19	-	-	-	-	-	1	-	-	-	-
PM20	-	1	1	-	-	-	-	-	-	-
PM21	1	-	-	-	-	-	-	-	-	-
PM22	-	-	-	-	1	-	-	-	-	-
PM23	4	2	-	-	-	-	-	-	-	-
PM24	-	-	1	3	-	-	-	-	-	-
PM25	1	-	-	-	-	-	-	-	-	-
PM26	-	-	-	-	4	-	-	-	-	-
PM27	-	-	-	-	1	-	-	-	-	-
PM28	3	-	2	-	-	-	-	-	2	-
PM29	-	-	-	-	-	-	-	-	1	-
PM30	1	2	-	3	-	-	-	-	-	-
PM31	-	-	-	1	1	2	1	-	-	-
PM32	1	2	-	1	1	1	-	-	-	-
PM33	4	3	-	1	-	1	1	1	-	-
PM34	-	-	-	2	2	-	-	-	-	-
PM35	-	-	-	-	1	-	-	-	-	-
PM36	-	-	-	2	-	2	- •	-	-	-
PM37	-	-	-	1	-	-	-	-	1	-
PM38	-	-	-	1	-	1	-	-	-	-
PM39	-	-	2	4	3	3	-	-	-	-
PM40	-	-	-	-	-	-	2	-	-	-
PM41	-	-	-	3	-	-	-	-	-	-
PM42	2	-	1	4	-	-	2	-	1	-
PM43	-	-	-	-	-	-	-	-	1	-
PM44	1	-	2	-	-	-	-	-	2	-
NO. OF SEQUENCES	19	10	16	38	15	12	7	1	12	2

Figure 5.2. Minimum spanning network (MSN) of proboscis monkey mtDNA control region haplotypes. Each circle represents a different haplotype and the diameter indicates haplotype frequency. The smallest circle represents a singleton. Mutational steps are represented by black bars on lines connecting haplotypes.
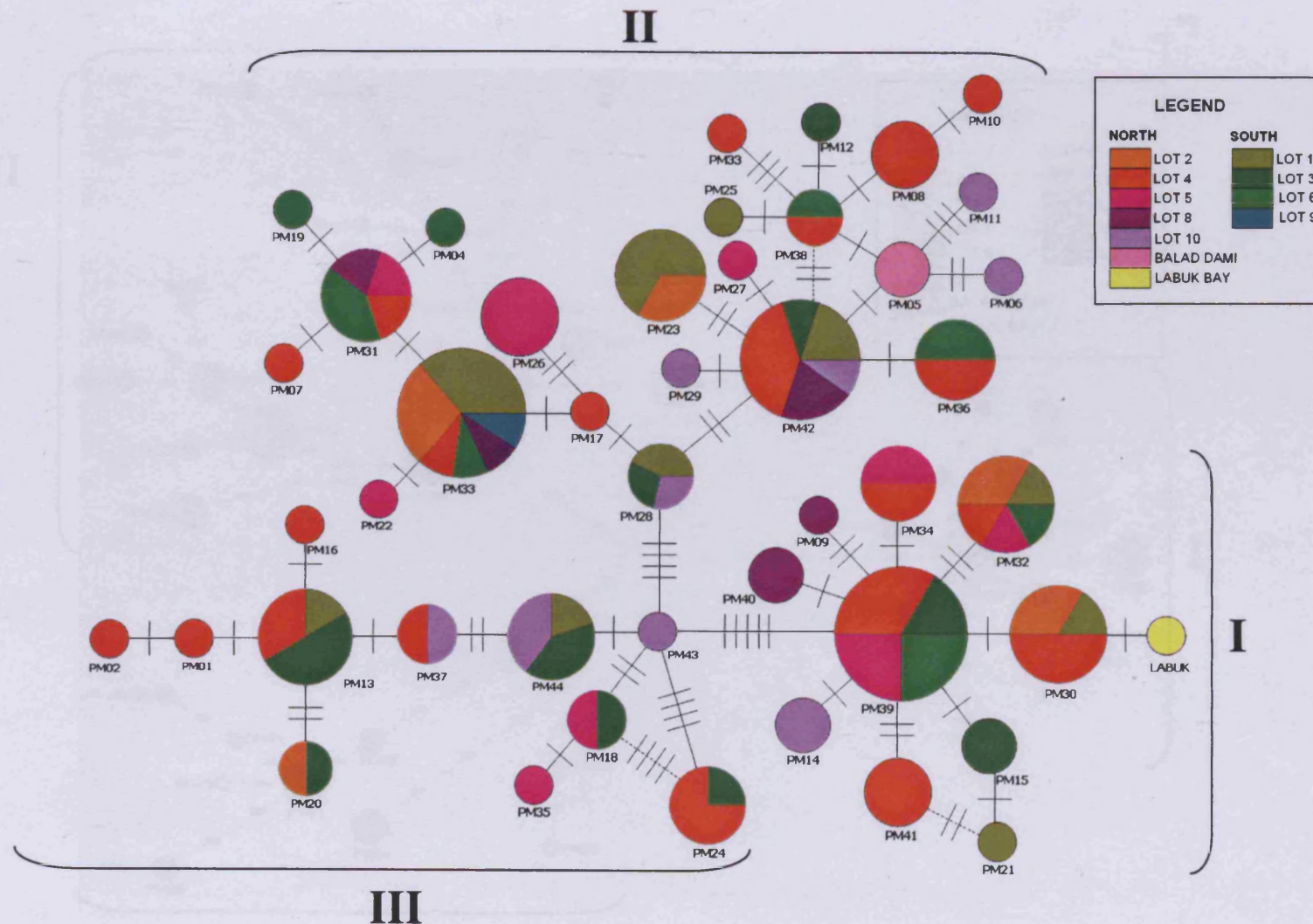
Figure 5.3. Median joining network (MJN) of proboscis monkey mtDNA control region haplotypes. Each circle represents a different haplotype and the diameter indicates haplotype frequency. The smallest circle represents a singleton. Black dots are median vectors presumed unsampled or missing intermediates and numbers indicate the sites that have undergone mutations.
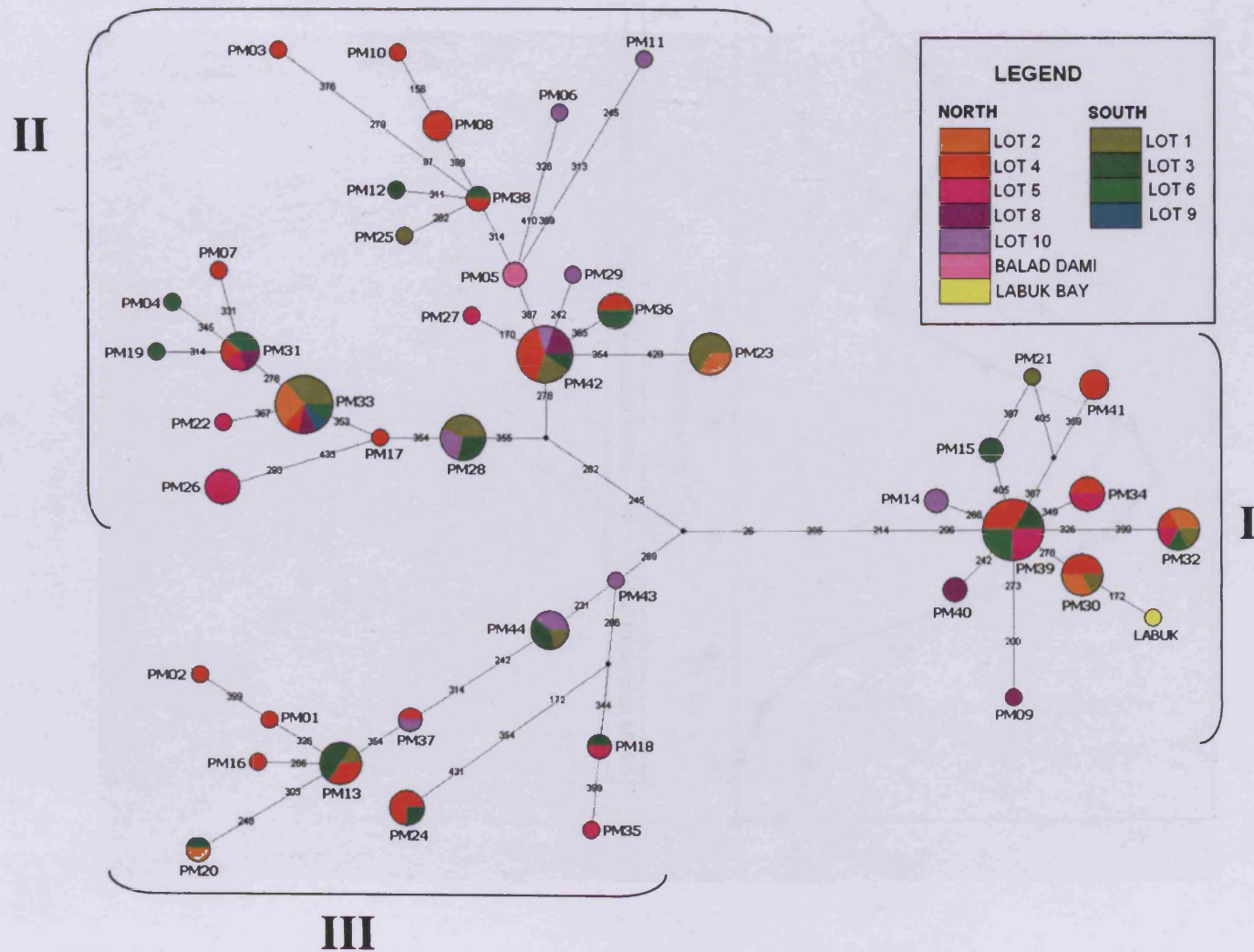
Figure 5.4. Mismatch distribution for each of the analysed Lower Kinabatangan Wildlife Sanctuary proboscis monkey samples. Observed (solid lines) and expected (discontinuous line) showing frequencies of pairwise differences.
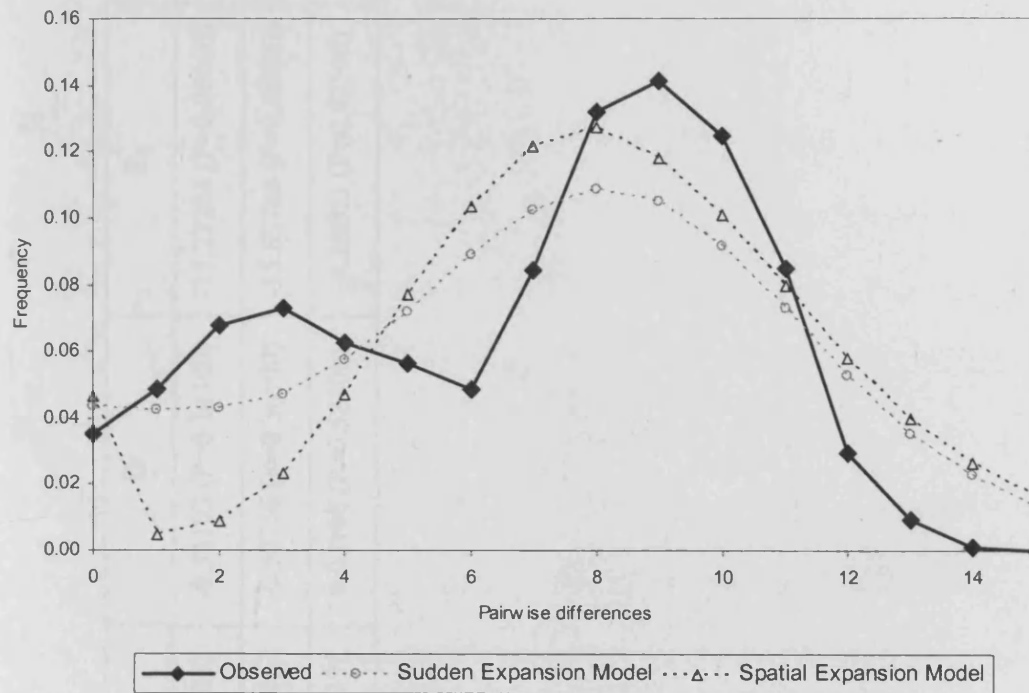
Table 5.3. Number of sequences and haplotypes, nucleotide diversity (π), haplotype diversity (h), Tajima's D and Fu's F_S and population (θ) and growth (g) parameters of proboscis monkey mitochondrial DNA control region sequences for all Lower Kinabatangan Wildlife Sanctuary Lots.

	Seq.	Hap.	π	h	D	F_S	θ	g
LKWS	132	44	0.015710 ± 0.008243	0.9647 ± 0.0052	-0.53103 (P=0.34300)	-17.22284 (P=0.00000)	0.0953 ± 0.0053	369.71 ± 31.73
NORTH	84	38	0.016277 ± 0.008551	0.9710 + 0.0063	-0.50756 (P=0.36500)	-15.69746 (P=0.00000)	-	-
SOUTH	48	21	0.014664 ± 0.007850	0.9495 + 0.0128	0.02908 (P=0.58300)	-4.25203 (P=0.07500)	-	-

variation between populations on either side of the river (P=0.000). Most of the haplotypes within each clade (I & II) were scattered throughout Kinabatangan, with the exception of the clade III which was absent from Lots 6, 8 and 9 (Figs. 5.2 & 5.3). This pattern of high haplotype and nucleotide diversity indicates either a large stable population with a long evolutionary history or secondary contact between differentiated lineages (Grant & Bowen, 1998).

Tajima's D value for the whole of LKWS population was -0.531 (Table 5.3). Only the northern side of the river showed a negative value of -0.508, the southern side of the river showed a positive value of 0.029, but none of the values were significant. In contrast, Fu's Fs revealed a highly significant value of -17.223 (P=0.001) for the whole LKWS and -15.698 (P=0.001) for the northern side of the river. The southern side of the river showed a smaller non-significant value of -4.252 (Table 5.3). Generally, it is accepted that Fu's Fs is more sensitive than Tajima's D to detect an excess of recent mutations, a pattern typical of both demographic expansion and/ or selective sweep (Fu, 1997; Ford, 2002). The mismatch distribution of all samples (Fig. 5.4) revealed a bimodal pattern, indicative of intermixed populations in Kinabatangan and demographic expansion. A coalescent approach to detect population expansion using FLUCTUATE indicated large positive estimates for the growth parameters for the Kinabatangan population (Table 5.3). Estimates of growth parameters (g $\pm$ SD) and theta (Θ $\pm$ SD) were 369.71 $\pm$ 31.73 and 0.095 $\pm$ 0.005, respectively. All these analyses concur, suggesting that the proboscis monkey population in the Lower Kinabatangan is expanding quite rapidly.

5.4   DISCUSSION

To date this is the first population genetic study on wild proboscis monkeys. The presence of 44 control region haplotypes from 133 samples is indicative of high haplotype

diversity in the proboscis monkey population of the Lower Kinabatangan Wildlife Sanctuary, Sabah. Mismatch distribution analysis revealed a bimodal pattern, indicating a mixture of two or more populations that subsequently expanded. It is clear from both haplotype networks and AMOVA that there was no geographical partitioning of haplotypes in the Kinabatangan. These results have two major implications for our understanding of Kinabatangan and Sabah proboscis monkey populations. Firstly, despite the small sampling coverage (only from Kinabatangan basin), the high level of genetic diversity among individual samples revealed a tri-phyletic pattern, suggesting a split of these three lineages or the retention of one or more ancestral haplogroups by chance (Masembe *et al.*, 2006). Secondly, the network and mismatch distribution results indicate that the current population contains a mixture of different groups indicating possible secondary contact between refugial populations in the Kinabatangan River basin. The fact that the proboscis monkey did not exhibit any degree of population genetic structuring between riversides of the Kinabatangan was not surprising as proboscis monkeys are excellent swimmers (Brandon-Jones, 1996), thus the Kinabatangan River dissecting the Lower Kinabatangan Wildlife Sanctuary did not pose a barrier for movement of this species, in contrast to other species studied in this thesis.

The apparent lack of geographically partitioning of Kinabatangan proboscis monkeys may also be related to the non-philopatric nature of the females (Murai, 2006; Murai *et al.*, 2006) indicating free movement of females. Similar trends have been observed within the colony of Cape fur seal (*Arctocephalus pusillus pusillus*) in South Africa/Namibian breeding colonies (Matthee *et al.*, 2006). Murai (2006) and Murai *et al.* (2006) studying the same population of proboscis monkey in Kinabatangan revealed the promiscuous behaviour of the females, which leave their groups to breed with other males before returning to their former groups. The peaceful nature of male proboscis monkeys probably supports this behaviour. As proboscis monkeys live in a harem, female offspring would increase the number females in

the harems, and male offspring might challenge the dominant males for its rank. Nevertheless, infanticide among wild proboscis monkeys has been observed in Labuk Bay, north of Kinabatangan (Agoramoorthy & Hsu, 2005b).

The genetic diversity of proboscis monkeys within the Kinabatangan was higher than we predicted, exceeding that of both long-tailed macaques (see Chapter 3) and orang-utans (see Chapter 4). The number of haplotypes even exceeds that recorded for Eritrean hamadryas baboon (*Papio hamadryas hamadryas*), with only 35 mtDNA control region haplotypes from 10 localities in Eritrea (Hapke *et al.*, 2001). Winney *et al.* (2004) studying the same baboon species in Saudi Arabia recorded only 26 mtDNA control region haplotypes from 107 individuals. Similar to our results, Hapke *et al.* (2001) did not find any structure within their populations. Another study on Asiatic colobine also found high diversity with little structure among the populations of *Trachypithecus cristatus* and *T. auratus* in Peninsula Malaysia, Sumatra and Java (Rosenblum *et al.*, 1997). The high genetic diversity observed in Kinabatangan proboscis monkey probably reflects the female biased dispersal and /or the large ancestral population that inhabited Borneo. The female biased and the size of the harem also contributed to the large sizes of the female lines. As mtDNA are inherited by females only, the high level of mtDNA diversity is proportioned to the large numbers of breeding females in the ancestral population before the onset of the climatic chaos. The lack of large predators in Borneo might also be one of the factors allowing for the population to retain its size. The unique diets of proboscis monkey, provides a special niche with very little competition thus enables them to secure their food source (Agoramoorthy & Hsu, 2005a).

The triphyletic lineage detected within Kinabatangan proboscis monkeys might have been the product of several secondary contacts (from several refugia) during the Plio-Pleistocene expansion and contraction of forests (Masembe *et al.,* 2006). Similar to the

current study, Masembe *et al.* (2006) also observed triphyletic lineages between African Oryx in Tanzania and Kenya and interpreted this as the effects of ancient hybridization and introgression. Birungi & Arctander (2000) suggested that the two divergent clades identified for African Kob (*Kobus kob*) were the result of range expansion and secondary contact of two ancestral lineages that evolved in allopatry.

Throughout the Miocene, Borneo was connected to mainland Asia, Sumatra and Java, (Hall, 1998) facilitating the colonisation of colobine primates (Sterner *et al.*, 2006). Asian colobines began to diversify sometime around the late Miocene and at the same time North Borneo was uplifted from the sea (Hutchison *et al.*, 2000; Noad, 2001). By the early Pliocene (~5 MYA) the Sunda Islands had obtained an essentially modern formation (Hall, 1998; Gorog *et al.*, 2004). The proboscis monkey probably became an endemic species in Borneo during the end of Pliocene (Harrison *et al.*, 2006), accounting for the lack of fossil records of this species elsewhere in Sundaland (Harcourt & Schwartz, 2001). The warmer climate together with the expanse of tropical forest during the Miocene probably facilitated expansion and colonisation of colobine primates, including the proboscis monkey, in South East Asia (Harrison *et al.*, 2006). Towards the end of Pliocene, climatic change resulted in the formation of more open and arid habitats (Meijaard & Groves, 2006). As proboscis monkeys are highly forest dependent with a restricted diet, open habitat and savanna represents a barrier for their dispersal, thus limiting their range towards the end of the Pliocene. At this time, tropical forest was limited to the highlands, i.e. Mount Kinabalu, which would have provided refugia for many forest dependent species (Tanaka *et al.*, 2001; Gathorne-Hardy *et al.*, 2002). So even though a land bridge existed between the Sunda Islands, unsuitable habitat is likely to have formed barriers for dispersal for proboscis monkeys confined to refugia (Hall, 1998). During the Pleistocene climate cycles, proboscis monkey refugia in Borneo were probably located within the mountainous areas of Western Sabah, Western-Central Borneo and Eastern

Borneo (Brandon-Jones, 1996, 1998; Gathorne-Hardy *et al.*, 2002, Gorog *et al.*, 2004) (see Fig. 5.5). Towards the end of the drought period (Brandon-Jones, 1998), the proboscis monkey populations probably began to increase in size as the forest expanded. Several previously isolated populations of proboscis monkey might have met in secondary contact zones, such as the Kinabatangan. The union of different refugial populations and the non-philopatric nature of proboscis monkeys might have facilitated interbreeding, thereby generating the current pattern of genetic diversity.

Figure 5.5. Possible routes of proboscis monkey colonization of Kinabatangan from three hypothetical Pleistocene refugia.
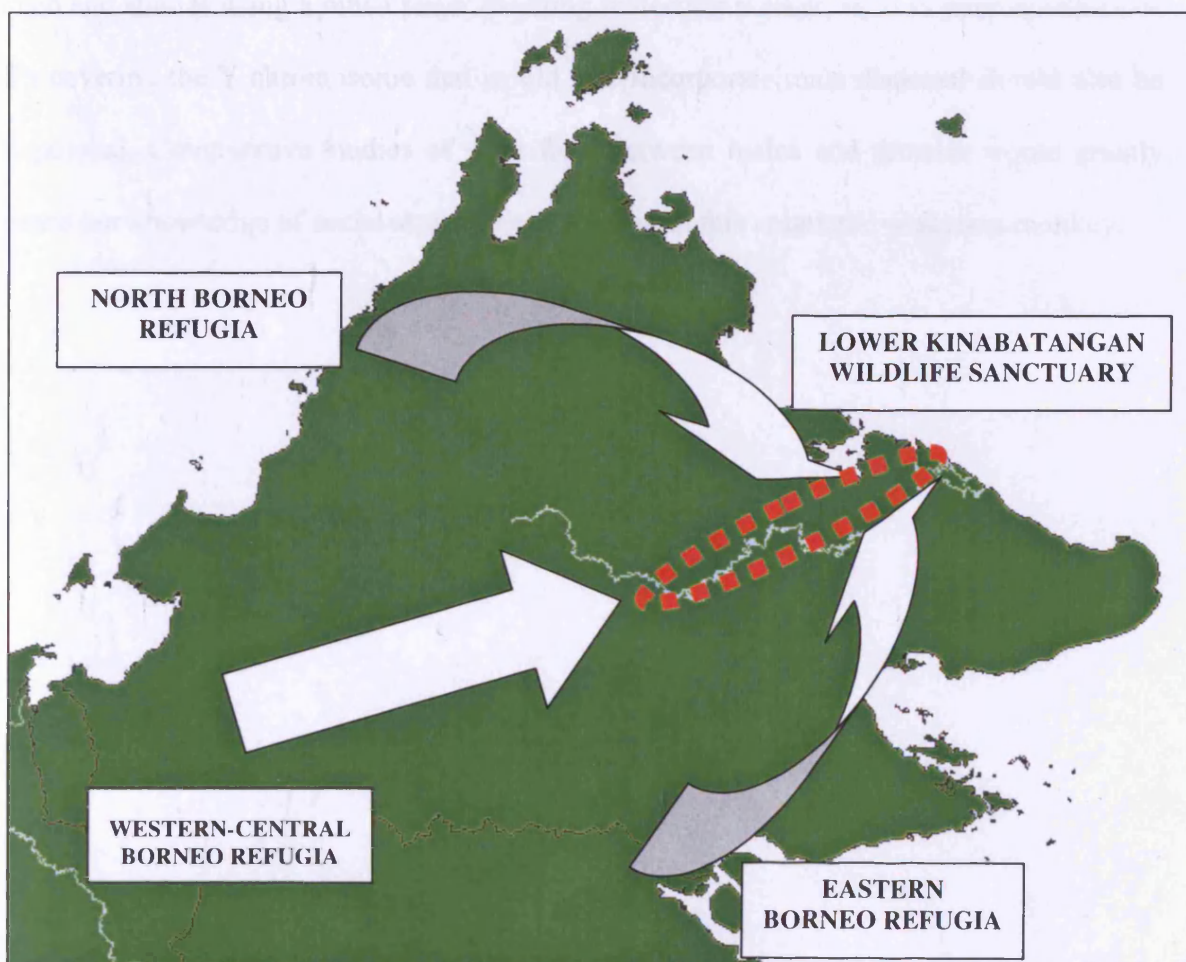
In conclusion, the current study revealed high genetic diversity in proboscis monkeys in the Kinabatangan, despite the fact that their habitat consists of fragmented, degraded forest.

Three distinct lineages were identified that are hypothesized to have evolved in allopatry in three separate glacial refugia, but having subsequently come into secondary contact following forest expansion (see Fig 5.5). Considering the Kinabatangan proboscis monkey population consists of only approximately 3,400 individuals (Goossens *et al.*, 2003) and yet still retains a very high level of genetic diversity, then the entire Borneo population is likely to be much higher. Thus, the present day proboscis monkey population in Kinabatangan is not facing immediate danger. However, this species is classified as an endangered species and pressures on habitat continue to slowly eradicate suitable habitats ideal for proboscis monkey. Ideally, future work should assess the genetic diversity of proboscis monkeys across their range in Borneo and studies using a much faster evolving molecular marker, such as microsatellites or SNPs covering the Y chromosome that would also incorporate male dispersal should also be incorporated. Comparative studies of gene flow between males and females would greatly enhance our knowledge of social structure and kinship in this enigmatic proboscis monkey.

CHAPTER 6: FINAL DISCUSSION

The current study is the first to compare phylogeographic patterns among primate taxa in Borneo in any geographic region. The species selected for this study (orang-utan, long-tailed macaque and proboscis monkey) vary in their dispersal ability, social organisation and longevity. The orang-utan is of particular interest to evolutionary biologists, being the only ape found outside Africa and, unlike other apes, it has a solitary nature (Groves, 2001). The proboscis monkey is endemic to Borneo and is a specialist feeder, living in relatively large harems. In contrast, the long-tailed macaque is widely distributed all over South East Asia, living in large hierarchical social systems and is considered a pest species due to its adaptability (Groves, 2001).

This study focussed on the Lower Kinabatangan Wildlife Sanctuary (LKWS) in Sabah, a 27,000 ha area of fragmented forested. Heavy reliance on logging (and now agriculture) has left much of the Sabah forest, including the Kinabatangan, in fragments, isolated from one another by a wasteland of scrubland, agriculture and human settlements (WWF, 1998; McMorrow & Talip, 2001; Jomo et al., 2004). However, despite the degraded nature of the habitat, the LKWS still contains a rich biodiversity and provides an ideal study site with which to assess the impact of forest fragmentation on primates (i.e. orang-utan see Goossens et al., 2005; Appendix II). In addition, 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 three primates.

The current study revealed moderate to high levels of genetic diversity for orang-utans, long-tailed macaques and proboscis monkeys. Furthermore, the Kinabatangan River does influence the population structure of the former two species. Orang-utans, due to their

inability to swim (Fleagle, 1998), exhibit the greatest genetic differences on either side of the river, whereas either smaller or no differences were detected for long-tailed macaques and proboscis monkeys, respectively. Both of these latter species are good swimmers (Fleagle, 1998), especially the proboscis monkey. This study also revealed that LKWS long-tailed macaques have experienced long term population stability, whereas both orang-utans and proboscis monkeys have experienced population bottlenecks and expansions.

In this Chapter, the observed genetic structure of three species of primates inhabiting the LKWS is discussed first in the context of the geology of the region and then in relation to conservation of the area, with recommendations given for management of the primates in LKWS. The study was based on data derived from mitochondrial DNA (mtDNA) that has a mutation rate of approximately 2% per million years in primates (Brown *et al.*, 1979). Thus, it is appropriate to begin our historical account of the chronological events that led to the current population genetic structure for each species in the Pliocene.

6.1   HISTORICAL PERSPECTIVES ON COLONIZATION

Eastern Sabah, where the Kinabatangan is located, was characterised in the early Miocene by argillaceous materials, pyroclastics and slump breccias, reflecting the unstable conditions of the deposition (Geological Survey of Malaysia, 1989; Noad, 2001). The Gomantong Limestone karst in the middle of Kinabatangan basin is of Chattian to Burdigalian age (Late Oligocene to Early Miocene) (Noad, 2001). The Upper Kinabatangan (Kuamut formation) slump deposits or melange contain fossils assemblages indicating that the deposition occurred in deep marine conditions (Geological Survey of Malaysia, 1989). Noad (2001) suggested that the Gomantong Limestone was deposited on an open shelf, where the shallower portions of the shelf were colonised by coralline red algae, forming algal flats. Subsequent limestone deposits occurred along a palaeo-shoreline running roughly east-west,

close to the Gomantong outcrops (Noad, 2001). By the Pliocene, most of Sabah had been

uplifted and rotated (Hall, 1998), and during the early Pleistocene, Kinabatangan was lifted

completely from the sea.

The Asian colobines are far more diverse than their African counterparts (Davis &

Oates, 1994), the two lineages split during the late Miocene and spread from Africa to

continental Asia (Sterner *et al.*, 2006). In Asia, the ancestral colobine began to diversify

around 6.7 million years ago (MYA) and at the same time, the ancestral proboscis monkey

began to colonize Borneo via the Malay Peninsula (see Fig 6.1A; Harrison *et al.*, 2006;

Meijaard & Groves, 2006).

Figure 6.1. The land distribution (A) during the late Miocene - early Pliocene, (B) late
Pliocene to Early Pleistocene, (C) Early to Middle Pleistocene, and (D) Middle
Pleistocene during periods of high sea level. Adapted from Meijaard & Groves
(2004).
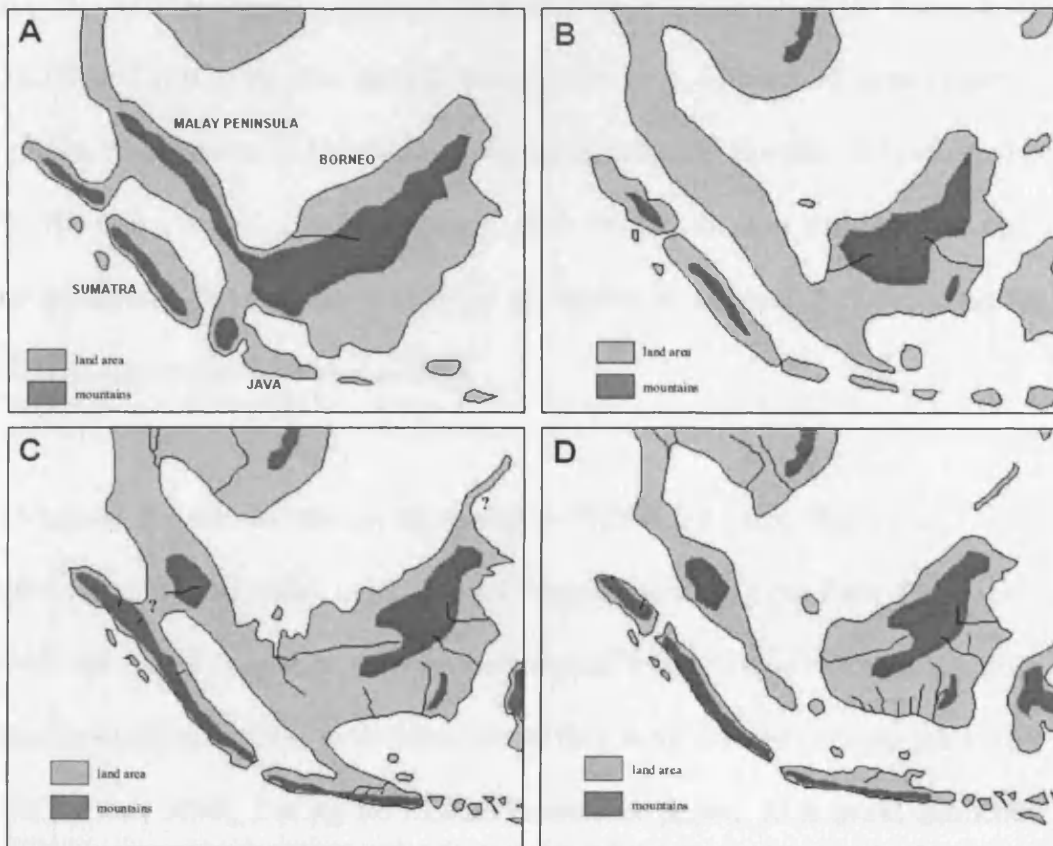
The beginning of the Pliocene was marked by warm and humid climates, and tropical and subtropical vegetation occurred all over South East Asia (Meijaard & Groves, 2006). The expanse of tropical habitats facilitated the expansion of colobine monkeys, which are specialized feeders (Yeager *et al.*, 1997; Wasserman & Chapman, 2003; Chapman *et al.*, 2006). As the climate became warmer, sea level rose by 20 - 60 m, separating Borneo from the Malay Peninsula. This effectively marooned the proboscis monkey on the island of Borneo. Towards the beginning of the Middle Pliocene, global climate changes initiated aridification and savannah habitats started to expand (Harrison *et al.*, 2006; Meijaard & Grooves, 2006).

The Late Pliocene began with several cold phases leading to more open vegetation. This isolated tropical and subtropical forest, creating refugia for forest dependent species (Meijaard & Groves, 2006). Towards the Early Pleistocene, with decreasing temperature, sea level dropped (120 - 135 m) exposing the land bridges between the islands of the Sunda Shelf again (Figs. 6.1B and 6.1C). At this period, orang-utans from continental Asia began to colonize the Sunda Island and could be found in Sumatra, Java and Borneo (Delgado & van Schaik, 2000; Harrisons *et al.*, 2006). However, their movement was restricted as open vegetation and savanna of the Late Pliocene acted as barriers to dispersal for colonizing the entire Sunda Shelf islands (Harrison *et al.*, 2006).

In Sundaland, Pleistocene human expansion (~70,000 YA, Excoffier *et al.*, 1992) probably caused extinctions of many large animals. Vegetation during the Early Pleistocene began to thin out and mixed vegetation with isolated tropical forest refugia was common (Fig. 6.2). Orang-utan populations started to differentiate as they were isolated on their respective islands (see Muir *et al.*, 2000). During the Middle Pleistocene period, widespread extinction

occurred throughout South East Asia (van de Bergh *et al.*, 2001; Tougard, 2001; Meijaard, 2003).

Figure 6.2 Land distributions in the Last Glacial Maxima where the data indicate that more open vegetation types may have existed. Adapted from Meijaard (2003).
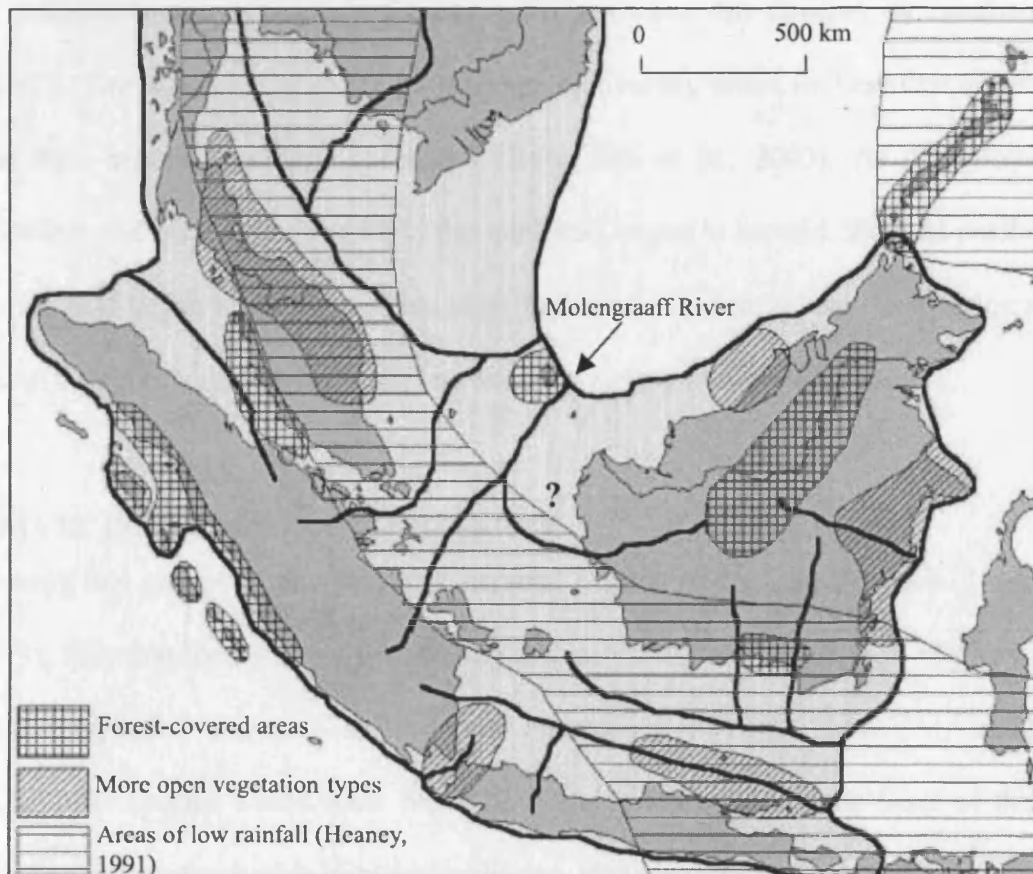
The Late Pleistocene was characterized by seasonal forest and open vegetation. This further isolated the tropical forest refugia (Meijaard, 2003; Harrison *et al.*, 2006). Eruption of Mount Toba in North Sumatra (74,000 YA), the second largest volcanic eruption producing 10 billion metric tons of ash resulted in a global drop of temperature contributing to the Wisconsin Glacial Epoch (Ambrose, 1998; Muir *et al.*, 2000). Gathorne-Hardy & Harcourt-Smith (2003) found no evidence of the Toba eruption causing a bottleneck of Mentawai Island primates. Ambrose (2003) suggested that the Mentawai primates were saved due to the downwind location of the Island from the volcanic eruption.

During the Late Pleistocene to Early Holocene, savanna persisted, thus forming a long-term barrier between the east (Sumatra) and west (Borneo) of Sundaland (Meijaard, 2003; Birds *et al.*, 2005). The separation was further enforced by several large rivers isolating the two sides. Even though the savanna persisted for thousands of years separating the two sides, West Sumatra and a large part of Eastern Borneo were still covered by rainforest (Meijaard, 2003). The high level of endemicity and plant diversity found in these two regions suggests that they served as refugia (Meijaard, 2003; Slik *et al.*, 2003). As the climate continued to warm and increase in humidity, the rainforest began to expand. Isolated patches of rainforest refugial began to merge to form large expanses of forest, where the meeting of animals from different refugia resulted in hybrid zones or zones of secondary contact.

6.2 PRIMATE POPULATION DEMOGRAPHY

Assuming that proboscis monkeys first invaded Borneo in the Late Pliocene (2.6-1.8 MYA; Fig 6.3), followed by the orang-utan during the early Pleistocene (1.8-0.78 MYA; Fig 6.3) and the long-tailed macaque in the late Pleistocene (0.13-0.012 MYA), extant populations of these species would most likely have stable structures on the basis of their early colonization during favourable climate conditions. However, the current study indicated that both the proboscis monkey and orang-utan have experienced bottlenecks and subsequent population expansion, and only the long-tailed macaque population has been stable. This difference is probably explained by the ecology of the animals.

Proboscis monkeys live in large harems, have a high reproduction rate related to the high female to male ratio and have excellent dispersal capabilities (Fleagle, 1998; Meijaard & Nijman, 2000). However, this is counteracted by their specific dietary requirements unlike the orang-utan which is omnivorous (Yeager *et al.*, 1997; Delgado & van Schaik, 2000). The orang-utan is a solitary species, with a low reproduction rate and especially females have limited dispersal capabilities (Rijksen & Meijaard, 1999; Delgado & van Schaik, 2000).

Long-tailed macaques, on the other hand, live in multimale-multifemale groups with a high reproduction rate (Fleagle, 1998). They have good dispersal capabilities, in fact Abegg & Thierry (2002) suggested that long-tailed macaques recolonised the islands of Sundaland by rafting. Unlike the proboscis monkey, the long-tailed macaque is not limited to specific diets but is an opportunistic feeder able to utilise most resources available, hence Harcourt & Schwartz (2001) described this primate as an excellent survivor candidate.

Figure 6.3. The proposed sequential events of colonization by the two primate species from Pliocene to early Pleistocene. (Red dot - proboscis monkey, green squares – orang-utan).
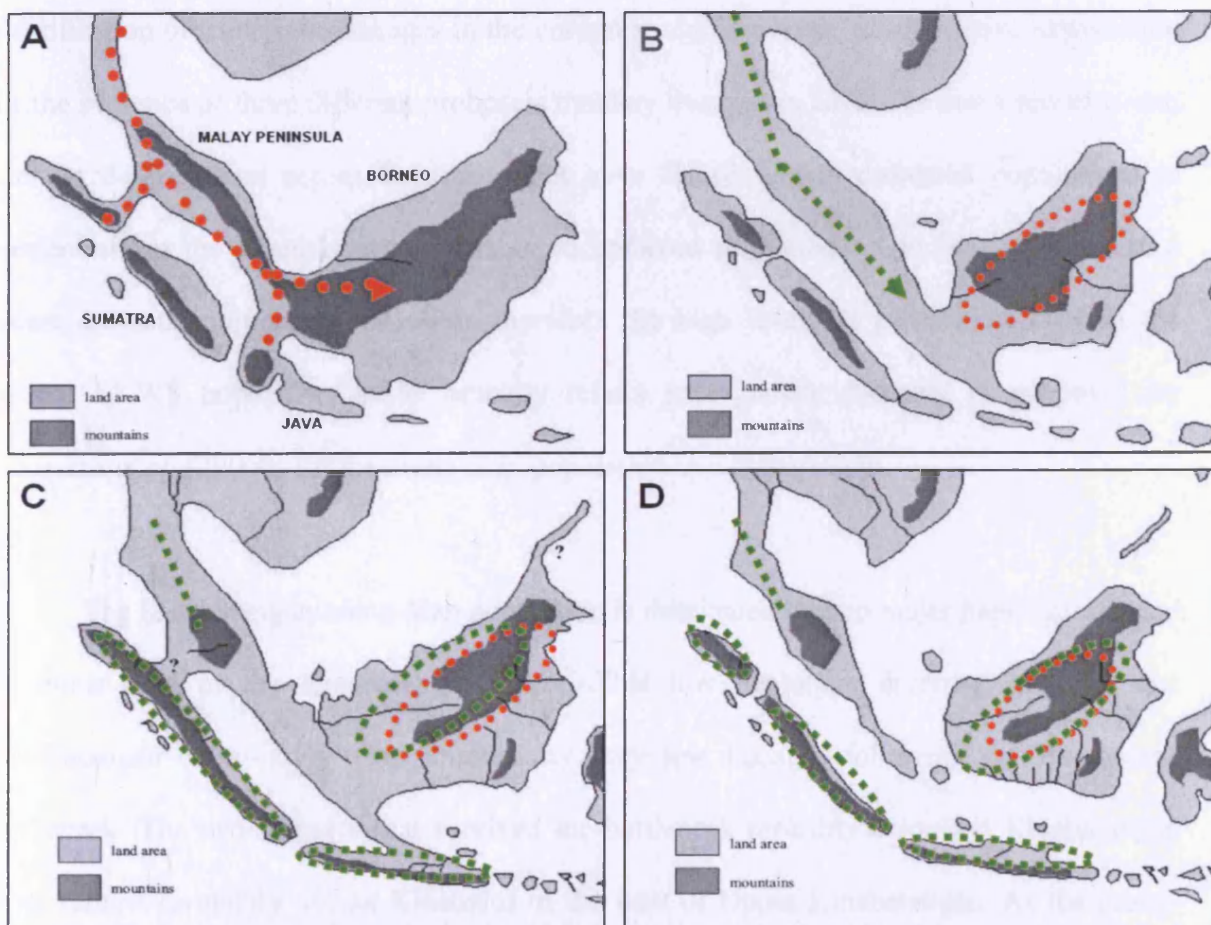
Both the orang-utan (current study; Goossens *et al.*, 2006b) and the proboscis monkey appear to have undergone bottlenecks. In the current study, the lowest genetic diversity was detected in the orang-utan and the highest in the proboscis monkey. How could animals with

previous histories of bottlenecks have such different levels of genetic diversity? The answer probably relates to their reproductive modes and dispersal abilities. The female biased population structure of proboscis monkeys, together with their age to maturity (5 years for females) and fecundity (probably around 5 offspring/female) results in a reproduction rate several times higher than that of the orang-utan (sexual maturity of female 11-15 years with an average fecundity of 5 offspring/female). Furthermore, the proboscis monkey has a much high dispersal capacity than the orang-utan. As the forest expanded and different fragments of rainforest merged during the late Pleistocene, isolated populations of proboscis monkey might have interbred, producing mixed lineages. Evidence for this hypothesis was provided by the identification of triphyletic lineages in the current study. However, an alternative explanation for the presence of three different proboscis monkey lineages in LKWS is that a recent event, such as deforestation across Borneo, might have forced widely dispersed populations to concentrate in the Kinabatangan. Such rapid enforced migration might have resulted in a recent dramatic population reduction, therefore the high levels of genetic diversity in the current LKWS population might actually reflect past genetic diversity as proposed by Goossens *et al.* (2006b) for the orang-utan population in Kinabatangan.

The Kinabatangan orang-utan population is dominated by two major haplotypes found on either side of the Kinabatangan River. This low haplotype diversity suggests that Kinabatangan orang-utans were founded by very few lineages following the Pleistocene bottleneck. The two lineages that survived the bottleneck probably colonized Kinabatangan from refugia (probably Mount Kinabalu) in the west of Upper Kinabatangan. As the orang-utan is a slow breeder with a very low density and being solitary animals, the expansion rate must have been much slower compared to the proboscis monkey (Harcourt & Schwart, 2001)

Climate fluctuation during the Pleistocene also influenced the distribution of forested areas. During the Pleistocene, the Sunda Shelf was exposed and based on accumulating

palynological evidence it is suggested that the Shelf was covered by cool, arid savannah-like vegetation (Goroj *et al.*, 2004). Between 190,000 to 130,000 years ago, during the dry glacial period, the low sea level (thus low river level) may have allowed freer movement within the Kinabatangan, possibly resulting in the slight the mixing of orang-utan haplotypes between riversides (Brandon-Jones, 1998; Goroj *et al.*, 2004). During these periods rain forest refugia were confined to northern and eastern Borneo according to a study based on termite community composition (Gathorne-Hardy *et al.*, 2002). If the orang-utan population expanded during and after this period this would suggest that the population had previously suffered a bottleneck as the forest contracted to savannah and orang-utan was confined in refugia. Due to the low density of the orang-utan population, their solitary nature and their dependence on rain-forest, the orang-utan is more prone to local extinction (Harcourt *et al.*, 2002). As conditions improved, the orang-utan might have colonized habitats including the Kinabatangan, but as the orang-utan reproduces slowly, the expansion would have been slow. Besides orang-utans, proboscis monkeys were also originally forest dependent species (Brandon-Jones, 1998). The genus evidently evolved in forest-woodland habitat and then when their habitats were overwhelmed by mangroves it began to adapt (Brandon-Jones, 1998).

Figure 6.4. The hypothetical colonization routes for orang-utan (red arrows), proboscis monkey (yellow arrows) and long-tailed macaques (white arrow).
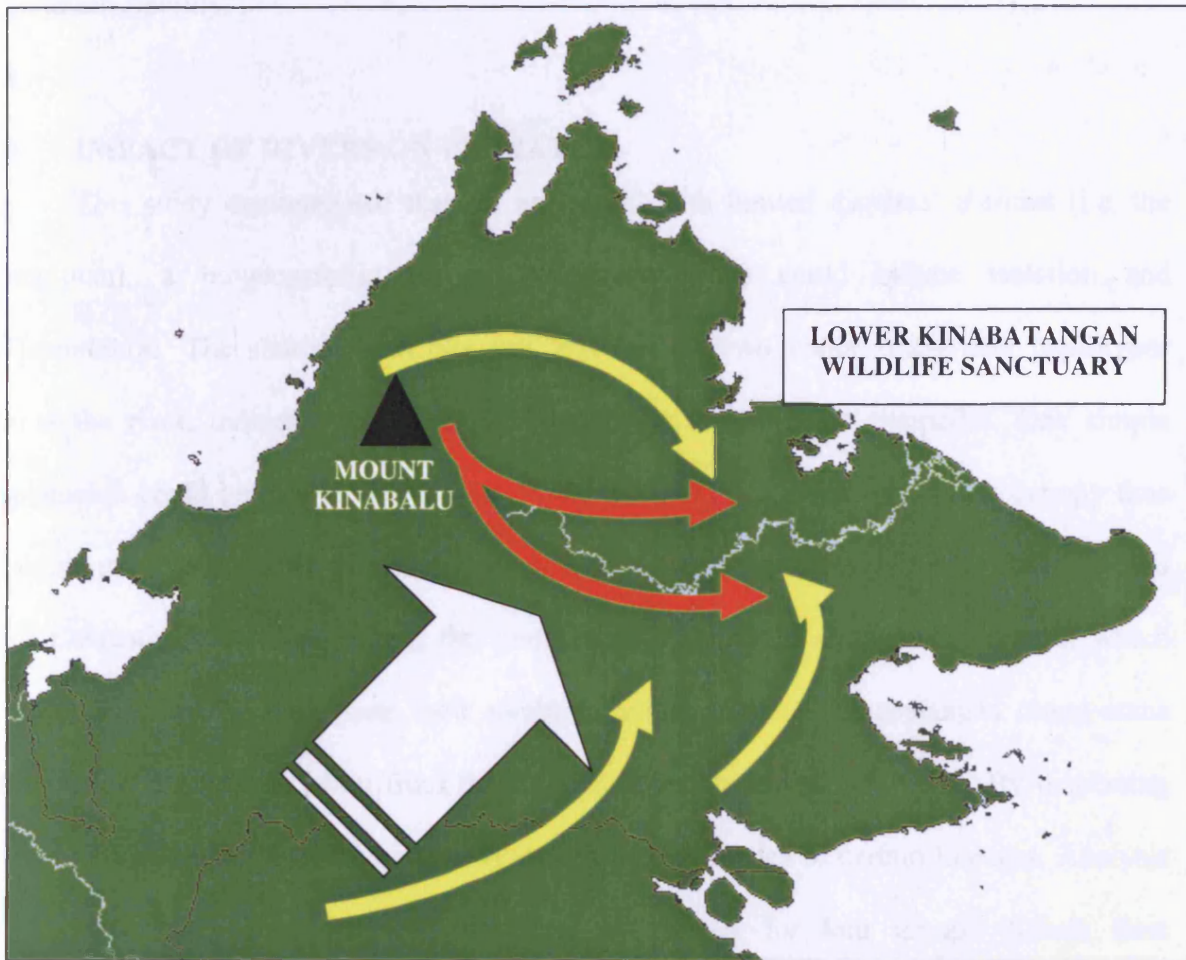
Unlike the orang-utan, there was no evidence of population bottlenecks in long-tailed macaques, instead our analyses indicated a large and stable population in the Kinabatangan. Long-tailed macaques are opportunistic feeders which can utilise a wide array of food and thus are adaptable to a variety of conditions (Fleagle, 1998). They can also adapt easily to human settlements, raiding plantations and houses for food (Wheatley, 1988). The 'super adaptability' of macaques makes them the most successful genus of primates extending from Japan to North Africa (Abegg & Thierry, 2002). They are included in the top 100 list of the most successful alien species (Lowe *et al.*, 2000; Long, 2003). With this flexibility on diet requirement and ability to adapt, the colonization of Kinabatangan would be easier for this species compared to the orang-utan and proboscis monkey (see Chapters 4 and 5). As long-tailed macaques live in a moderate to large social groups and their fast reproduction rates,

unlike orang-utans, enables them to spread and colonize more areas, generating the signal for population stability.

6.3 IMPACT OF RIVERS ON PRIMATES

This study demonstrated that for an animal with limited dispersal abilities (i.e. the orang-utan), a biogeographic barrier (i.e. large river) could induce isolation and differentiation. The sharing of haplotypes between the two major orang-utan haplotypes across the river, indicates some historical movement between the riversides. One simple explanation could be that the past forest of Kinabatangan formed a continuous canopy thus enabling movement within the canopy across the river aided by lower water levels associated with fluctuating sea levels during the Pleistocene. Orang-utan social behaviour, in which females will usually settle near their mother's range, suggests Kinabatangan orang-utans dispersed in a linear direction from the headwater to the sea or vice versa. By dispersing closed to one another this would have increased the dominance of certain lineages. Analysis of Bornean orang-utans, revealed reciprocal monophyly for four groups: Sabah, East Kalimantan, Central and Southwest Kalimantan, and Northwest Kalimantan and Sarawak. These four groups correspond to the three subspecies of orang-utan in which Sabah and East Kalimantan were *Pongo pygmaeus morio*, Central and Southwest Kalimantan were *P. p. wurmbii* and Northwest Kalimantan and Sarawak were *P. p. pygmaeus*. Warren *et al.* (2001) suggested that the separation was due to river barriers and this is strongly supported by the current study. Between the four major samples sites lie major rivers in Kalimantan, Sarawak and Sabah. For example, the two populations of *P. p. morio* are separated by the the Kayan River. The Mahakam and Barito rivers isolate *P. p. morio* from *P. p. wurmbii* in Central and Southwest Kalimantan. *P. p. pygmaeus* is isolated from subspecies *wurmbii* by the Kapuas River and from subspecies *morio* in the north by the Rajang and Baram rivers.

Within Borneo, the Kapuas River has been shown to act as a geographic barrier for orang-utans (Uchida, 1998). The current study also revealed the importance of rivers

impeding movement of orang-utans in Kinabatangan and Borneo (see Chapter 4). Historical records have revealed that orang-utans were absent from Brunei (area surrounded by Baram River) and South Kalimantan (between the Barito and Mahakam Rivers). This was probably related to the orang-utans inability to cross rivers, as these three large big rivers prevent arboreal movement through the canopy.

6.4   INFLUENCES OF OX-BOW LAKES

Ox-bow lakes are common in the LKWS, with the highest concentration occurring between Lots 3 and 4 (Fig. 6.5) and Lots 6 and Lot 7. Boonratana (cited in Bennett & Davies, 1994) reported that in Sukau, proboscis monkeys occur in abundance near ox-bow lakes. Based on the sampling map generated from GPS recordings (Figs. 3.1, 4.1 and 5.1 in Chapters 3, 4 and 5, respectively), primate sightings and sample collection did appear to be higher in the region of ox-bow lakes, but this requires further investigation.

Figure 6.5. The possible influences of ox-bow lake formation (Lot 4) in distribution of primates along the Kinabatangan River especially between North and South banks.

6.6   SOCIAL BEHAVIOUR INFLUENCES POPULATION STRUCTURE

Long-tailed macaques exhibit high haplotypic diversity with groups of closely related haplotypes that are connected to each other by long mutational steps. Each of the closely related haplotypes were distributed on the same side of the river, clearly indicating female bonding and a strong philopatric nature, characteristic of the macaque (Melnick, 1987). Similarly, the structure observed from network analysis in the current study may reflect the highly philopatric nature of female macaques (Melnick, 1987) and there was no indication to suggest a barrier to gene flow for this species.

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 there is extensive gene flow between demes in LKWS (see Chapter 5). Unlike the orang-utan, proboscis monkeys were found in higher densities, with a higher proportion of females than males and being peaceful enough to form large congregations of unimale groups (Kawabe & Mano, 1972). Such higher densities coupled with a much shorter generation time results in an increased reproduction rate. Free movement between females increases the number of maternal lines as indicated by the number of haplotypes and homogenising effects on both sides of the river (Murai, 2006; Murai *et al.*, 2006).

Unlike long-tailed macaques or proboscis monkeys, orang-utans are solitary animals. With a history of recent bottlenecks (Goossens *et al.*, 2006b), the long longevity and slow reproduction rates would hinder an immediate expansion of the orang-utan population in Kinabatangan. The population would slowly recover with evidence of founder events as observed in the current study (see Chapter 4).

6.7 CONSERVATION ISSUES FOR KINABATANGAN PRIMATES

The current study indicates that the primates in Kinabatangan must be managed according to their dispersal abilities and population structure. Moreover, this study and Goossens *et al.* (2005, 2006a, b) suggests that the future of the primates in Kinabatangan depends on management of the sanctuary. Further development or encroachment of the LKWS forest is likely to hasten the extinction process of vulnerable species (i.e. the orang-utan).

Currently LKWS is surrounded by ever encroaching oil-palm plantation and villages. The protection status of LKWS as a wildlife sanctuary should be enforced by local authorities to ensure protection of the species in LKWS, particularly against hunting and habitat degradation. Orang-utans and proboscis monkeys are listed on Schedule 1 (Section 2) of Sabah's totally protected species Part 1 in the Sabah Wildlife Conservation Enactment No. 6 of 1997. Long-tailed macaques, on the other hand, are listed on the same enactment under Schedule 2 (Section 2) of Sabah's protected species of animals with limited hunting and collection under licence. Hunting is one of the major obstacles in conservation project in Malaysia. Intensive hunting can exterminate low reproducing rate species and alter species composition (Laurance *et al.*, 2000). Similarly, in Amazonia, hunting pressures is one of the major threats for large animal and bird declines in fragmented forests (Laurance *et al.*, 2000). The Lots of LKWS are easily accessible for hunters due to the edge effect of fragmented habitats. During the current study, no evidence of hunting was detected, however this activity may be more apparent at the borders between the forest and the plantation. Orang-utans are still captured though, especially when they stray inside the plantation areas. Most of them are relocated into other forest reserves (i.e. Tabin Wildlife Reserve), but some juveniles are sent to a rehabilitation centre in Sepilok. This small forest reserve is run by Sabah Wildlife Department funded by the Government.

Volunteer projects, such as the Kinabatangan Orang-utan Conservation Project, a non-governmental organisation project of HUTAN, a French NGOs, have been successfully carried out in Sukau to educate the local people on the importance of conserving wildlife and forest. The same project could be applied to other areas in Kinabatangan or Sabah to manage the forest sustainably. Only through understanding the importance of maintaining our biodiversity will our campaign for conservation be successful. The cooperation of local people is vital to support any conservation efforts undertaken by government or non-government agencies. In the Lower Kinabatangan, the Model Ecologically Sustainable Community Tourism Project or MESCOT, centred at Batu Putih, on the extreme west of Sukau village is the other successful nature-tourism based project, which also involves local people. Their most successful project, the Miso Walai Homestay Programme, combined homestay with transport services and recreational activities.

Currently, LKWS is divided into 10 forest fragments (Lots), isolated or partially isolated from each other either by the Kinabatangan River or its tributaries, agricultural lands, degraded secondary forest, forest reserve and villages. In general, conservation biologists agree that landscape connectivity enhances population viability (Beier & Noss, 1998). Many studies have shown that many animals, birds and invertebrates are unable or unwilling to cross even small forest clearings (Bierregaard & Stouffer, 1997; Laurance et al., 2000). As orang-utans and proboscis monkeys require a large tract of land (home ranges of 0.42 - 7.77 km^2 and 9 km^2, respectively), it is imperative that connections between these fragments of protected land are re-established. Ideally, the Government should provide funding to purchase the land between forest fragments to create forest corridors. In addition, the heavily degraded forest could be re-planted with suitable plants (i.e. fruiting trees etc.) and managed to provide corridors of mature trees between fragments. As Lots 7 and 8 are divided by a major road

from Sandakan to Lahad Datu, arboreal corridors (such as artificial arboreal bridge) could be created to connect the forest between the two Lots. Recently, arboreal corridors have been installed in some sites in LKWS by local people with the help of KOCP and HUTAN. Other projects undertaken by local people, like re-planting forest on the heavily degraded forest fragments, have been funded by the WWF Malaysia.

Future Direction

Animal conservation projects should be supported by the analyses of several independent data sets (Wayne *et al.*, 1994). The use of different molecular markers often provides additional information unobtainable just by observational or ecological studies. The current study illustrates that a single molecular marker is insufficient to provide information for conservation purposes as highlighted by the information-rich datasets (mtDNA and microsatellites, see Appendix II, Goossens *et al.*, 2005) for the orang-utan compared to the information-poor data (mtDNA only) for proboscis monkeys and long-tailed macaques. Further information on male-inherited DNA (Y-chromosome) might provide interesting data with which to compare the female-inherited DNA to understand the contribution of each sex in population genetic diversity (Domingo-Roura *et al.*, 2001)

Future work on orang-utans should include larger sampling areas as the current study indicates that orang-utan populations seem to be influenced by biogeographic barriers such as rivers. This provides an interesting question on the distribution of orang-utans, especially why for some parts of Borneo such as Brunei, there are no records of orang-utans. Are rivers (Baram and Tutoh) that surround Brunei responsible for preventing colonization of this area by orang-utans?

Nevertheless, the information currently obtained by mtDNA should provide a motivation for a larger study to investigate the population structure in Kinabatangan. With more sample coverage from all over Borneo, the question remains as to whether the proboscis monkey retains the triphyletic clade patterns currently observed and if so, how is spatially organised. The triphyletic clades observed in Kinabatangan would be an interesting to investigate over a larger scale.

The current study, even with its limitations, has revealed the conservation value of the Lower Kinabatangan Wildlife Sanctuary as an important place for primates in Sabah. However, more studies are needed to accurately answer questions on dispersal and genetic structure of the threatened primate populations living in the Kinabatangan. Also, further research (i.e. kinship, genetic diversity within Lots etc.) is needed before we can predict the fate of these three primates in Kinabatangan.

REFERENCES

Abegg, C. & Thierry, B. 2002. *Macaque evolution and dispersal in insular south-east Asia.* Biological Journal of the Linnean Society 75:555–576.

Agoramoorthy, G. & Hsu, M.J. 2005a. *Borneo's proboscis monkey – a study of its diet of mineral and phytochemical concentration.* Current Science 89:454-457.

Agoramoorthy, G. & Hsu, M.J. 2005b. *Occurrence of infanticide among wild proboscis monkeys (Nasalis larvatus) in Sabah, Northern Borneo.* Folia Primatologica 76:177-179.

Ahston, P.S. 1988. *Dipterocarp biology as a window to the understanding of tropical forest structure.* Annual Review of Ecology and Systematics 19:347-370.

Allen, M., Engström, A., Meyers, S., Handt, O., Saldeen, T., von Haeseler, A., Pääbo, S. & Gyllensten, U. 1998. *Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities.* Journal of Forensic Science 43:453-464.

Ambrose, S.H. 1998. *Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans.* Journal of Human Evolution 34: 623–651.

Ambrose, S.H. 2003. *Did the super-eruption of Toba cause a human population bottleneck? Reply to Gathorne-Hardy and Harcourt Smith.* Journal of Human Evolution 45: 231–237.

Ancrenaz, M., Calaque, R. & Lackman-Ancrenaz, I. 2004. *Orang-utan nesting behaviour in disturbed forest of Sabah, Malaysia: implications for nest census.* International Journal of Primatology 25:983-1000.

Ancrenaz, M., Gimenez, O., Ambu, L., Ancrenaz, K., Andau, P., Goossens, B., Payne, J., Sawang, A., Tuuga, A. & Lackman-Ancrenaz, I. 2005. *Aerial surveys give new estimates for orang-utans in Sabah, Malaysia.* Public Library of Science Biology 3:e3.

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 7:375-385.

Anonymous. 2006. What is the Aquatic Ape Theory (AAT/H)? (http://www.aquaticape.org/whataat.html). Assessed online on 29 September 2006.

Ascunce, M.S., Cortes-Ortiz, L. & Mudry, M.D. 2003. *The mitochondrial control region of the black howler monkey, Alouatta caraya (Primates, Platyrrhini), and the development of new primers.* Molecular Ecology Notes 3:372-375.

Avise, J.C. 2000. *Phylogeography. The history and formation of species.* Harvard University Press.

Avise, J.C. 2004. *Molecular markers, natural history and evolution.* 2nd Edition. Sinauer Associates, Sunderland, Massachusetts.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. 1987. *Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics.* Annual Review of Ecology and Systematics 18:489-522.

Ayres, J.M. & Clutton-Brock, T.H. 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. An introduction to its natural vegetation including a preliminary plant checklist of the region.* Project MYS359/96. WWF Malaysia. PP87.

Baker, C.S., Lambertsen, R.H., Weinrich, M.T., Calambokidis, J., Early, G. & O'Brien, S.J. 1991. *Molecular genetic identification of the sex of humpback whales (Megaptera novaeangliae).* International Whaling Commission (Special Issue 13): 105-111.

Ballard, J.W.O & Rand, D.M. 2005. *The population biology of mitochondrial DNA and its phylogenetic implications.* Annual Review of Ecology, Evolution and Systematic 36:621-642.

Bandelt, H.J., Forster, P. & Röhl, A. 1999. *Median-joining networks for inferring intraspecific phylogenies.* Molecular Biology and Evolution 16:37-48.

Barkman, T.J. & Simpson, B.B. 2001. *Origin of high-elevation Dendrochilum species (Orchidaceae) endemic to Mount Kinabalu, Sabah, Malaysia.* Systematic Botany 26:658-669.

Bawa, K. S. & Seidler, R. 1998. *Natural forest management and conservation of biodiversity in tropical forests.* Conservation Biology 12:46-55.

Beaman, J.H. 2005. *Mount Kinabalu: hotspot of plant diversity in Borneo.* Biologiske Skrifter 55:103-127.

Becker, P.H., Voight, C.C., Arnold, J.M. & Nagel, R. 2006. *A non-invasive technique to bleed incubating birds without trapping: a blood-sucking bug in a hollow egg.* Journal of Ornithology 147:115-118.

Beier, P. & Noss, R.F. 1998. *Do habitat corridors provide connectivity?* Conservation Biology 12:1241-1252.

Bennett, E.L. & Davis, G. 1994. *The ecology of Asian colobines.* In: Colobine Monkeys: their ecology, behaviour and evolution. Davies, A.G. & Oates, J.F. (eds.) Cambridge University Press, Cambridge. UK. pp.129-171.

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

Bermingham, E. & Avise, J.C. 1986. *Molecular zoogeography of freshwater fishes in the southeastern United States.* Genetics 113:939-965.

Bernard, H. 1997. *A study on the distribution and abundance of proboscis monkey (Nasalis larvatus) in the Klias peninsula.* Journal of Wildlife Management and Research 1:1-12.

Bierregaard, R.O. & Stouffer, P.C. 1997. *Understorey birds and dynamic habitat mosaics in Amazonian rainforests.* In: Tropical Forest Remnants. Ecology, management and conservation of fragmented communities. Laurance, W.F. & Bierregaard, R.O. (eds.). Chicago University Press, Chicago. USA. pp.138-155.

Bigoni, F., Stanyon, R., Wimmer, R. & Schempp, W. 2003. *Chromosome painting shows that the proboscis monkey (Nasalis larvatus) has a derived karyotype and is phylogenetically nested within asian colobines* American Journal of Primatology 60:85-93.

Bird, M.I., Taylor, D. & Hunt, C. 2005. *Palaeoenvironment of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland?* Quaternary Science Reviews 24:2228-2242.

Birungi, J. & Arctander, P. 2000. *Large sequence divergence of mitochondrial DNA genotypes of the control region within populations of the African antelope, kob (Kobus kob).* Molecular Ecology 9:1997-2008.

Boonratana, R. 1993. *The ecology and behavior of the proboscis monkey (Nasalis larvatus) in the Lower Kinabatangan, Sabah.* PhD. Thesis. Mahidol University, Bangkok.

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.

Boonratana, R. 2003. *Feeding ecology of proboscis monkey (Nasalis larvatus) in the Lower Kinabatangan, Sabah, Malaysia.* Sabah Parks Nature Journal 6:1-26.

Brandon-Jones, D. 1996. *The Asian Colobinae (Mammalia: Cercopithecidae) as indicators of Quartenary climatic change.* Biological Journal of the Linnean Society 59:327-350.

Brandon-Jones, D., 1998. *Pre-glacial Bornean, primate impoverishment and Wallace's line.* In: Biogeography and geological evolution of Southeast Asia. Hall, R. & Holloway, J.D. (eds.). Backhuys Publishers, Leiden, The Netherlands. pp.393-404.

Brooks, T.M., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Rylands, A.B., Konstant, W.R., Flick, P., Pilgrim, J., Oldfield, S., Magin, G. & Hilton-Taylor, C. 2002. *Habitat loss and extinction in the hotspot of biodiversity.* Conservation Biology 16:909-923.

Brown, G.K., Nelson, G. & Ladiges, P.Y. 2006. *Historical biogeography of Rhododendron section Vireya and the Malesian archipelago.* Journal of Biogeography 33: 1929-1944.

Brown, W.M., George, M., Jr. & Wilson, A.C. 1979. *Rapid evolution of animal mitochondrial DNA.* Proceedings of the National Academy of Sciences USA 76:1967-1971.

Cannon, C.H. & Manos, P.S. 2003. *Phylogeography of the Southeast Asian stone oaks (Lithocarpus).* Journal of Biogeography 30:211-226.

Cannon, C.H., Peart, D.R. & Leighton, M. 1998. *Tree species diversity in commercially logged Bornean rainforest.* Science 281:1366-1368.

Chapman, C. A., Balcomb, S. R. Gillespie, T. R. Skorupa, J. P. & Struhsaker, T. T. 2000. *Long-term effects of logging on African Primate Communities: a 28 year comparison from Kibale National Park, Uganda.* Conservation Biology 14: 207-217.

Chapman, C.A., Wasserman, M.D., Gillespie, T.R., Speirs, M.L., Lawes, M.J., Saj, T.L. & Ziegler, T.E. 2006. *Do food availability, parasitism and stress have synergistic effects on red colobus populations living in forest fragments?* American Journal fo Physical Anthropology 131:525-534.

Clifford, S.L., Anthony, N.M., Bawe-Johnson, M., Abernethy, K.A., Tutin, C.E.G., White, L.J.T., Bermejo, M., Goldsmith, M.L., McFarland, K., Jeffery, K J., Bruford, M.W. &

Wickings, E. J. 2004. *Mitochondrial DNA phylogeography of western lowland gorillas (Gorilla gorilla gorilla)*. Molecular Ecology 13:1551-1561.

Cockerham, C. C. 1969. *Variance of gene frequencies*. Evolution 23:72-83.

Cockerham, C. C. 1973. *Analysis of gene frequencies*. Genetics 74: 679-700.

Collura, R.V., Auerbach, M.R. & Stewart, C.B. 1996. *A quick direct method that can differentiate expressed mitochondrial genes from their nuclear pseudogenes*. Current Biology 6:1337-1339.

Colón, C.P. 2002. *Ranging behaviour and activity of the Malay civet (Viverra tangalunga) in a logged and an unlogged forest in Danum Valley, East Malaysia*. Journal of Zoology 257:473-485.

Cowlishaw, G. & Dunbar, R. 2000. *Primate conservation biology*. Chicago University Press. Chicago. USA. pp.93-118.

Curran, L.M., Trigg, S.N., McDonald, A.K., Astiani, D., Hardiono, Y.M., Siregar, P., Caniago, I & Kasischke, E. 2004. *Lowland forest loss in protected areas of Indonesian Borneo*. Science 303:1000-1003.

Davies, A.G. & Oates, J.F. 1994. *Colobine monkey: their ecology, behaviour and evolution*. Cambridge University Press, Cambridge. UK.

Davies, S.J., Palmiotto, P.A., Ashton, P.S., Lee, H.S. & Lafrankie, J.V. 1998. *Comparative ecology of 11 sympatric species of Macaranga in Borneo: tree distribution in relation to horizontal and vertical resource heterogeneity*. Journal of Ecology 86:662-673.

de Ruiter J.R. & Geffen, E. 1998. *Relatedness of matrilines, dispersing males and social groups in long-tailed macaques (Macaca fascicularis)*. Proceeding of the Royal Society of London B 265:79-87.

de Ruiter, J.R. 1994. *Behaviour and genes in natural populations of long-tailed macaques (Macaca fascicularis)*. PhD Thesis. Universiteit Utrecht.

Delgado Jr., R.A. & Van Schaik, C.P. 2000. *The behavioural ecology and conservation of the orang-utan (Pongo pygmaeus): a tale of two islands*. Evolutionary Anthropology 9:201-218.

Delport, W., Ferguson, J.W.H. & Bloomer, P. 2002. *Characterization and evolution of the mitochondrial DNA control region in hornbills (Bucerotiformes)*. Journal of Molecular Evolution 54:794-806.

Dennis, R.A. & Colfer, C.P. 2006. *Impacts of land use and fire on the loss and degradation of lowland forest in 1983-2000 in East Kutai District, East Kalimantan, Indonesia*. Singapore Journal of Tropical Geography 27:30-48.

DiFiore, A. & Randall, D. 1994. *Evolution of social organisation: a reappraisal for primates by using phylogenetic methods*. Proceeding of National Academy of Science USA 91:9941-9945.

Disotell, T.R. 1996. *The phylogeny of old world monkeys*. Evolutionary Anthropology 5:18-24.

Domingo-Roura, X., Marmi, J., López-Giráldez, J.F. & Garcia-Franquesa, E. 2001. *New molecular challenges in animal conservation*. Animal Biodiversity and Conservation 24:19-29.

Dudgeon, D. 2000. *The ecology of tropical Asian rivers and streams in relation to biodiversity.* Annual Review of Ecology and Systematics 31:239-263.

Eizirik, E., Bonatto, S.L., Johnson, W.E., Crawshaw, P.G., Vie, J.C., Brousset, D.M., O'Brien, S.J. & Salzano, F.M. 1998. *Phylogeographic patterns and evolution of the mitochondrial DNA control region in two neotropical cats (Mammalia, Felidae).* Journal of Molecular Evolution 47:613-624.

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

Excoffier L. 2004. *Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model.* Molecular Ecology 13:853-864.

Excoffier, L., & Smouse, P. 1994. *Using allele frequencies and geographic subdivision to reconstruct gene genealogies within a species. Molecular variance parsimony.* Genetics 136:343-59.

Excoffier, L., Laval, G. & Schneider, S. 2005. *Arlequin ver. 3.0: An integrated software package for population genetics data analysis.* Evolutionary Bioinformatics Online 1:47-50.

Excoffier, L., Smouse, P. & Quattro, J. 1992 *Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data.* Genetics 131:479-491.

Fahrig, L. 2003. *Effects of habitat fragmentation on biodiversity.* Annual Review of Ecology and Systematics 34:487-515.

FAO 2001. *Global Forest Resources Assessment 2000.* Main Report. FAO Forestry Paper 140, Rome 2001, Italy, pp. 479.

Felton, A.M., Engström, L.M., Felton, A. & Knott, C.D. 2003. *Orang-utan population density, forest structure and fruit availability in hand-logged and unlogged peat swamp forests in West Kalimantan, Indonesia.* Biological Conservation 114:91-101.

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

Frankham, R. 1996. *Relationship of genetic variation to population size in wildlife.* Conservation Biology 10:1500-1508.

Frankham, R. 2005. *Genetics and extinction.* Biological Conservation. 126.131-140.

Frankham, R., Ballou, J.D. & Briscoe, D.A. 2002. *Introduction to conservation genetics.* Cambridge University Press.

Fu, Y.X. & Li, W.H. 1993. *Statistical tests of neutrality of mutations.* Genetics 133:693-709.

Fu, Y.X. 1997. *Statistical tests of neutrality against population growth, hitchhiking and background selection.* Genetics 147:915-925.

Fu, Y.X. 1997. *Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.* Genetics 147: 915-925.

Fumagalli, L. Taberlat, P., Favre, L. & Hausser, J. 1996. *Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews.* Molecular Biology and Evolution 13:31-46.

Galdikas, B.M.F. 1988. *Orang-utan diet, range and activity at Tanjung Putting, Central Borneo.* International Journal of Primatology 9:1-31.

Garthorne–Hardy, F.J., Syaukani, Davies, R.G., Eggleton, P. & Jones, D.T. 2002. *Quarternary rainforest refugia in southeast Asia: using termites (Isoptera) as indicators.* Biological Journal of the Linnean Society 75:453-466.

Gathorne-Hardy, F.J. & Harcourt-Smith, W.E.H. 2003. *The super-eruption of Toba, did it cause a human bottleneck?* Journal of Human Evolution 45:227–230

Geological Survey of Malaysia, 1989. *Geological Survey of Malaysia Annual Report 1989.* Ministry of Primary Industries.

Goldstein, D.B. & Schlotterer, C. (eds.) 1999. *Microsatellites: evolution and applications.* Oxford University Press. Oxford. UK.

Goossens, B., Antony, N., Jeffery, K., Johnson-Bawe, M. & Bruford, M.W. 2003. Collection, storage and analysis of non-invasive genetic material in primate biology. In: *Field and Laboratory Methods in Primatology: A Practical Guide.* Setchell, J.M. & Curtis, D.J. (eds). Cambridge University Press, Cambridge, UK. pp 295-308.

Goossens, B., Chicki, L., Ancrenaz, M., Lackman-Ancrenaz, I., Andau & Bruford, M.W. 2006b. *Genetic signature of anthropogenic population collapse in orang-utan.* Public Library of Science Biology 4:285-291

Goossens, B., Chikhi, L., Jalil, M.F., Ancrenaz, M., Lackman-Ancrenaz, I., Mohamed, M., Andau, P. & Bruford, M.W. 2005. *Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (Pongo pygmaeus) populations from Sabah, Malaysia.* Molecular Ecology 14:441-456.

Goossens, B., Chikhi, L., Utami, S.S., De Ruiter, J. & Bruford, M.W. 2000. *A multi-sample, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape.* Conservation Genetics 1:157-162.

Goossens, B., Setchell, J.M., Abulani, D.M.A., Jalil, M.F., James, S.S., Aris, S., Lakim, M., Seventri, A.D., Sariningsih, S.S. & Ancrenaz, M. 2003. *A boat survey of primates in the lower Kinabatangan Wildlife Sanctuary. Lower Kinabatangan Scientific Expedition.* Universiti Malaysia Sabah, Malaysia. pp.37-46

Goossens, B., Setchell, J.M., James, S., Funk, S.M., Chikhi, L., Abulani, A., Ancrenaz, M., Lackman-Ancrenaz, I. & Bruford, M.W. 2006a. *Philopatry and reproductive success in Bornean orang-utans (Pongo pygmaeus).* Molecular Ecology 15:2577-2588.

Gorog, A.J., Sinaga, M.H. & Engstrom, M.D. 2004. *Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (Maxomys surifer, Leopoldamys sabanus and Maxomys whiteheadi).* Biological Journal of the Linnean Society 81:91-109.

Grant, W.S. & Bowen, B.W. 1998. *Shallow population histories in deep evolutionary lineages of marine fishes: insight from sardines and anchovies and lessons for conservation.* Journal of Heredity 89:415-426

Groves, C.P. 1986. *Systematics of the great apes.* In Swindler, D.R. & Erwin, J. (eds). Comparative primate biology 1: Systematics, evolution and anatomy. New York. pp.187-217.

Groves, C.P. 2001. *Primate taxonomy.* Smithsonian Institution. pp. 298-300.

Groves, C.P., Westwood, C. & Shea, B.T. 1992. *Unfinished business: Mahalanobis and a clockwork orang.* Journal of Human Evolution 22:327-340.

Grytnes, J.A. & Beaman, J.H. 2006. *Elevational species richness patterns for vascular plants on Mount Kinabalu, Borneo.* Journal of Biogeography 33:1838-1849.

Hall, R. 1996. *Reconstructing Cenozoic SE Asia.* In: Tectonic Evolution of SE Asia. Hall, R. & Blundell, D. J. (eds.). Geological Society of London Special Publication, 106:153-184.

Hall, R. 1998. *The plate tectonics of Cenozoic South East Asia and the distribution of land and sea.* In: Biogeography and geological evolution of Southeast Asia. Hall, R. & Holloway, J.D. (eds.), Backhuys Publishers, Leiden, The Netherlands. pp. 99-131.

Halos, S.C., Chu, J.Y., Ferreon, A.C.M. & Magno, M.M.F. 1999. *Philippine population database at nine microsatellite loci for forensic and paternity applications.* Forensic Science International 101:27-32.

Hancock, J.M. 1999. *Microsatellites and other simple sequences: genomic context and mutational mechanisms.* In: Microsatellites: evolution and applications. Goldstein, D.B. & Schlotterer, C. (eds.). Oxford University Press. Oxford. UK. pp.1-9.

Hanke, M. & Wink, M. 1994. *Direct DNA sequencing of PCRr-amplified vector inserts following enzymatic degradation of primer and dNTPs.* Biotechniques 17:858-860.

Hapke, A., Zinner, D. & Zischler, H. 2001. *Mitochondrial DNA variation in Eritrean hamadryas babbons (Papio hamadryas hamadryas): life history influences population genetic structure.* Behavioural Ecology and Sociobiology 50:483-492.

Harcourt, A.H. & Schwartz, M.W. 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.

Harcourt, A.H. 1999 *Biogeographic relationships of primates on south-east Asian islands.* Global Ecology and Biogeography 8:55-61.

Harcourt, A.H., Coppeto, S.A. & Parks, S.A. 2002. *Rarity, specialization and extinction in primates.* Journal of Biogeography 29:445-456.

Harrison, T., Krigbaum, J. & Manser, J. 2006. Chapter 12: *Primate Biogeography and ecology on the Sunda Shelf islands: a paleontological and zooarchaeological perspecitive.* In: Primate Biogeography. Lehman, S.M. & Fleagle, J.G. (eds.). Springer, New York. USA. pp.331-372.

Hauser, N., Peckham, H. & Clapham, P. 2000. *Humpback whales in the Southern Cook Islands, South Pacific.* Journal of Cetacean Research and Management 2:159-164.

Hayaishi, S. & Kawamoto, Y. 2006. *Low genetic diversity and· biased distribution of mitochondrial DNA haplotypes in the Japanese macaque (Macaca fuscata yakui) on Yakushima Island.* Primates 47:158-164.

Hayashi, J.I., Tagashira, Y. & Yoshida, M.C. 1985. *Absence of extensive recombination between inter- and intraspecies mitochondrial DNA in mammalian cells.* Experimental Cellular Research 160:387-395.

Heaney, L.R. 1991. *A synopsis of climatic and vegetational change in Southeast Asia.* Climatic Change 19:53-61.

Hewitt, G. 2000. *The genetic legacy of the quaternary ice ages.* Nature 405:907-913.

Heydon, M.J. & Bulloh, P. 1997. *Mousedeer densities in a tropical rainforest: the impact of selective logging.* Journal of Applied Ecology 34:484-496.

Hoech, W.R., Blakley, K.H. & Brown, W.M. 1991. *Heteroplasmy suggests limited biparental inheritance of Mytilus mitochondrial DNA*. Science 251:1488-1490.

Hoelzer, G. A., Dittus, W. P. J., Ashley, M. V. & Melnick, D. J. 1994. *The local distribution of highly divergent mitochondrial DNA haplotypes in toque macaques Macaca sinica at Polonnaruwa, Sri Lanka*. Molecular Ecology 3: 451-458.

Hutchison, C.S., Bergman, S.C., Swauger, D.A. & Graves, J.E. 2000. *A Miocene collisional belt in North Borneo : uplift mechanism and isostatic adjustment quantified by thermochronology*. Journal of the Geological Society 157:783-793.

Inger, R. F. & Voris, H.K. 2001. *The biogeographical relations of the frogs and snakes of Sundaland*. Journal of Biogeography 28:863-891.

IUCN 2006. *2006 IUCN Red List of Threatened Species*. (www.iucnredlist.org). Downloaded on 21 August 2006.

IUCN, 2000. *Red list of threatened species*. Available from: www.iucn.org/redlist/2000.

Jefferey, K. 2003. *Application of forensic genetics to the population biology of western lowland gorillas at Lopé, Gabon*. Ph.D. Thesis. Cardiff University.

Johns, A. D. 1986. *Effects of selective logging on the behavioural ecology of west Malaysian primates*. Ecology 67:684-694.

Jomo, K.S., Chang, Y. T., & Khoo, K.J. 2004. *Deforesting Malaysia: the political economy and social ecology of agricultural expansion and commercial logging*. Zed Books in association with United Nations Research Institute for Social Development. pp.96-146.

Kanthaswamy, S. & Smith, D.G. 2002. *Population subdivision and gene flow among wild orangutans*. Primates 43:315-327.

Kanthaswamy, S., Kurushima, J.D. & Smith, D.G. 2006. *Inferring Pongo conservation units: a perspective based on microsatellite and mitochondrial DNA analyses*. Primates 47:310-321.

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

Kerley, L.L., Goodrich, J.M., Miquelle, D.G., Smirnov, E.N., Nikolaev, I.G., Quigley, H.B. & Hornocker, M.G. 2002. *Effects of roads and human disturbance on Amur tigers*. Conservation Biology 16:1-12.

Kern, J. 1964. *Observations on the habits of the Proboscis Monkey made in the Brunei Bay Area, Borneo*. Zoologica, 49:183-192.

Kinnaird, M.F., Sanderson, E.W., O'Brien, T.G. Wibisono, H.T. & Woolmer, G. 2003. *Deforestation trends in a tropical landscape and implications for endangered large mammals*. Conservation Biology 17:245-257.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Paabo, S., Villablanca, F.X. & Wilson, A.C. 1989. *Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers*. Proceeding of the National Academy of Science USA 86:6196-6200.

Kuhner, M. K., Yamato, J. & Felsenstein, J. 1998 *Maximum likelihood estimation of population growth rates based on the coalescent*. Genetics 149:429-434.

Lansman, R.A., Shade, R.O., Shapira, J.F. & Avise, J.C. 1981. *The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III. Techniques and potential applications.* Journal of Molecular Evolution 17:214-226.

Launhardt, K., Borries, C., Hardt, C., Epplen, T. & Winkler, P. 2001. *Paternity analysis of alternative male reproductive routes among the langurs (Semnopithecus entellus) of Ramnagar.* Animal Behaviour 61:53-64.

Laurance, W.F., Vasconcelos, H.L. & Lovejoy, T.E. 2000. *Forest loss and fragmentation in the Amazon: implication for wildlife conservation.* Oryx 34:39-45.

Lawler, S.H., Sussman, R.W. & Taylor, L.L. 1995. *Mitochondrial DNA of the Mauritian macaques (Macaca fascicularis): an example of the founder effect.* American Journal of Physical Anthropology 96:133-141.

Li, H., Meng, S., Men, Z., Fu, Y. & Zhang, Y. 2003. *Genetic diversity and population history of golden monkey (Rhinopithecus roxellana).* Genetics 164:269-275.

Li, M., Wei, F., Goossens, B., Feng, Z., Tamate, H.B., Bruford, M.W. & Funk, S.M. 2005. *Mitochondrial phylogeography and subspecific variation in the red panda (Ailurus fulgens): implications for conservation.* Molecular Phylogenetics and Evolution 36:78-89.

Long, J. C. 1986. *The allelic correlation structure of Gainj and Kalam speaking people. I. The estimation and interpretation of Wright's F-statistics.* Genetics 112: 629-647.

Long, J.L. 2003. *Introduced mammals of the world: their history, distribution and influence.* CSIRO Publishing.

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), pp. 12.

Lutz, S., Weisser, H.J., Heizmann, J. & Pollak, S. 1998. *Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany.* International Journal of Legal Medicine 112:145-150.

Macdonald, D.W. 1982. *Notes on the size and composition of groups of proboscis monkey, Nasalis larvatus.* Folia Primatologica 37:95-98.

MacKinnon, J. & MacKinnon, K.S. 1974. *The behaviour and ecology of wild orang-utan (Pongo pygmaeus).* Behaviour 22:3-74.

Marmi, J., Bertranpetit, J., Terradas, J., Takenaka, O. & Domingo-Roura, X. 2004. *Radiation and phylogeography in the Japanese macaque, Macaca fuscata.* Molecular Phylogenetics and Evolution 30:676-685.

Masembe, C., Muwanika, V.B., Nyakaana, S., Arctander, P. & Siegismund, H.R. 2006. *Three genetically divergent lineages of the oryx in eastern Africa: evidence for an ancient introgression hybridization.* Conservation Genetics 7:551-562.

Matthee, C.A., Fourie, F., Oosthuizen, W.H., Meyër, M.A. & Tolley, K.A. 2006. *Mitochondrial DNA sequence data of the Cape fur seal (Arctocephalus pusillus pusillus) suggest that population numbers may be affected by climatic shifts.* Marine Biology 148:899-905.

Maudet, C., Miller, C., Bassano, B., Breitenmoser-Würsten, C., Gauthier, D., Obexer-Ruff, G., Michallet, J., Taberlet, P. & Luikart, G. 2002. *Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex (Capra ibex (ibex))*. Molecular Ecology 11:421-436.

McMorrow, J. & Talip, A.M. 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, & Groves, C.P. 2006. *Chapter 11: The geography of mammals and rivers in mainland southeast Asia*. In: Primate Biogeography. Lehman, S.M. & Fleagle, J.G. (eds.) Springer, New York. USA. pp. 305-329.

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

Meijaard, E. 2003. *Mammals of south-east Asian islands and their Late Pleistocene environments*. Journal of Biogeography 30:1245-1257.

Meijaard, E., Sheil, D., Nasi, R. & Stanley, S.A. 2006. *Wildlife conservation in Bornean timber concessions.* Ecology and Society 11:47. (Online. http://www.ecologyandsociety.org/vol11/iss1/art47/).

Melnick, D.J. & Hoelzer, G.A. 1996. *The population genetic consequences of macaque social organisation and behaviour*. In: Evolution and ecology of macaque societies. Fa, J.E. & Lindburg, D.G. (eds.) Cambridge University Press. Cambridge. UK. pp.413-443.

Melnick, D.J. 1987. *The genetic consequences of primate social organization: a review of macaques, babbons and vervet monkeys*. Genetica 73:117-135.

Melnick, D.J., Hoelzer, G.A., Absher, R. & Ashley, M.V. 1993. *mtDNA diversity in rhesus monkeys reveals overestimates of diversgence time and paraphyly with neighbouring species*. Molecular Biology and Evolution 10:282-295.

Modolo, L., Salzburger, W. & Martin, R.D. 2005. *Phylogeography of barbary macaques (Macaca sylvanus) and the origin of the Gibraltar colony*. Proceedings of the National Academy of Science USA 102:7392-7397.

Moore, W.S. 1995. *Inferring phylogenies from mtDNA variation: mitochondrial-gene trees vs. nuclear-gene trees*. Evolution 49:718-726.

Morgan. E. 1999. Aquatic Ape Hypothesis. Souvenir Press (New Edition).

Morin, P.A., Chambers, K.E., 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.

Morin, P.A., Wallis, J., Moore, J.J., Chakraborty, R. & Woodruff, D.S. 1993. *Non-invasive sampling and DNA amplification for paternity exclusion, community structure and phylogeography in wild chimpanzees*. Primates 34:347-356.

Moritz, C. & Faith, D.P. 1998. *Comparative phylogeography and the identification of genetically divergent areas for conservation*. Molecular Ecology 7:419-429.

Moritz, C., Dowling, T.E. & Brown, W.M. 1987. *Evolution of animal mitochondrial DNA: relevance for population biology and systematics*. Annual Review of Ecology and Systematic 18:269-292.

Morley, R.J. 1998. *Palynological evidence for Tertiary plant dispersals in the SE Asian region in relation to plate tectonics and climate.* In: Biogeography and geological evolution of SE Asia. Holloway, J.D. & Hall, R. (eds.) Backhuys. Leiden. The Netherlands. pp. 211-234.

Muir, C.C., Galdikas, B.M.F. & Beckenbach, A.T. 1998. *Is there sufficient evidence to elevate the orang-utan of Borneo and Sumatra to separate species?* Journal of Molecular Evolution 46:378-381.

Murai, T. 2004. *Social behaviours of all-male proboscis monkeys when joined by females.* Ecological Research 19:451-454.

Murai, T. 2006. *Mating behaviours of the proboscis monkey.* American Journal of Primatology 68:832-837.

Murai, T., Mohamed, M., Bernard, H., Andau, P.M., Saburi, R. & Higashi, S. 2006. *Female transfer between one-male groups of proboscis monkey (Nasalis larvatus).* Primates: in press.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. 2000. *Biodiversity hotspots for conservation priorities.* Nature 403:853-858.

Na-Nakorn, U., Sukmanomon, S., Nakajima, M., Taniguchi, N., Kamonrat, W., Poompuang, S. & Nguyen, T.T.T. 2006. *MtDNA diversity of the critically endangered Mekong giant catfish (Pangasianodon gigas Chevey, 1913) and closely related species: implications for conservation.* Animal Conservation 9: 483-494.

Nei, M. 1987. *Molecular Evolutionary Genetics.* Columbia University Press, New York, NY, USA.

Newsome, J. & Flenley J.R. 1988. *Late quarternary vegetation history of the central highlands of Sumatra II. Palaeopalynology and vegetational history.* Journal of Biogeography 15:555-578.

Noad, J. 2001. *The Gomantong Limestone of eastern Borneo: a sedimentological comparison with the near-contemporaneous Luconia Province.* Palaeogeography, Palaeoclimatology, Palaeoecology 175:273-302.

Nyakaana, S. & Arctander, P. 1999. *Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: Evidence for male-biased gene flow.* Molecular Ecology 8: 1105-1115.

Page, S.L., Chiu, C.H. & Goodman, M. 1999. *Molecular phylogeny of old-world monkeys (Cercopithecidae) as inferred from γ-Globin DNA sequences.* Molecular Phylogenetics and Evolution 13:348-359.

Palumbi, S.R. 1996. *Chapter 7: Nucleic Acids II: The Polymerase Chain Reaction.* In: Molecular Systematics. Hillis, D.M., Moritz, C & Mable, B.K. (eds.) Sinauer Associates Inc. USA. pp. 205-246.

Payne, J. & Francis, C.M. 1998. *A field guide to the mammals of Borneo.* The Sabah Society, Kota Kinabalu.

Perwitasari-Farajallah, D., Kawamoto, Y. & Suryobroto, B. 1999. *Variation in Blood Proteins and Mitochondrial DNA Within and Between Local Populations of Longtail Macaques, Macaca fascicularis on the Island of Java, Indonesia.* Primates 40:581-596.

Posada, D. & Crandall, K.A. 2001. *Intraspesific gene genealogies: tree grafting into networks*. Trends in Ecology and Evolution 16:37-45.

Pribil, S. & Houlahan, J.E. 2003. *Life-history strategies associated with local population variability confer regional stability*. Proceeding of the Royal Society of London B 270:1419-1423.

Primmer, C.R., Moller, A.P. & Ellengren, H. 1996. *A wide-range survey of cross species microsatellite amplification in birds*. Molecular Ecology 5:365-378.

Purvis, A., Agapow, P., Gittleman, J.L. & Mace, G.M. 2000. *Nonrandom extinction and the loss of evolutionary history*. Science 288:328-330.

Ramos-Onsins, S.E. & Rozas, J. 2002. *Statistical properties of new neutrality tests against population growth*. Molecular Biology and Evolution 19:2092-2100.

Ray, N., Currat, M. & Excoffier, L. 2003. *Intra-Deme Molecular Diversity in Spatially Expanding Populations*. Molecular Biology and Evolution 20:76-86.

Richard, A.F. 1985. *Sympatry, competition, and the niche*. In: Primates in nature. Richard A.F. (ed.). W.H. Freeman and Co.

Rijksen, H.D. & Meijaard, E. 1999. *Our vanishing relative: the status of wild orang-utans at the close of the twentieth century*. Chapman & Hall.

Robertson, J.M.Y. & van Schaik, C.P. 2001. *Causal factors underlying the dramatic decline of the Sumatran orang-utan*. Oryx 35:26-38.

Rodman, P.S. & Mitani, J.C. 1987. *Orangutans: sexual dimorphism in a solitary species*. In: Primate societies. Smuts, B., Cheney, D.L., Seyfarth, R.M., Struhsaker, T. & Wrangham, R.W. (eds.). University of Chicago Press. Chicago. USA. pp.146-154.

Roeder, A. D., Archer, F.I., Poinar, H.N. & Morin, P.A. 2004. *A novel method for collection and preservation of faeces for genetic studies*. Molecular Ecology Notes 4:761-764.

Rogers, A. R. & Harpending, H. 1992 *Population growth makes waves in the distribution of pairwise genetic differences*. Molecular Biology and Evolution 9:552-569.

Rosenblum, L.L., Supriatna, J., Hassan, M.N. & Melnick, D.J. 1997. *High mitochondrial DNA diversity with little structure within and among leaf monkey populations (Trachypithecus cristatus and Trachypithecus auratus)*. International Journal of Primatology 18:1005-1027.

Rozas, J., Sa'nchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. 2003. *DnaSP, DNA polymorphism analyses by the coalescent and other methods*. Bioinformatics 19: 2496-2497.

Salas, A., Lareu, V., Calafell, F., Bertranpetit, J. & Carracedo, A. 2000. *mtDNA hypervariable region II (HVII) sequences in human evolution studies*. European Journal of Human Genetics 8:964-974.

Salter, R.E. & Mackenzie, N.A. 1985. *Conservation status of the proboscis monkey in Sarawak*. Biological Conservation 33:119-132.

Salter, R.E., MacKenzie, N.A., Nightingale, N., Aken, K.M. & Chai, P.K.P. 1985. *Habitat use, ranging behaviour and food habits of the proboscis monkey, Nasalis larvatus (van Wurmb), in Sarawak*. Primates 26:436-451.

Scheffrahn, W., de Ruiter, J.R. & van Hooff, J.A.R.A.M. 1996. *Genetic relatedness within and between populations of Macaca fascicularis on Sumatra and off-shore islands*. In:

Evolution and ecology of macaque societies. Fa, J.E. & Lindburg, D.G. (eds.) Cambridge University Press. Cambridge. UK. pp.20-42.

Sheil, D., Sayer, J.A. & O'Brien, T. *Tree species diversity in logged rainforests.* Science 284:1587a.

Slatkin, M. & Hudson, R.R. 1991. *Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations.* Genetics 129:555-562.

Slatkin, M. 1987. *Gene flow and the geographic structure of natural populations.* Science 236:787-792.

Slatkin, M. 1993. *Isolation by distance in equilibrium and non-equilibrium populations.* Evolution 47:264-279.

Slik, J.W.F., Poulsen, A.D., Ashton, P.S., Cannon, C.H., Eichhorn, K.A.O., Kartawinata, K., Lanniari, I., Nagamasu, H., Nakagawa, M., van Nieuwstadt, M.G.L., Payne, J., Purwaningsih, Saridan, A., Sidiyasa, K., Verburg, R.W., Webb, C.O. & Wilkie, P. 2003. *A floristic analysis of the lowland dipterocarp forests of Borneo.* Journal of Biogeography 30:1517-1531.

Smith, D.G. & McDonough, J. 2005. *Mitochondrial DNA variation in Chinese and Indian rhesus macaques (Macaca mulatta).* American Journal of Primatology 65:1-25.

Sodhi, N.S., Koh, L.P., Brook, B.W. & Ng, P.K.L. 2004. *Southeast Asia biodiversity: an impending disaster.* Trends in ecology and evolution 19:654-660.

Sorensen, M.D., Ast, J.C., Dimcheff, D.E., Yuri, T. & Mindell, D.P. 1999. *Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates.* Molecular Phylogenetic and Evolution 12:105-114.

Sterner, K.N., Raaum, R.L., Zhang, Y., Stewart, C. & Disotell, T.R. 2006. *Mitochondrial data support an odd-nosed colobine clade.* Molecular Phylogenetic and Evolution 40:1-7.

Stibig, H.J. & Malingreau, J.P. 2003. *Forest cover of insular Southeast Asia mapped from remote satellite images of coarse spatial resolution.* AMBIO 32:469-475.

Straub, R.E., Speer, M.C., Luo, Y., Rojas, K., Overhauser, J., Ott, J. & Gilliam, T.C. 1993. *A microsatellite genetic linkage map of human chromosome 18.* Genomics 15:48-56.

Strausberger, B.M. & Ashley, M.V. 2001. *Eggs yield nuclear DNA from egg-laying female cowbirds, their embryos and offspring.* Conservation Genetics 2:385-390.

Stuijts, I.M. 1993. *Late Pleistocene and Holocene vegetation of West Java, Indonesia.* Modern Quarternary Research in Southeast Asia 12:1-173.

Sun, X., Li, X., Luo, Y. & Chen, X. 2000. *The vegetation and climate at the last glaciation on the emerged continental shelf of the South China Sea.* Palaeogeography, Palaeoclimatology, Palaeoecology 160:301-316.

Supriatna, J.A., Yanuar, Martarinza, Wibisono, H.T., 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. & Bouvet, J. 1991. *A single plucked feather as a source of DNA fro bird genetic studies.* The Auk 108:959-960.

Taberlet, P. & Luikart, G. 1999. *Non-invasive genetic sampling and individual identification*. Biological Journal of the Linnean Society 68:41-55.

Tajima, F. 1983. *Evolutionary relationship of DNA sequences in finite populations*. Genetics 105:437-460.

Tajima, F. 1989. *Statistical method for testing the neutral mutation hypothesis by DNA polymorphism*. Genetics 123:585-595.

Tajima, F. 1989. *The effect of change in population size on DNA polymorphism*. Genetics 123:597-601.

Tanaka, H., Roubik, D.W., Kato, M., Liew, F. & Gunsalam, G. 2001. *Phylogenetic position of Apis nuluensis of northern Borneo and phylogeography of A. cerana as inferred from mitochondrial DNA sequences*. Insectes Sociaux 48:44-51.

Tautz, D. & Renz, M. 1984. *Simple sequences are ubiquitous repetitive components of eukaryotic genomes*. Nucleic Acids Research 25:4127-4138.

Tautz, D. 1989. *Hypervariability of simple sequences as a general source for polymorphic DNA markers*. Nucleic Acids Research 17:6463-6471.

Telfer, P.T., Souquière, S., Clifford, S.L., Abernethy, K.A., Bruford, M.W., Disotell, T.R., Sterner, K.N., Roques, P., Marx, P.A. & Wickings, E.J. 2003. *Molecular evidence for deep phylogenetic divergence in Mandrillus sphinx*. Molecular Ecology 12:2019-2024.

Tosi, A. J., Morales, J. C. & Melnick, D. J. 2000. *Comparison of Y-chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history*. Molecular Phylogenetic and Evolution: 17:133–144.

Tougard, C. 2001. *Biogeography and migration routes of large mammal faunas in South-East Asia during the Late Middle Pleistocene: focus on the fossil and extant faunas from Thailand*. Paleogeography, Paleoclimatology, Plaeoecology 168:337-358.

Trigg, S.N., Curran, L.M. & McDonald, A.K. 2006. *Utility of landsat 7 satellite data for continued monitoring of forest cover change in protected areas in southeast Asia*. Singapore Journal of Tropical Geography 27:49-66.

Turner, I.M. & Corlett, R.T. 1996. *The conservation value of small, isolated fragments of lowland tropical rain forest*. Trends in Ecology and Evolution 11:330-333.

Uchida, A. 1998. *Variation in tooth morphology of Pongo pygmaeus*. Journal of Human Evolution 34:71-79.

Umapathy, G., Singh, M. & Mohnot, S.M. 2003. *Status and distribution of Macaca fascicularis umbrosa in the Nicobar Islands, India*. International Journal of the Primatology 24: 281-93.

Utami, S.S. & van Hoof, J.A.R.A.M. 1997. *Meat eating behaviour of adult female orang-utans*. American Journal of Primatology 43:159-165.

Utami, S.S., Goossens, B., Bruford, M.W., de Ruiter, J.R. & van Hoof, J.A.R.A.M. 2002. *Male bimaturism and reproductive success in Sumatran orang-utans*. Behavioural Ecology 13:643-652.

Valsecchi, E., Glockner-Ferrari, D., Ferrari, M. & Amos, W. 1998. *Molecular analysis of the efficiency of sloughed skin sampling in whale population genetics*. Molecular Ecology 7:1419-1422.

Van den Bergh, G.D., de Vos, J. & Sondaar, P.Y. 2001. *The late quartenary palaeogeography of mammal evolution in the Indonesian archipelago.* Palaeogeography, Palaeoclimatology, Palaeoecology 171:385-408.

van Schaik, C.P. 1999. *The socioecology of fission-fusion sociality in Sumatran orangutans.* Primates 40: 73-90.

van Schaik, C.P., Amerongen, A.J.J. van, & Noordwijk, M.A. van 1996. *Riverine refuging in Sumatran longtailed macaques (Macaca fascicularis).* In: Evolution and ecology of macaque societies. Fa, J. & Lindburg, D. (eds.) Cambridge University Press. Cambridge. UK. pp. 160-181.

van Schaik, C.P., Monk, K.A. & Robertson, J.M.Y. 2001. *Dramatic decline in orang-utan numbers in the Leuser Ecosystem.* Oryx 35:14-25.

Verstappen, H.T. 1997. *The effect of climatic change on Southeast Asian geomorphology.* Journal of Quaternary Science 12:413-418.

Vigilant, L. 1999. *An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs.* Biological Chemistry 380:1329-1331.

Vigilant, L., Pennington, R., Harpending, H., Kocher, T.D. & Wilson, A.C. 1989. *Mitochondrial DNA sequences in single hairs from a southern African population.* Proceeding of the National Academy of Science USA 86:9350-9354.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A.C. 1991. *African populations and the evolution of human mitochondrial DNA.* Science 253:1503-1507.

Voris, H.K. 2000. *Maps of the Pleistocene sea levels in Southeast Asia: shorelines, river systems and time duration.* Journal of Biogeography 27:1153-1167.

Waits, L.P. & 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.

Wan, Q., Wu, H., Fujihara, T. & Fang, S. 2004. *Which genetic marker for which conservation genetic issue?* Electrophoresis 25:2165-2176.

Warren, K.S., Nijman, I.J., Lenstra, J.A., Swan, R.A., Heriyanto & den Boer, M. 2000. *Microsatellite DNA variation in Bornean orang-utans (Pongo pygmaeus).* Journal of Medical Primatology 29:57-62.

Warren, K.S., Verschoor, E.J., Langenhuijzen, S., Heriyanto, Swan, R.A., Vigilant, L. & Heeney, J.L. 2001. *Speciation and intrasubspecific variation of Bornean orang-utans, Pongo pygmaeus pygmaues.* Molecular Biology and Evolution 18:472-480.

Wasserman, M.D. & Chapman, C.A.. 2003. *Determinants of colobus monkey abundance: The importance of food energy, protein, and fiber content.* Journal of Animal Ecology 72:650-659.

Wayne, R.K., Bruford, M.W., Girman, D., Rebholz, W.E.E., Sunnucks, P. & Taylor, A.C. 1994. *Molecular genetics of endangered species.* In: Creative conservation. Interactive management of wild and captive animals. Olney, P.J.S., Mace, G.M. & Feistner, A.T.C. (eds.). Chapman & Hall. UK. pp.92-117.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. 2002. *Phylogenies and community ecology.* Annual Review of Ecology and Systematics 33:475-505.

Weber, J.L. & May, P.E. 1989. *Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction*. American Journal of Human Genetics 44:388-396.

Weir, B.S. & Cockerham, C.C. 1984. *Estimating F-statistics for the analysis of population structure*. Evolution 38:1358-1370.

Wennik, P.W., Baker, A.J. & Tilanus, M.G.J. 1994. *Mitochondrial control-region sequences in two shorebird species, the Turnstone and the Dunlin, and their utility in population genetic studies*. Molecular Biology and Evolution 11:22-31.

Wheatley, B.P. 1988. *Cultural behaviour and extractive foraging in Macaca fascicularis*. Current Anthropology 29:516-519.

Wheatley, B.P., Harya-Putra, D.K. & Gonder, M.K. 1996. *A comparison of wild and food-enhanced long-tailed macaques (Macaca fascicularis)*. In: Evolution and ecology of macaque societies. Fa, J. & Lindburg, D. (eds.) Cambridge University Press. Cambridge. UK. pp. 182-206.

Wich, S.A., Singleton, I., Utami-Atmoko, S.S., Geurts, M.L., Rijksen, H.D. & van Schaik, C.P. 2003. *The status of the Sumatran orang-utan Pongo abelii: an update*. Oryx 37:49-54.

Wich, S.A., Utami-Atmoko, S.S., Mitra-Setia, T., Rijksen, H.D., Schürmann, C., van Hooff J.A.R.A.M. & van Schaik, C.P. 2004. *Life history of wild Sumatran orangutans (Pongo abelii)*. Journal of Human Evolution 47:385-398.

Wilson, M.E.J. & Moss, S.J. 1999. *Cenozoic evolution of Borneo-Sulawesi*. Palaeogeography, Palaeoclimatology and Palaeooceanography. 145:303-337.

Winney, B.J., Hammond, R.L., Macasero, W., Flores, B., Boug, A., Biquand, V. Biquand, S. & Bruford, M.W. 2004. *Crossing the Red Sea: phylogeography of the hamadryas babbon, Papio hamadryas hamadryas*. Molecular Ecology 13:2819-2827.

Wolfheim, J.H. 1983. *Primates of the world: distribution, abundance and conservation*. Seattle. University of Washington Press.

Wright, S.J. 2005. *Tropical forests in a changing environment*. Trends in Ecology and Evolution 20:553-560.

WWF. 1998. *Kinabatangan*. Joint Publication GTZ, WWF and Sabah Government.

Xia, X. & Xie, Z. 2001. *DAMBE: Data analysis in molecular biology and evolution*. Journal of Heredity 92:371-373.

Xu, X. & Arnason, U. 1996. *The mitochondrial DNA molecule of Sumatran and a molecular proposal for two (Bornean and Sumatran) species of orang-utan*. Journal of Molecular Evolution 43:431-437.

Yeager, C.P. 1989. *Feeding ecology of the proboscis monkey (Nasalis larvatus)*. International Journal of Primatology 10:497-529.

Yeager, C.P. 1992. *Proboscis monkey (Nasalis larvatus) social organization: Nature and possible functions of intergroup patterns of association*. American Journal of Primatology 26:133-137.

Yeager, C.P. 1995. *Does intraspecific variation in social systems explain reported differences in the social structure of the proboscis monkey (Nasalis larvatus)?* Primates 36:575-582.

Yeager, C.P., Silver, S.C. & Dierenfeld, E.S. 1997. *Mineral and phytochemical influences on foliage selection by the proboscis monkey (Nasalis larvatus).* American Journal of Primatology 41:117-128.

Zane, L., Bargelloni, L. & Patarnello, T. 2002. *Strategies for microsatellite isolation: a review.* Molecular Ecology 11:1-16.

Zhan, X., Li, M., Zhang, Z., Goossens, B., Chen, Y., Wang, H., Bruford, M.W. & Wei, F. 2006. *Molecular censuring doubles giant panda population estimate in a key nature reserve.* Current Biology 16:451-452.

Zhang, Y., Ryder, O.A. & Zhang, Y. 2001. *Genetic divergence of orang-utan subspecies (Pongo pygmaeus).* Journal of Molecular Evolution 52:516-526.

Zhi, L., Karesh, W.B., Janczewski, D.N., Frazier-Taylor, H., Sajuthi, D., Gombek, F., Andau, M., Martenson, S. & O'Brien, S.J. 1996. *Genomic differentiation among natural populations of orang-utan (Pongo pygmaeus).* Current Biology 6:1326-1336.

Appendix I
Map of Sabah showing the forest reserves and conservation areas. The colour indicates the type of forest reserves as indicated by the Legends.
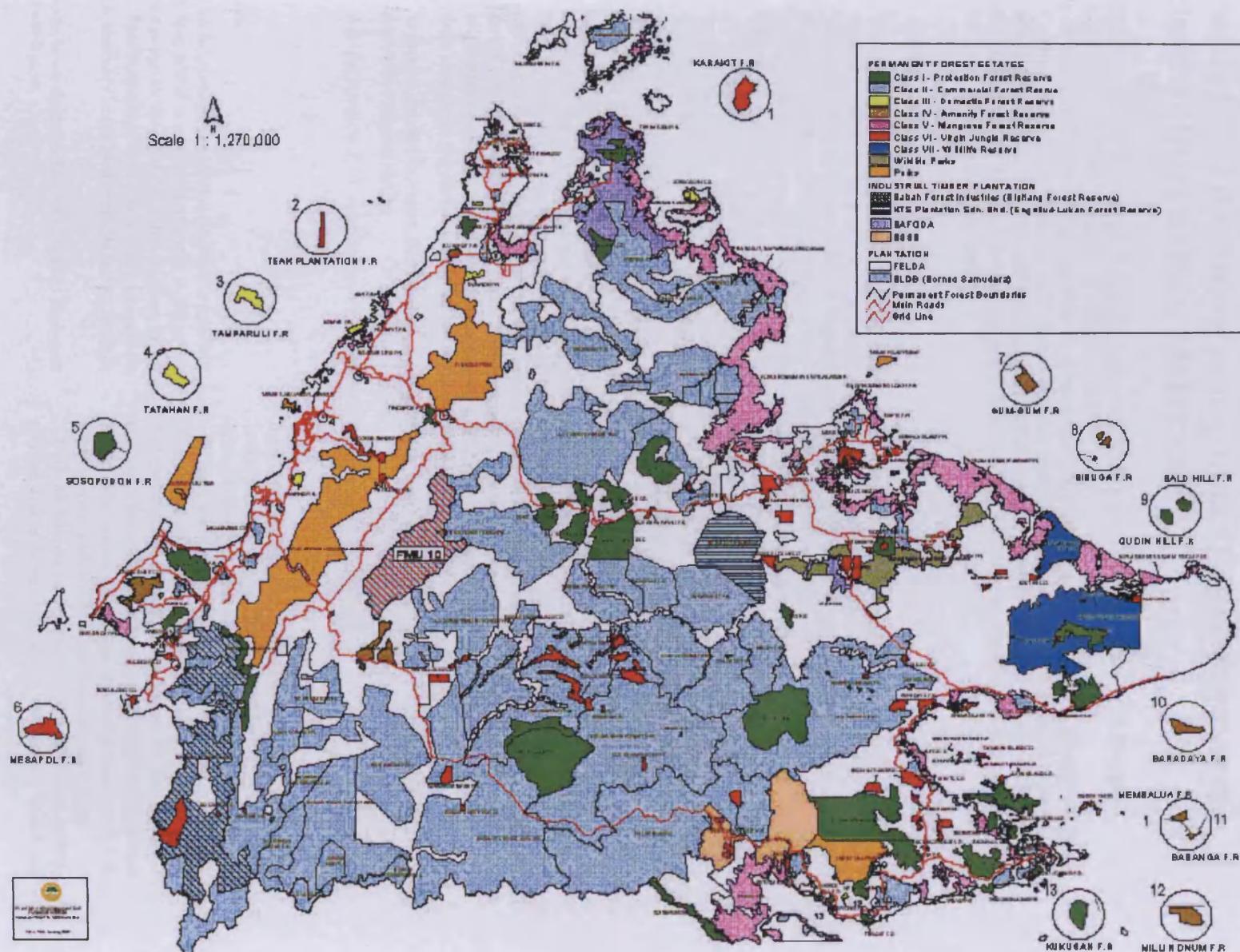
Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia

B. GOOSSENS,*†‡§ L. CHIKHI,§ M. F. JALIL,*‡ M. ANCRENAZ,† I. LACKMAN-ANCRENAZ,†
M. MOHAMED,‡ P. ANDAU¶ and M. W. BRUFORD*
*Biodiversity and Ecological Processes Group, Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3TL, UK, †Kinabatangan Orang-utan Conservation Project, PO Box 3109, 90734 Sandakan, Sabah, Malaysia, ‡Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Locked Bag 2073, 88999 Kota Kinabalu, Sabah, Malaysia, §UMR 5174 Evolution et Diversité Biologique, Bâtiment 4R3b2, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cédex 4, France, ¶Sabah Wildlife Department, Muis Complex, 88100 Kota Kinabalu, Sabah, Malaysia

Abstract

We investigated the genetic structure within and among Bornean orang-utans (*Pongo pygmaeus*) in forest fragments of the Lower Kinabatangan flood plain in Sabah, Malaysia. DNA was extracted from hair and faecal samples for 200 wild individuals collected during boat surveys on the Kinabatangan River. Fourteen microsatellite loci were used to characterize patterns of genetic diversity. We found that genetic diversity was high in the set of samples (mean H_E = 0.74) and that genetic differentiation was significant between the samples (average F_{ST} = 0.04, $P < 0.001$) with F_{ST} values ranging from low (0.01) to moderately large (0.12) values. Pairwise F_{ST} values were significantly higher across the Kinabatangan River than between samples from the same river side, thereby confirming the role of the river as a natural barrier to gene flow. The correlation between genetic and geographical distance was tested by means of a series of Mantel tests based on different measures of geographical distance. We used a Bayesian method to estimate immigration rates. The results indicate that migration is unlikely across the river but cannot be completely ruled out because of the limited F_{ST} values. Assignment tests confirm the overall picture that gene flow is limited across the river. We found that migration between samples from the same side of the river had a high probability indicating that orang-utans used to move relatively freely between neighbouring areas. This strongly suggests that there is a need to maintain migration between isolated forest fragments. This could be done by restoring forest corridors alongside the river banks and between patches.

Keywords: genetic diversity, immigration, microsatellites, noninvasive sampling, *Pongo pygmaeus*, population fragmentation

Received 3 August 2004; revision received 19 October 2004; accepted 22 October 2004

Introduction

Current trends in great ape populations indicate a dramatic ongoing decline, which is predicted to result in the extinction of ape species in the wild for entire regions in the near future. Recent findings have particularly focused on African apes, and have implicated multiple factors, such as

Correspondence: Benoît Goossens, Fax: + 33 5 61 55 73 27; E-mail: goossensbr@cardiff.ac.uk

deforestation, hunting and disease (Walsh *et al.* 2003; Leendertz *et al.* 2004; Leroy *et al.* 2004). Less well publicised, but equally dramatic, has been the decline in Asia's only great ape, the orang-utan species of Sumatra and Borneo (*Pongo abelii* and *Pongo pygmaeus*). Current trends suggest that extinction is potentially imminent for the Sumatran species in the wild and although anthropogenic pressures are equally severe in parts of the orang-utan's range in Borneo, some potentially viable populations remain.

On both islands, orang-utans exist now mainly in fragmented and isolated populations, the sizes of which are

only now being accurately estimated. Using wild reserve data, Rijksen & Meijaard (1999) estimated that the number of orang-utans may have dropped from *c.* 315 000 around 1900 to *c.* 27 000 in 1997 and have recently been estimated to be as low as 3500 in Sumatra (Wich *et al.* 2003). In Borneo, orang-utans appear to be widely distributed across Indonesia (Central, West, and East Kalimantan) and Malaysia (Sarawak and Sabah). Still, the situation appears critical as the population is estimated to have dropped from 23 000 in 1996 to 15 000 individuals in 1997 (a reduction of some 33% in one year, Rijksen & Meijaard 1999) because of drought and fires. Despite the uncertainties existing on these population size estimates (see Payne 1987, 1988), there is general agreement that populations have decreased at least 10-fold in the last 100 years (Delgado & van Schaik 2000).

There appears to be several causes for the dramatic decline of the Bornean orang-gutan in the last century, but it is known that processes threatening orang-utan populations include hunting, habitat loss, habitat degradation and forest fragmentation (Delgado & van Schaik 2000; Robertson & van Schaik 2001). In the Malaysian state of Sabah in the northern part of Borneo, human pressure has steadily increased. Indeed, in the last 20 years very large areas of forest have been logged and converted into oil palm plantations. In 2001, it was estimated that between a third and half of the original forest area had disappeared (McMorrow & Talip 2001).

Throughout their range, great apes, in common with many other species are increasingly affected by anthropogenic forest fragmentation and this problem is a major issue in Malaysia and Borneo (e.g. Laidlaw 2000; Kinnaird *et al.* 2003). Conservation planning for these increasingly isolated forest fragment communities presents a demanding set of challenges, ranging from determining viable population sizes (which may be in the tens of thousands for orang-utans, Harcourt 2002), assessing the potential for and importance of dispersal among populations (e.g. Travis & Dytham 1999) and estimating the relative importance of different ecological and life history parameters in predicting extinction risk (e.g. Brashares 2003). Large-bodied, slow-reproducing species, such as the orang-utan, have been shown in many studies to be more prone to extinction (Webb *et al.* 2002; Cardillo 2003), especially in closed or fragmented habitats (e.g. Laurance 1991; Davies *et al.* 2000). However, predicting population persistence is complex and, for example, those species which are able to effectively utilize modified habitats may remain stable or even increase in fragmented landscapes (Laurance 1991). Further, dispersal behaviours may modify according to habitat availability and persistence, potentially affecting, for example, predictions of extinction/recolonization in metapopulations (Travis & Dytham 1999; Reed 2004).

One trait which has rarely been examined in large-bodied species living in fragmented populations is genetic diversity.

While the link between genetic diversity and population persistence has been demonstrated in smaller bodied vertebrates, invertebrates and plants (e.g. Saccheri *et al.* 1998; Madsen *et al.* 1999; Pryor *et al.* 2001) studies are much less common in large vertebrates, probably because of the fact that their slower vital rates are not expected to result in measurable reductions in genetic diversity over the timescale (in generations) of most anthropogenic habitat fragmentation (but see, for example, Miller & Waits 2003).

However, while genetic diversity may not obviously be affected in such species using standard gene diversity measures, such effects may be discernible in the genealogical data found in their allele distributions (e.g. O'Ryan *et al.* 1998; Goossens *et al.* submitted) and, regardless, genetic diversity in present day populations still needs to be managed judiciously in these species to guarantee their persistence in the future. For example, in the absence of direct behavioural observation, genetic methods can be used to infer dispersal and immigration events which can have profound consequences for population viability (e.g. Keller *et al.* 2001), allow the assignment of sexed individuals to their natal populations (e.g. Berry *et al.* 2004; Möller & Beheregaray 2004) and permit the development of a better understanding of how geographical features in different landscapes correlate with dispersal and genetic differentiation among local populations (e.g. Palsson 2004).

Orang-utans are large-bodied, semisolitary and slow-reproducing species, with extreme sexual dimorphism in body size and appearance. Orang-utans also show a pronounced bimaturism among sexually mature males and matings seem to be promiscuous, with both morphs (flanged and unflanged males) being reproductively successful in the populations (Rodman & Mitani 1987; Delgado & van Schaik 2000; Utami *et al.* 2002). Sexual maturity is variable and difficult to determine in the wild, particularly for males. In females it may vary between 7 and 15 years and is probably greater than 10 years in males (Leighton *et al.* 1995; Delgado & van Schaik 2000). Females care for dependent offspring for at least six years and the interbirth interval is about 8 years (Leighton *et al.* 1995). Slow growth and development contribute to a long lifespan, estimated to be about 45 years for both sexes in the wild (Leighton *et al.* 1995). Little is known about dispersal. Maturing females tend to remain near the natal area (philopatry), while males move away (Mitani 1989; Galdikas 1995; Singleton & van Schaik 2001). However, they seem to be very poor dispersers and they can be confined in isolated populations (van Schaik *et al.* 2001).

Within this context we studied the genetic diversity of an important remaining orang-utan 'population' in Sabah, found in the forests of the Lower Kinabatangan flood plain (*c.* 1100 individuals, Ancrenaz *et al.* 2004). In this area, conversion of forest into oil palm plantations has resulted in a highly fragmented forest structure (Fig. 1, Rijksen & Meijaard
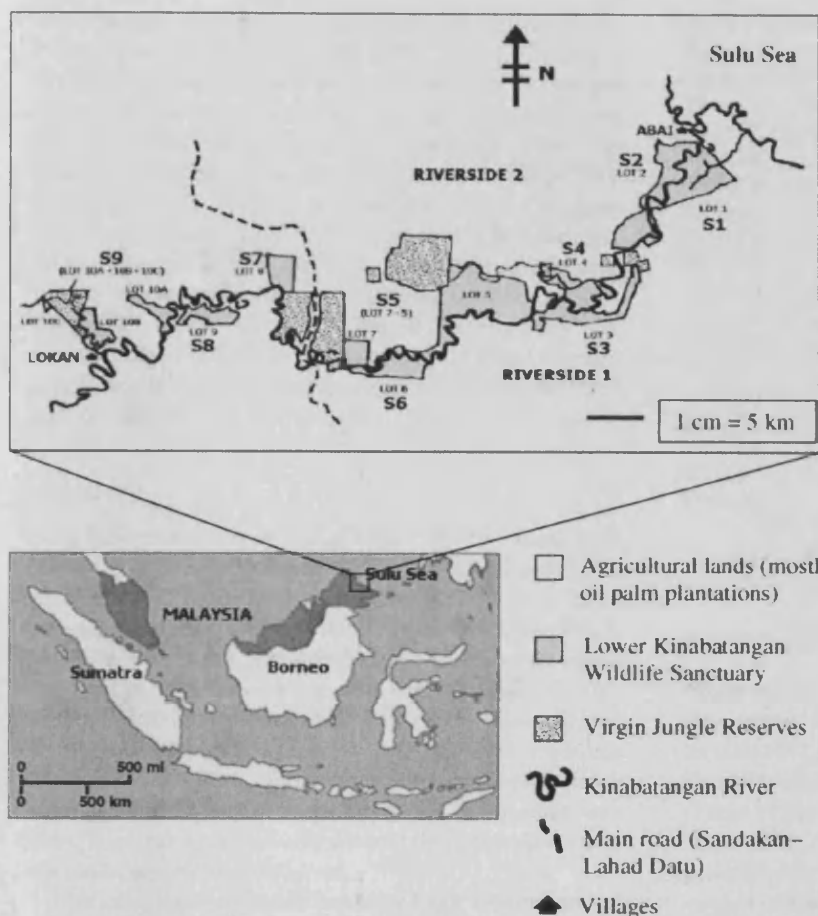
Fig. 1 Map of the Lower Kinabatangan Wildlife Sanctuary (LKWS) showing the location of the 10 lots of forests and the virgin jungle forest reserves alongside the Kinabatangan River. The inside map shows the location of the LKWS in Borneo Island.

1999; McMorrow & Talip 2001). In 2002, the state government of Sabah gazetted 27 000 ha of these forests as a wildlife sanctuary, with the ultimate aim of creating a corridor for wildlife along the Lower Kinabatangan flood plain, between the remaining virgin forest reserves. The impact of habitat fragmentation on the long-term survival of isolated orangutan subpopulations is the main focus of current ecological and behavioural surveys in the region (Lackman-Ancrenaz *et al.* 2001). Population densities are unusually high for secondary forest, perhaps a result of recent habitat loss and consecutive concentration of individuals in the remaining forests (Ancrenaz *et al.* 2004).

While Bornean orang-utans have already been genetically studied (e.g. Zhi *et al.* 1996; Warren *et al.* 2000, 2001), the present study is the first to be carried out on wild animals (as opposed to individuals mostly sampled in zoos or in and around rehabilitation centres). Two other important specificities of the present work are: (i) the large number of individuals and loci typed (14 loci typed for 200 individuals), and (ii) the high proportion — *c.* 20% — that these individuals represent compared to the estimated number

of individuals present in the sampled region (Ancrenaz *et al.* 2004). Specifically, we examine genetic structure within and among the remaining sampled forest fragments and determine the effect of natural barriers such as the Kinabatangan River, on isolation. We estimated diversity within and migration rates among forest patches on the same and different sides of the river. Our analysis includes an attempt to assess the genetic effects of both past and ongoing dispersal. The applicability of these data to be incorporated in conservation assessment in a management plan for *P. pygmaeus* in the region is discussed.

Materials and methods

The Lower Kinabatangan flood plain and the Lower Kinabatangan Wildlife Sanctuary

The Lower Kinabatangan flood plain (5°20′–5°45′ N, 117°40′–118°30′ E) is located in eastern Sabah, Malaysia. The flood plain is a patchwork of different habitat types: riverine forest, seasonally flooded forest, swamp forest,

dry dipterocarp forest, nipa palms, and mangrove (Azmi 1998). However, since the mid 1950s, the whole Lower Kinabatangan region has been subjected to large-scale commercial timber exploitation and agriculture. During the past 20 years, postlogging land conversion to oil palm plantations has been extensive (McMorrow & Talip 2001).

On 16 January 2002, the proposed Lower Kinabatangan Wildlife Sanctuary (LKWS) was gazetted and now comprises 10 sectors or lots (lots 1–10, with lot 10 divided into 10A–C) chosen to increase connectivity between remaining forest reserves (Fig. 1). The aim of this sanctuary is to transform the 27 000 ha of flood plain into a forest corridor connecting the coastal mangrove swamps with dry land forests upriver.

Sampling

Shed hair in nests and faeces were collected from wild orang-utans during boat surveys carried out alongside the Kinabatangan River (between Abai village and Lokan village, corresponding to a 280 km river tract, see Fig. 1) between January and May 2001. When a fresh nest (between one and five days old – see Goossens et al. 2004) was spotted, shed hairs were collected. Shed hairs were also collected during line transects made to estimate nest densities (Ancrenaz et al. 2004). Faecal samples found below fresh nests were collected as well. When an orang-utan was directly encountered it was followed until defecation and the faecal sample was collected.

Hair samples were stored in plastic bags, whereas faecal samples were stored in 50 mL BD Falcon™ tubes with 90% ethanol. Precautions were taken to avoid human contamination during the sampling by using sterile gloves and implements (sterilized forceps). GPS coordinates were taken for each sample.

Shed hairs from 176 different nests, and faecal samples from 71 orang-utans were collected and could be assigned to nine sampling regions S1–S9 (Fig. 1), which corresponds mostly to the lots described above, except that samples in lots 5 and 7 were grouped into S5. In the 32 cases where faecal samples were collected below fresh nests, they were used instead of shed hairs collected in the nest. Thus, of a total of 279 samples collected, 247 samples were selected for genetic analyses.

DNA extraction

For shed hairs (144 samples), DNA was extracted using a polymerase chain reaction (PCR) buffer method (Vigilant 1999). Faecal extractions (103 samples) were carried out in a Class I microbiological safety hood, using the QIAamp DNA Stool Mini Kit (QIAGEN) and following a protocol for orang-utans detailed in Goossens et al. (2000a) and Utami et al. (2002).

Table 1 Characteristics of 14 human-derived microsatellite loci used in *Pongo pygmaeus*

Locus ID	T_a (°C), time (s)	Size (bp)
D5S1457	49, 45	111–139
D5S1470	51, 30	208–236
D1S550	60, 30	128–166
D2S1326	60, 30	200–224
D3S2459	60, 45	200–216
D4S1627	55, 30	188–208
D4S2408	64, 45	274–306
D5S1505	64, 30	211–243
D6S501	60, 30	153–181
D13S321	60, 45	200–216
D13S765	60, 45	185–205
D12S375	60, 30	172–188
D2S141	64, 45	138–150
D16S420	58, 30	178–194

T_a = optimal PCR annealing temperature.

DNA amplification and microsatellite genotyping

Fourteen human-derived microsatellite loci were used: 2 dinucleotide loci D2S141 and D16S420; and 12 tetranucleotide loci D5S1457, D5S1470, D1S550, D2S1326, D3S2459, D4S1627, D4S2408, D5S1505, D6S501, D13S321, D13S765 and D12S375 (Table 1) (see also Coote & Bruford 1996; Goossens et al. 2000b, 2002; Zhang et al. 2001; Utami et al. 2002). All forward primers were fluorescently labelled. All PCR reactions were carried out in 12.5 μL total containing 2.5 μL DNA extract. A multiple-tube procedure was conducted for each faecal extract according to Taberlet et al. (1996). For each extract, three amplifications were performed using the D5S1457 locus (Goossens et al. 2000a). After that, the most successful extract (three positive PCRs) for each sample was amplified seven times for each locus to avoid typing errors (see Taberlet et al. 1999 for a review). Amplifications were carried out in 12.5 μL [10 mM Tris-HCl (pH 9.0), 200 mM (NH$_4$)$_2$SO$_4$, 50 μM each dNTP, 1.5 mM MgCl$_2$, 5 ng of BSA, 0.1 U AmpliTaq Gold DNA polymerase (Perkin Elmer), 0.5 μM reverse primer, 0.5 μM fluorescent (TET, FAM or HEX) forward primer, 2.5 μL of DNA extract]. PCR amplification of 50 cycles was carried out for each locus separately (initial denaturation 94 °C for 10 min, 94 °C for 15 s, 45 °C to 52 °C for 15–30 s, 72 °C for 30–60 s). The annealing temperature was optimized for each locus (Table 1). All PCR products were separated on an acrylamide gel using an ABI PRISM 377 DNA sequencer. Gels were analysed using GENESCAN ANALYSIS 2.0 and GENOTYPER 1.1 software.

Genetic diversity and population structure

Genetic diversity was measured as the mean number of alleles per locus (MNA), observed (H_O), and expected (H_E)

heterozygosities (Nei 1978). Linkage disequilibrium (LD) was estimated across all pairs of loci using the correlation coefficient of Weir (1979). A permutation approach was used to determine which LD values were significant. Wright's F statistics were estimated according to Weir & Cockerham (1984) and their departure from the null hypothesis (no genetic differentiation for F_{ST}, and Hardy–Weinberg equilibrium for F_{IS} and F_{IT}) was tested using permutations. Analyses were performed using the GENETIX software (Belkhir *et al.* 1996/1997, available at http://www.University-montp2.fr/~genetix/genetix/genetix.htm).

Differentiation between river sides

As the Kinabatangan River represents a natural barrier to the movement of orang-utans, we divided the samples into two sets called River Side 1 — RS1 = (S1, S3, S6, S8) — and River Side 2 — RS2 = (S2, S4, S5, S7, and S9). We then looked at the distribution of pairwise F_{ST} values between samples belonging (i) to the same side of the river (RS1 vs. RS1 and RS2 vs. RS2) and (ii) to different sides (RS1 vs. RS2). We then compared them to the set of all pairwise F_{ST} values (i.e. regardless of the river side). For simplicity of notation we shall refer to these sets as $F_{ST(RS1)}$, $F_{ST(RS2)}$, $F_{ST(RS1-2)}$, and $F_{ST(TOT)}$, respectively. We were interested in two different statistical tests. In the first, we compared the $F_{ST(RS1)}$, $F_{ST(RS2)}$ and $F_{ST(RS1-2)}$ distributions to the $F_{ST(TOT)}$ distribution. This allowed us to determine whether each subset was significantly different from a random subsample of all the F_{ST} values. This was done by repeatedly permuting the set $F_{ST(TOT)}$ values to create three sets of F_{ST} values containing the same number of F_{ST} values as $F_{ST(RS1)}$, $F_{ST(RS2)}$ and $F_{ST(RS1-2)}$, respectively. At each permutation and for each set we calculated the mean F_{ST} value. The observed (real) average F_{ST} of each subset was then compared to the distributions obtained. In the second set of tests, we compared the $F_{ST(RS1)}$ and $F_{ST(RS2)}$ distributions to the distribution of $F_{ST(RS1-2)}$ values. The randomization was done by sampling from the distribution of $F_{ST(RS1-2)}$ values one subset of F_{ST} values with the same size as the $F_{ST(RS1)}$ (or $F_{ST(RS2)}$) set. The distribution of means was then compared to the real mean. This allowed us to test whether F_{ST} values within each river side were significantly lower than those observed between river sides.

In order to further assess the effect of the river in the patterns of genetic differentiation and to test the correlation between genetic and geographical distance, a series of two-way Mantel tests was carried out (Mantel 1967). In each of these tests the matrix of pairwise F_{ST} values was used against four different matrices of geographical distances. The four geographical distances were built in order to account for the potential role of the Kinabatangan as a geographical barrier. In the first case, the river was ignored and a simple Euclidian distance was computed among all

samples. In the three other cases three different assumptions were made regarding the point at which orang-utans were possibly able to cross the river, whereas distances between samples from the same side were computed by following the river. The three hypothetical crossing points were assumed to be (i) at the level of S8 and S9 (the most upstream samples used) where the river is approximately 200 m wide, (ii) approximately 150 km upstream of S8 and S9, which is probably the closest location where the river is narrow enough to allow orang-utans to use fallen trees to cross the river, and (iii) at the Kinabatangan source, approximately 260 km upstream of S8 and S9. All the permutation tests above, including the Mantel test, were performed using the R statistical package.

Immigration between river sides

Wilson & Rannala (2003) recently developed a Bayesian method to estimate rates of recent immigration in a set of linked populations using multilocus data. The method is based on a simple model where individuals are exchanged between populations over generations. The probability of observing a particular genotype in a given population can be expressed as a function of the model's demographic parameters (this probability is the likelihood). These parameters include the allele frequencies, the immigration rates (m_{ij}, the proportion of individuals in population j that originate from population i), the inbreeding level in each population (F_i being the inbreeding in population i), and the time at which the immigration event took place (the method currently accounts for immigration events taking place either at the sampling generation t_1, or one generation before, t_2). Based on this likelihood function, Wilson & Rannala (2003) use a Markov chain Monte Carlo (MCMC) approach to explore the parameter space and obtain samples from the posterior distributions of the parameters of interest. One interesting property of this method is that, contrary to most methods currently available, it does not require samples to be at Hardy–Weinberg equilibrium (HWE). Also, an advantage over assignment methods is that migration events are accounted for in the calculation of allele frequencies, and hence in the likelihood. This is typically ignored by assignment methods. Finally, it is important to note that the method allows to estimate immigration rather than migration rates.

The method is implemented in the software BAYESASS (http://www.rannala.org/labpages/software.html). The software allows the user to change parameters affecting the proposal distributions, namely *deltap*, *deltam*, and *deltaF*, which define the manner in which the parameter space is explored during the MCMC (details on the proposal distributions can be found in Wilson & Rannala 2003). Using different values as we did can be crucial as some choices could produce sticky Markov chains that take a long time to

converge (e.g. Gilks et al. 1996). Different summaries of MCMC runs and some parameter distributions can be saved and examined. One reason for saving only summaries rather than the chains is that the number of parameters of potential interest can grow very quickly with the number of populations. However, we found that not having access to the chains could be problematic (see below). In particular, it can be critical to check that equilibrium has been reached before using summaries (e.g. Chikhi et al. 2001). The code was therefore modified and recompiled to produce an output with the F_i and the m_{ij}, thus producing outputs with $2n$ columns for data from n populations (i.e. for each step of the MCMC a line with the n following numbers is produced: $F_1 \ldots F_n\ m_{11}\ m_{22} \ldots m_{nn}$). The modifications of the code and the corresponding executable were provided to G. Wilson and can be obtained from him. The results presented here for the comparison of the two river sides were obtained from five independent runs with different starting values for the random number generator and different values of the proposal distributions (the parameters deltap, deltam and deltaF varied between 0.05 and 0.35). Wilson & Rannala (2003) ran their data for 3×10^6 iterations discarding the first 10^6 as a burn-in. We first used these conditions but decided to increase the total number of iterations to check convergence on very long runs. The number of iterations were 10, 15 (two runs) and 20×10^6. In order to test whether our modifications of the code did not make any change, the data were run for 20×10^6 iterations and the results compared to the original ones from the 5 runs. They were not distinguishable from them.

Assignment tests

Assignment tests were also performed on the data to determine whether it was possible to assign individuals to their population or river side of origin. Different approaches have been proposed to estimate assignment probabilities (e.g. Paetkau et al. 1995; Cornuet et al. 1999). We applied the method of Rannala & Mountain (1997) because it has been shown to provide the best assignment results by simulations (Cornuet et al. 1999). Note that most methods usually provide similar results unless sample sizes (in terms of individuals and number of loci typed) are small. Because we are interested in the role of the Kinabatangan River as a barrier to orang-utans movement, the exact method chosen was not crucial. This point is discussed below (see Cegelski et al. 2003; Berry et al. 2004; and references therein for in-depth analysis and comparison of assignment methods). Rannala and Mountain's method is implemented in the GENECLASS software. Simulations were used to rank 'populations' (i.e. samples S1 to S9) and determine the most probable sources for all 200 individuals. For each individual, we then checked whether the most probable, second

most probable, ... sample was from the same river side as the individual analysed. Using only the most probable source is not necessarily a good choice given that the second most probable could be from the other side. We thus decided to apply a 'majority rule' algorithm and check the most probable river side among the k most probable samples. The value of k could in principle be any value between 1 and 9. However, the value $k = 1$ corresponds to choosing the most probable sample. Taking $k = 9$ will lead us to take all populations which would not make sense either. Given that there are four samples from RS1 and five from RS2, even if the assignment was perfect, the most probable sample from the opposite river side would necessarily appear on the 5th and 6th rank, respectively. This also means that the majority rule must account for the higher probability of having an individual assigned to RS2 by chance. For individual from RS1 the majority rule applies if more than 4/9th of k comes from RS1, whereas for individuals from RS2, the rule applies if more than 5/9th of k individuals are assigned to RS2. Because k cannot be too small or too large for the reasons given above, we decided to apply the majority rule to the first five and six samples in the assignment ranking.

Results

Genetic typing

We were able to reliably amplify DNA from 201 out of 247 samples. Two individuals had the same genotype at all 14 loci and corresponded to samples taken from two fresh nests c. 100 m apart. With the exclusion of this case, no other pair of samples had the same genotypes at the 14 loci. We found one pair of individuals identical at 13 loci and another pair identical for 12 loci. Two pairs were identical at 11 loci and three pairs had 10 loci in common. While we cannot exclude the possibility that the two genotypes above were from different individuals, they were considered to be from the same individual, leaving a total of 200 different individuals, corresponding to a total of $200*14 = 2800$ genotypes. Of these only seven genotypes (0.25%) were not reliable using the multiple-tubes approach and were therefore coded as missing genotypes.

Genetic diversity, departure from Hardy–Weinberg equilibrium (HWE) and LD

All the 14 loci used in the study were polymorphic, with between five and nine alleles per locus across all samples (Table 2). The mean number of alleles (MNA) per locus ranged between 4.1 (S3) and 4.9 (S4, S5 and S9); the lowest was 3.3 (S7, which only has a sample size of five). Average H_E values were high (0.66–0.75, Table 2). Average H_O values were slightly higher (0.67–0.77), generating slightly

Table 2 Average number of alleles across samples (N_a), observed (H_O) and expected (H_E) heterozygosities and departures from Hardy–Weinberg proportions (F_{IS}) for all samples and for all loci

Locus	Sample n	S1 27	S2 26	S3 22	S4 20	S5 27	S6 33	S7 5	S8 24	S9 16	N_a
D5S1457	H_E	0.766	0.773	0.637	0.663	0.718	0.734	0.778	0.699	0.804	7
	H_O	0.852	0.654	0.727	0.600	0.593	0.818	1.000	0.667	0.750	
	F_{IS}	0.115	0.157	−0.145	0.097	0.178	−0.116	−0.333	0.047	0.070	
		NS	NS	NS	NS	NS	NS	***	NS	NS	
D5S1470	H_E	0.750	0.724	0.802	0.685	0.744	0.733	0.644	0.756	0.752	8
	H_O	0.963	0.769	0.727	0.800	0.667	0.758	0.600	0.750	0.562	
	F_{IS}	0.290	−0.064	0.096	−0.174	0.105	−0.034	0.077	0.008	0.258	
		***	NS	NS	NS	NS	NS	NS	NS	NS	
D1S550	H_E	0.757	0.744	0.646	0.745	0.789	0.772	0.689	0.551	0.754	7
	H_O	0.778	0.692	0.773	0.750	0.556	0.697	0.600	0.458	0.750	
	F_{IS}	0.027	0.070	−0.202	−0.007	0.300	0.098	0.143	0.172	0.006	
		NS	NS	NS	NS	**	NS	NS	NS	NS	
D2S1326	H_E	0.746	0.762	0.780	0.700	0.723	0.792	0.711	0.840	0.720	7
	H_O	0.815	0.731	0.864	0.526	0.556	0.849	0.800	0.708	0.625	
	F_{IS}	0.094	0.041	−0.110	0.253	0.235	−0.072	−0.143	0.159	0.135	
		NS	NS	NS	NS	*	NS	NS	NS	NS	
D3S2459	H_E	0.640	0.782	0.624	0.696	0.735	0.731	0.733	0.726	0.740	5
	H_O	0.778	0.654	0.682	0.600	0.630	0.697	0.800	0.667	0.750	
	F_{IS}	0.220	0.167	−0.096	0.141	0.146	0.047	−0.103	0.083	−0.014	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
D4S1627	H_E	0.665	0.647	0.680	0.768	0.693	0.717	0.689	0.719	0.796	6
	H_O	0.815	0.769	0.727	0.750	0.704	0.758	0.800	0.875	0.812	
	F_{IS}	0.231	−0.193	−0.072	0.024	−0.016	−0.058	−0.185	−0.223	−0.021	
		**	NS	NS	NS	NS	NS	NS	*	NS	
D4S2408	H_E	0.658	0.624	0.537	0.735	0.639	0.693	0.533	0.662	0.704	5
	H_O	0.630	0.577	0.636	0.700	0.741	0.667	0.800	0.708	0.625	
	F_{IS}	0.043	0.076	−0.190	0.048	−0.162	0.039	−0.600	−0.071	0.115	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
D5S1505	H_E	0.736	0.768	0.724	0.817	0.778	0.788	0.600	0.840	0.827	9
	H_O	0.926	0.769	0.773	0.800	0.808	0.758	0.800	0.792	0.625	
	F_{IS}	0.265	−0.001	−0.069	0.021	−0.039	0.039	−0.391	0.059	0.250	
		**	NS	NS	NS	NS	NS	***	NS	**	
D6S501	H_E	0.547	0.622	0.506	0.760	0.698	0.626	0.622	0.622	0.706	8
	H_O	0.556	0.654	0.545	0.850	0.704	0.636	0.600	0.708	0.875	
	F_{IS}	0.016	−0.052	−0.079	−0.122	−0.008	−0.017	0.040	−0.143	−0.250	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
D13S321	H_E	0.800	0.737	0.627	0.796	0.792	0.792	0.600	0.798	0.772	5
	H_O	0.808	0.731	0.636	0.800	0.741	0.758	0.800	0.792	0.875	
	F_{IS}	0.010	0.008	−0.016	−0.005	0.066	0.044	−0.391	0.008	−0.138	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
D13S765	H_E	0.752	0.581	0.688	0.717	0.705	0.707	0.467	0.728	0.514	6
	H_O	0.731	0.577	0.773	0.800	0.667	0.697	0.600	0.750	0.438	
	F_{IS}	0.029	0.007	−0.126	−0.120	0.055	0.015	−0.333	−0.031	0.153	
		NS	NS	NS	NS	NS	NS	***	NS	NS	
D12S375	H_E	0.639	0.606	0.575	0.679	0.680	0.671	0.733	0.662	0.724	5
	H_O	0.741	0.731	0.455	0.850	0.704	0.697	0.800	0.833	0.812	
	F_{IS}	0.162	−0.212	0.213	−0.259	−0.036	−0.039	−0.103	−0.265	−0.127	
		NS	NS	NS	*	NS	NS	NS	*	NS	
D2S141	H_E	0.677	0.594	0.809	0.768	0.722	0.760	0.778	0.802	0.782	6
	H_O	0.704	0.577	0.818	0.800	0.778	0.758	0.800	0.913	0.812	
	F_{IS}	0.040	0.028	−0.012	−0.043	−0.079	0.004	−0.032	−0.142	−0.040	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
D16S420	H_E	0.686	0.664	0.638	0.759	0.804	0.744	0.778	0.784	0.855	8
	H_O	0.680	0.577	0.727	0.800	0.852	0.697	1.000	0.833	0.938	
	F_{IS}	0.009	0.133	−0.143	−0.056	−0.060	0.064	−0.333	−0.065	−0.100	
		NS	NS	NS	NS	NS	NS	NS	NS	NS	
Total											
	H_E	0.701	0.688	0.662	0.735	0.730	0.733	0.668	0.728	0.746	
	H_O	0.770	0.676	0.705	0.745	0.693	0.732	0.771	0.747	0.732	
		0.099	0.017	−0.065	−0.014	0.052	0.002	−0.177	−0.027	0.020	
	F_{IS}	***	NS	NS	NS	NS	NS	*	NS	NS	
	MNA	4.500	4.714	4.143	4.857	4.857	4.714	3.286	4.786	4.929	

NS = nonsignificant, * = $P < 0.05$, **$P < 0.01$, ***$P < 0.001$. n = sample size.

Table 3 Pairwise F_{ST} values

	S1	S2	S3	S4	S5	S6	S7	S8	S9
S1 ($n = 27$)	0.000	0.051	0.029	0.056	0.040	0.015	0.075	0.022	0.045
S2 ($n = 26$)	***	0.000	0.074	0.038	0.014	0.037	0.022	0.061	0.042
S3 ($n = 22$)	***	***	0.000	0.092	0.067	0.028	0.120	0.027	0.079
S4 ($n = 20$)	***	***	***	0.000	0.013	0.046	0.029	0.065	0.019
S5 ($n = 27$)	***	**	***	*	0.000	0.028	0.014	0.054	0.015
S6 ($n = 33$)	***	***	***	***	***	0.000	0.049	0.018	0.033
S7 ($n = 5$)	***	NS	***	*	NS	**	0.000	0.079	0.020
S8 ($n = 24$)	***	***	***	***	***	***	***	0.000	0.038
S9 ($n = 16$)	***	***	***	**	*	***	NS	***	0.000

NS = non significant, * = $P < 0.05$, **$P < 0.01$, ***$P < 0.001$. n = sample size.

negative F_{IS} values (average $F_{IS} = -0.019$, NS) (Table 2). When all samples were considered, a significant departure from Hardy–Weinberg was observed with an average F_{IT} of 0.024 ($P < 0.01$). However, as the non significant F_{IS} values show, this appears to be mostly caused by differentiation between samples ($F_{ST} = 0.040$, $P < 0.001$) and is therefore probably the result of a Wahlund effect.

We found 95 significant LD values (at $\alpha = 0.05$) across all sampled regions (Appendix 1). S7 exhibited only two significant LD values, most probably because of its small sample size ($n = 5$). Most samples appear to have between six and 10 significant LD values, but S1 and S3 have 18 and 22 significant LD values, respectively. Despite this difference, we could not see any clear pattern across samples. For instance, no pair of loci was exhibiting a significant LD in more than three samples, the average being 1.07 population per locus pair with some variation across regions. For instance S1 and S3 only share three pairs of loci in LD. These results indicate that LD is most likely a result of the demographic history of the populations including events such as admixture or drift, rather than linkage.

Genetic differentiation between samples and between river sides

Overall, we found a limited but significant level of genetic differentiation among the samples (average $F_{ST} = 0.04$, $P < 0.001$). Pairwise F_{ST} values range between 0.01 and 0.12 and most are significant (Table 3, Fig. 2). As described in the Materials and methods section, it is possible to divide the samples into two sets, RS1 = (S1, S3, S6, S8) and RS2 = (S2, S4, S5, S7, S9), to test whether the Kinabatangan River represents a natural barrier to the movement of orang-utans. The distribution of pairwise F_{ST} values between samples belonging (i) to the same side of the river ($F_{ST(RS1)}$ and $F_{ST(RS2)}$) and (ii) to different sides ($F_{ST(RS1-2)}$) can be compared to the set of all pairwise F_{ST} values. Histograms of these values are represented in Fig. 2. The figure shows that, on average, F_{ST} values between samples from

the same side of the river (second and third panel from Fig. 2, average $F_{ST} = 0.023$ and 0.026, respectively) are lower than F_{ST} values between samples from different sides (lower panel of Fig. 2, average $F_{ST} = 0.058$). The permutation tests (all tests were significant at 0.1%, with the exception of the $F_{ST(RS1)}$ vs. $F_{ST(TOT)}$, which was significant at 5%) we performed allow us to demonstrate that (i) the three lower panels are not random sets of F_{ST} values, (ii) the F_{ST} values within each river side are significantly lower than the average F_{ST} across all samples, (iii) the F_{ST} values between river sides are significantly larger than the average F_{ST} across all samples, and (iv) the F_{ST} values within each river side are significantly lower than those observed between river sides.

Results of the Mantel tests performed with different measures of geographical distances (see Materials and methods) indicate that there is significant correlation between geographical and genetic distance when the river is considered to be a barrier ($r = 0.54, 0.72, 0.73, P < 0.01$, for the three distances used), but there is no correlation when the simple Euclidian geographical distance is used for all samples regardless of the river side ($r = -0.07$, NS). The correlation greatly increases between the case where we assume that orang-utans could cross at the level of S8 and S9 (where the river is still 200 m wide $r = 0.54$) and 150 km upstream (and the river starts to be reasonably narrow, $r = 0.72$). The correlation does not, however, increase with greater distance (i.e. when we assume that orang-utans could only cross the river at its source, $r = 0.73$).

Immigration rates

When samples were analysed using one river side vs. the other side, we found that the method of Wilson & Rannala (2003) produced highly consistent outputs across the runs, with clear indications that immigration rates are extremely low. For RS1, the posterior mean was 0.988, with the most probable value at 0.998 (Fig. 3). For RS2, the posterior mean was 0.971, with the most probable value at 0.982 (Fig. 3).
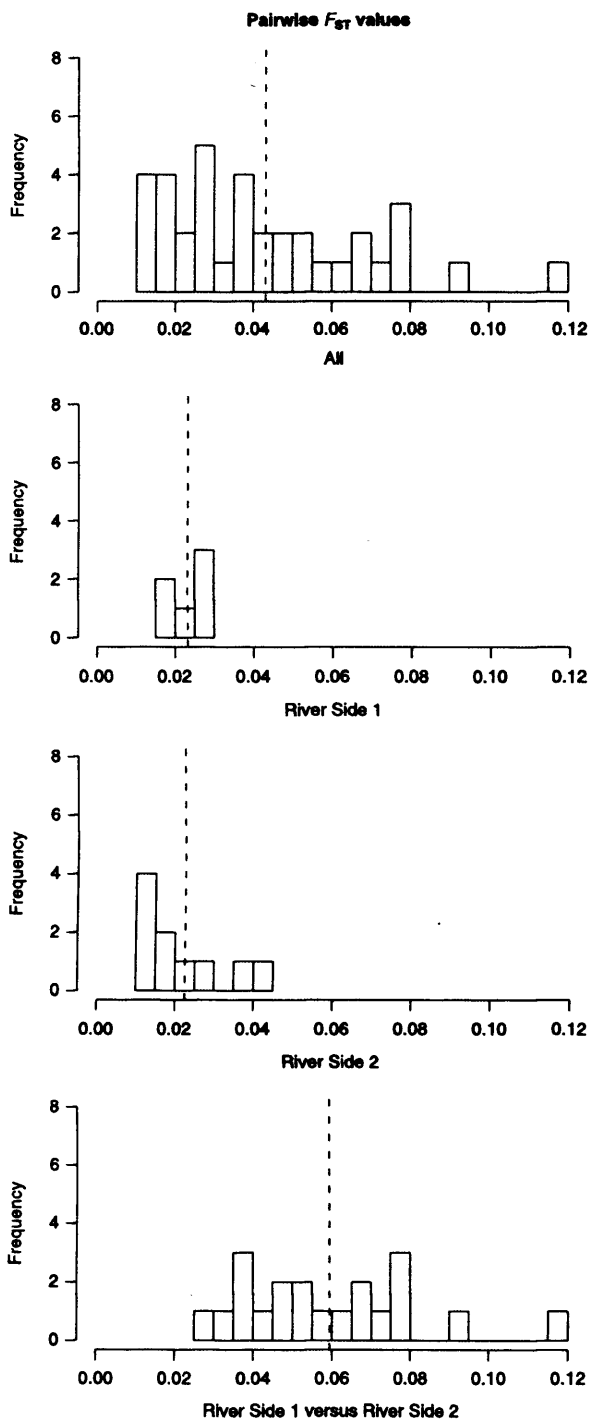
Pairwise F_{ST} values

Fig. 2 Pairwise F_{ST} values. The top panel represents pairwise F_{ST} values between all populations. The second and third panels represent pairwise F_{ST} values either between samples from RS1 or from RS2, respectively (see text). The last panel corresponds to pairwise F_{ST} between samples from opposite river sides.
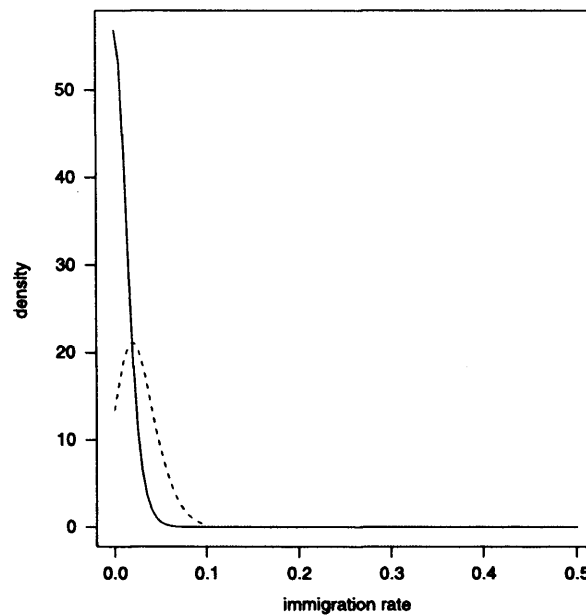

Fig. 3 Posterior distributions for the immigration rates for RS1 and RS2. The solid and dashed curves correspond to 1-m11 and 1-m22, and represent the immigration rate for RS1 and RS2, respectively.

Therefore, while it is impossible to completely reject the existence of movement across the river, there is very strong support for very low levels of migration despite the limited genetic differentiation. In fact, the method allows the determination for each individual (i.e. genotype) a posterior probability of being an immigrant. In such cases, it is also possible to estimate whether the multilocus genotype is an immigrant or the descendant of an immigrant in the preceding generation. These results show that for most individuals (156 out of 200), the probability of being a local rather than an immigrant is greater than 95%. For 41 of the remaining 44 individuals, this probability was not as high but remained larger than that of being an immigrant. For three individuals, all of which from RS2 (S4, S7 and S9), the probability of being an immigrant was larger than that of being local, even though the latter probability was still non-negligible (20, 40 and 44%, respectively). It is difficult to determine whether these individuals could indeed be immigrants or locals because of the limited F_{ST} values observed. In other words, most individuals are more likely to come from the river side they were sampled in, and for three individuals, the odds are that they could come from both sides.

When the method was applied using the nine samples (S1–S9) independently (i.e. allowing the estimation of immigration rates both within and between river sides), we found that the method produced inconsistent results across runs for some parameters, such as the F_i. In order to

determine whether these inconsistencies were a result of the lack of convergence of some parameter for some values of *deltap, deltam* or *deltaF* (i.e. inefficient proposal distributions), we modified the outputs (see Materials and methods). This allowed us to determine that migration rates were probably too high between samples from the same side to allow the method to work (this was confirmed by discussions with G Wilson). The method was therefore not used further at this scale. We note however, that the method is expected to become inefficient when immigration rates are greater than 66% (Wilson & Rannala 2003). Our results could thus indicate that individuals sampled in the different lots of each river side have relatively high probability of being immigrants from other neighbouring lots. This probability cannot be safely estimated because of the lack of convergence but is likely to be high.

The assignment analysis confirms these results but it indicates as well that for 11 individuals (five in RS1 and six in RS2), the most probable sample of origin is from the other river side. Applying the majority rule defined above, we find that among the five most probable samples, the river side of origin is more often represented in 168 individuals out 200. Applying the same rule to the six most likely sources increases this number to 181. In other words, there are between *c.* 20 and *c.* 35 individuals which are most probably assigned to the opposite river side.

Discussion

Genetic diversity, individual identification and population sizes

The results presented here show that the orang-utans sampled in the Lower Kinabatangan flood plain exhibit a high level of genetic variability despite the fragmentation of their environment. The diversity exhibited by the 14 human-derived microsatellite loci was high enough to permit an individual genetic identification of all 200 orang-utans typed in the study, which could prove particularly valuable for future studies of paternity assessment and relatedness.

Given the current census size estimates of approximately 1100 individuals (Ancrenaz *et al.* 2004), the genetic diversity observed in the Lower Kinabatangan orang-utans is surprisingly high, suggesting that orang-utans are not at mutation–drift equilibrium. This is supported by the relative lack of rare alleles, typically observed in populations that have been subject to a demographic bottleneck (e.g. Nei *et al.* 1975). Indeed, such populations are expected to first lose their rare alleles. As H_E is little affected by rare alleles (the square of their frequency is negligible), high H_E values can be observed long after the bottleneck has taken place (Nei *et al.* 1975; Chikhi & Bruford in press). These results are also in agreement with the very large numbers thought to have existed in the last centuries and millennia

across Borneo (Rijksen & Meijaard 1999; Delgado & van Schaik 2000), but do not allow us to determine whether the decrease in orang-utan numbers is recent or ancient.

One possibility is that the signal we detect corresponds to the slow decrease of orang-utans since the Pleistocene because of a combination of climate change and prehistoric hunting (Delgado & van Schaik 2000). It is possible, as a rough approximation, to estimate the long-term effective size, N_e, compatible with the observed level of genetic diversity as measured by H_E. Under the stepwise mutation model and with a mutation rate between 10^{-3} and 10^{-4}, we find that N_e would have to be between *c.* 1500 and 17 000 (Ohta & Kimura 1973). While such N_e estimates should not be taken at their face value, they are higher than the census size estimates and therefore confirm that the variability present in the Lower Kinabatangan is 'surprisingly high'. This result is further confirmed by the fact that in using human-derived microsatellites, there is likely to be an ascertainment bias towards underestimating genetic diversity (Ellegren *et al.* 1997), and hence N_e in orang-utans.

Another possibility is that the pattern of high diversity with few rare alleles and a small census size corresponds to much more recent changes, namely the anthropogenic destruction and fragmentation of the habitat that has taken place in the last century and in particular during the last decades. In the latter case, the high level of genetic diversity currently observed could be explained by three factors: (i) the presence of very large numbers throughout the Kinabatangan area over long periods of time, as noted above, (ii) the very recent habitat loss and degradation, which may have led to the concentration of the surviving individuals in the remaining forest patches along the river, and (iii) the long generation time and lifespan of the species which allowed populations to retain diversity for long periods after habitat loss. One consequence of this would be that high genetic diversity is transient and may only be present for a short time, as it would be 'concentrated' in adults which may soon be unable to reproduce.

A detailed exploration of past orang-utan demography has been carried out by Goossens *et al.* (submitted) and thus will not be developed here. We only note that when formal statistical tests are performed a strong and highly significant signal for a past demographic bottleneck is demonstrated confirming that the high level of genetic diversity observed in the Lower Kinabatangan is the remnant of an ancient large population.

Population structure, isolation by distance, and immigration rates

The analysis of the population structure showed a moderate (but significant) level of genetic differentiation between samples that are not geographically distant (average F_{ST} ~0.04). When only samples from the same side of the river

were analysed the average F_{ST} was significantly lower (F_{ST} ~0.025) than when samples from across the river were analysed (F_{ST}~0.06). These differences indicate that the Kinabatangan River represents a significant barrier to gene flow. The role of the Kinabatangan as a barrier is confirmed by our analysis of the correlation between pairwise F_{ST} values and the four geographical distances. When we assumed that orang-utans were able to cross the river in far upstream regions where the river's width becomes much smaller, we found that there was a significant correlation between geographical and genetic distance. This correlation disappeared when the river was artificially ignored. We also found that the correlation increased when we assumed that crossing the river was equivalent to travelling approximately 300 km. We thus recalculated the correlation between the two matrices by incrementing the distance from S8 and S9 to the crossing point by multiples of 10 km (corresponding to an increase of 20 km by going upstream for 10 km and back). We find that the correlation increases rapidly for the first 100 km (corresponding to an increase in distance of 200 km) from $r = 0.52$–0.69 but not after c. 150 km (i.e. 300 km in total).

Finally, the analysis of immigration rates allowed us to determine that rates of recent immigration were most probably close to zero across the river. We could not exclude the possibility that some individuals could have crossed the river and even found that assignment tests were sometimes favouring the opposite river side as the most probable area of origin. Practical knowledge of the sampled area suggests that it is extremely unlikely if not impossible for orang-utans to cross the Kinabatangan. The only bridge that could be used corresponds to the very frequented Sandakan-Lahad Datu road and is thus difficult to cross. Moreover, it would require the orang-utans to cross a village on one end. This suggests that the results obtained using the assignment method are either a result of the fact that intermediate genotypes can be 'generated' by both river sides (i.e. their likelihood is non-negligible using both riversides frequencies) or that they may come from other nonsampled regions. Another possibility is that they reflect the uncertainty resulting from the limited F_{ST} values between samples and river sides. In the immigration analysis this possibility is the most probable because of the decrease in average F_{ST} values obtained by pooling all samples from either river sides (the F_{ST} decreases from 0.058 to 0.036). Wilson & Rannala (2003) applied their method to two data sets exhibiting much higher differentiation levels. For example, in the wolf data used it appears that out of 36 pairwise F_{ST} values, only five were below 0.04, and 25 were larger than 0.058 (the average F_{ST} between the river sides) with values up to 0.188 (Carmichael et al. 2001). This explains the finer resolution obtained by Wilson & Rannala (2003).

These results are compatible with a model in which orang-utans move between neighbouring areas but do not cross the river, at least in the study area. In such a model, gene flow between the two river sides is maintained over generations through individuals crossing the river somewhere upstream. We cannot identify where orang-utans are most likely to cross the river, but in an isolation by distance model, the correlation between genetic and geographical distance increased when the crossing point was moved upstream until it reached values of 100–150 km. Put in a different way, the average F_{ST} observed between the river sides is equivalent to travelling approximately 200–300 km. Interestingly, these distances do correspond to regions where the river becomes narrower and crossing more plausible.

Previous studies on Borneo orang-utans have mostly used animals from rehabilitation centres. Warren et al. (2001) analysed mitochondrial DNA data from 41 individuals originating from six locations across Borneo including a sample from the Sepilok Orangutan sanctuary in Sabah. They found very large pairwise F_{ST} values between the samples (with two exceptions all values were larger than 0.48), and suggested that at least four biogeographical regions could be defined, namely (i) Southwest and Central Kalimantan, (ii) Northwest Kalimantan and Sarawak, (iii) Sabah, and (iv) East Kalimantan. Their study also suggested that the differentiation between these four regions could be very old (on the order of 860 000 years) and could therefore be a result of geographical barriers such as ancient river systems that separated populations during the colonization of the island from Sumatra. Our study confirms the potential role of rivers in isolating orang-utans at a much finer geographical scale. In an earlier study Warren et al. (2000), analysed orang-utans from East and West Kalimantan (the Indonesian part of Borneo) using five microsatellites with sample sizes between 10 and 43 individuals depending on the locus. They found that Nei's distance between East and West Kalimantan samples were small and had a large variance. They concluded that there was no significant differentiation at this scale. This result is at odds with both our results and those of Warren et al. (2001). One possible reason for the apparent discrepancy is that the conclusion of Warren et al. (2000) was based on the calculation of Nei's genetic distance and used small sample sizes. Comparison with our data is difficult, as they did not estimate F_{ST} values. For instance, their data indicated that diversity was higher within than between samples, which should not be interpreted as a lack of genetic differentiation and is indeed in agreement with our results (an F_{ST} of 0.02 indicates that 98% of the diversity is within sampled regions). We thus estimated F_{ST} values by using their Tables 1 and 2. The first provides allele frequencies and the second sample sizes for the different loci. Based on these tables we find that the single locus F_{ST}s are 0.061, 0.044, 0.004, 0.009 and −0.012 (i.e. 0.000). Thus, three loci essentially show no sign of

452 B. GOOSSENS *ET AL.*

genetic differentiation and two show values similar to those observed across the river or between the most differentiated samples from RS2 (Fig. 2). It is difficult to make strong conclusions from these calculations, and we can only note that more loci and more samples would be needed to have a better understanding of genetic differentiation at wider geographical scales.

Overall, our results show that significant genetic differentiation exists among orang-utan groups separated by less than 200 km. 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 road is a very frequented road and may provide a significant barrier to current and future gene flow.

Some consequences for the conservation of orang-utans

In the present study we have shown that LK orang-utans have maintained relatively high levels of genetic variability despite the increasing fragmentation of their habitat. While this may be seen as good news for the conservation of orang-utans, some caution should be taken. The maintenance or increase of current population sizes, including gene flow (through translocation for instance), are required to mitigate against significant loss of genetic diversity. Our results suggest that orang-utans move rather freely between lots from the same side of the river and that little, if any, movement seems to take place across the Kinabatangan River in the study area. Current orang-utan populations may continue to decrease in many of the forest lots investigated even if forest fragmentation stops. For example, in some lots the number of individuals estimated to survive is already low, as in lot 8 (corresponding to S7, see Fig. 1) where Ancrenaz *et al.* (2004) estimated the census size to be approximately 22. In such lots, genetic drift is going to reduce genetic diversity very quickly. We simulated genetic drift in this lot and found that two alleles will be lost every three generations for the next 10 generations at least. Given that these simulations optimistically assume that the census size is equal to the effective size, the situation is likely to be much worse. There is therefore an urgent need to maintain, and even increase, migration between lots. This could be done, for instance, by restoring forest corridors alongside the river banks and between lots. Translocation between lots from opposite sides of the river may be feasible because the differentiation is limited and a number of individuals were assigned to the opposite river side. However, we believe that such translocations should be avoided until other regions are sampled both upstream and away from the Kinabatangan River. Indeed, whereas we cannot exclude that nonsampled 'populations' could account for these individuals, we have good reasons to

think that orang-utans cannot cross the river in the area sampled. Of course, would population size keep decreasing, as could potentially happen in lot 8, translocation from any viable population would certainly be considered as a positive practical action. Whenever possible translocation between lots from the same side should be favoured. Moreover, wildlife surveys highlighted the importance of several areas in Sabah which need to be reconnected to each other (Ancrenaz *et al.* 2005). The time-frame for achieving corridor development may be hundreds of years, given the logistical challenges. Nonetheless, such systems are required if we are to conserve orang-utans and biodiversity in general in the long term.

Acknowledgements

References

Ancrenaz M, Gimenez O, Ambu L *et al.* (2005) Aerial surveys give new estimates for orang-utans in Sabah, Malaysia. *PLos Biology*, **3**, xx.

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**, 375–385.

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

Belkhir K, Borsa P, Chikhi L, Goudet J, Bonhomme F (1996/1997) GENETIX 3.07, Windows™ *Software for Population Genetics*. Laboratoire Génome et Populations, University of Montpellier II, Montpellier, France.

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology*, **13**, 551–561.

Brashares JS (2003) Ecological, behavioral, and life-history correlates of mammal extinctions in West Africa. *Conservation Biology*, 17, 733–743.

Cardillo M (2003) Biological determinants of extinction risk: why are smaller species less vulnerable? *Animal Conservation*, 6, 63–69.

Carmichael LE, Nagy JA, Larter NC, Strobeck C (2001) Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. *Molecular Ecology*, 10, 2787–2798.

Cegelski CC, Waits LP, Anderson NJ (2003) Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology*, 12, 2907–2918.

Chikhi L, Bruford MW (2005) Mammalian population genetics and genomics. In: *Mammalian Genomics* (eds Ruvinsky A, Marshall Graves J), pp. 539–584. CABI Publishers, UK.

Chikhi L, Bruford MW, Beaumont MA (2001) Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics*, 158, 1347–1362.

Coote T, Bruford MW (1996) Human microsatellites applicable for analysis of genetic variation in Apes and Old World monkeys. *Journal of Heredity*, 87, 406–410.

Cornuet J, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, 153, 1989–2000.

Davies KF, Margules CR, Lawrence KF (2000) Which traits of species predict population declines in experimental forest fragments? *Ecology*, 81, 1450–1461.

Delgado RA, van Schaik CP (2000) The behavioral ecology and conservation of the orangutan (*Pongo pygmaeus*): a tale of two islands. *Evolutionary Anthropology*, 9, 201–218.

Ellegren H, Moore S, Robinson N *et al.* (1997) Microsatellite evolution — a reciprocal study of repeat lengths at homologous loci in cattle and sheep. *Molecular Biology and Evolution*, 14, 854–860.

Galdikas GMF (1995) Social and reproductive behavior of wild adolescent female orangutans. In: *The Neglected Ape* (eds Nadler RD, Galdikas BMF, Sheeran LK, Rosen N), pp. 163–182. Plenum Press, New York, USA.

Gilks WR, Richarson S, Spiegelhalter DJ (1996) *Markov Chain Monte Carlo in Practice.* Chapman & Hall, New York.

Goossens B, Abdullah ZB, Sinyor JB, Ancrenaz M (2004) Which nests to choose: collecting shed hairs in wild orang-utans. *Folia Primatologica*, 75, 23–26.

Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau M, Bruford MW (submitted) Genetic signature of anthropogenic population collapse in orang-utans.

Goossens B, Chikhi L, Utami SS, de Ruiter J, Bruford MW (2000a) A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Conservation Genetics*, 1, 157–162.

Goossens B, Funk SM, Vidal C *et al.* (2002) Measuring genetic diversity in translocation progammes: principles and application to a chimpanzees release project. *Animal Conservation*, 5, 225–236.

Goossens B, Latour S, Vidal C *et al.* (2000b) Twenty new microsatellite loci for use with hair and faecal samples in the chimpanzee (*Pan troglodytes troglodytes*). *Folia Primatologica*, 71, 177–180.

Harcourt AH (2002) Empirical estimates of minimum viable population sizes for primates: tens to tens of thousands? *Animal Conservation*, 5, 237–244.

Keller LF, Jeffery KJ, Arcese P *et al.* (2001) Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proceedings of the Royal Society London*, 268, 1387–1394.

Kinnaird MF, Sanderson EW, O'Brien TG, Wibisono HT, Woolmer G (2003) Deforestation trends in a tropical landscape and implications for endangered large mammals. *Conservation Biology*, 17, 245–257.

Lackman-Ancrenaz I, Ancrenaz M, Saburi R (2001) The Kinabatangan Orang-utan Conservation Project (KOCP). In: *Proceedings of a Conference on the Apes: Challenges for the 21st Century* (ed. Chicago Zoological Society), pp. 262–265. Brookfield Zoo, Illinois, USA.

Laidlaw RK (2000) Effects of habitat disturbance and protected areas on mammals of Peninsular Malaysia. *Conservation Biology*, 14, 1639–1648.

Laurance WF (1991) Ecological correlates of extinction proneness in Australian tropical rainforest mammals. *Conservation Biology*, 5, 79–89.

Leendertz FH, Ellerbrok H, Boesch C *et al.* (2004) Anthrax kills wild chimpanzees in a tropical rainforest. *Nature*, 430, 451–452.

Leighton M, Seal US, Soemarna K *et al.* (1995) Orangutan life history and vortex analysis. In: *The Neglected Ape* (ed. Nadler RD, Galdikas BMF, Sheeran LK, Rosen N), pp. 97–108. Plenum Press, New York, USA.

Leroy EM, Rouquet P, Formenty P *et al.* (2004) Multiple ebola virus transmission events and rapid decline of Central African wildlife. *Science*, 303, 387–390.

Madsen T, Shine R, Olsson M, Wittzell H (1999) Conservation biology — Restoration of an inbred adder population. *Nature*, 402, 34–35.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27, 209–220.

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.

Miller CR, Waits LP (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 4334–4339.

Mitani JC (1989) Orangutan activity budgets: monthly variations and the effects of body size, parturition, and sociality. *International Journal of Primatology*, 18, 87–100.

Möller LM, Beheregaray LB (2004) Genetic evidence for sex-biased dispersal in resident bottlenose dolphins (*Tursiops aduncus*). *Molecular Ecology*, 13, 1607–1612.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583–590.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1–10.

O'Ryan C, Harley EH, Bruford MW *et al.* (1998) Microsatellite analysis of genetic diversity in fragmented South African buffalo populations. *Animal Conservation*, 1, 124–131.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetic Research*, 22, 201–204.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 4, 347–354.

Palsson S (2004) Isolation by distance, based on microsatellite data, tested with spatial autocorrelation (SPAIDA) and assignment test (SPASSIGN). *Molecular Ecology Notes*, 4, 143–145.

454 B. GOOSSENS *ET AL.*

Payne J (1987) Surveying orang-utan populations by counting nests from a helicopter; a pilot survey in Sabah. *Primate Conservation*, 8, 92–103.

Payne J (1988) *Orang-utan Conservation in Sabah*. WWF Malaysia, Kuala Lumpur.

Pryor KV, Young JE, Rumsey FJ *et al.* (2001) Diversity, genetic structure and evidence of outcrossing in British populations of the rock fern *Adiantum capillus-veneris* using microsatellites. *Molecular Ecology*, 10, 1881–1894.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 9197–9201.

Reed DH (2004) Extinction risk in fragmented habitats. *Animal Conservation*, 7, 181–191.

Rijksen HD, Meijaard E (1999) *Our Vanishing Relative: the Status of Wild Orang-utans at the Close of the 20th Century*. Kluwer Academic Publishers, Dordrecht.

Robertson JMY, van Schaik CP (2001) Causal factors underlying the dramatic decline of the Sumatran orang-utan. *Oryx*, 35, 26–38.

Rodman PS, Mitani JC (1987) Orangutans: sexual dimorphism in a solitary species. In: *Primate Societies* (eds Smuts B, Cheney DL, Seyfarth RM, Struhsaker T, Wrangham RW), pp. 146–154. University of Chicago Press, Chicago, USA.

Saccheri I, Kuussaari M, Kankare M *et al.* (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, 392, 491–494.

van Schaik CS, Monk KA, Robertson JMY (2001) Dramatic decline in orang-utan numbers in the Leaser Ecosystem, Northern Sumatra, *Oryx*, 35, 14–25.

Singleton I, van Schaik CP (2001) Orangutan home range size and its determinants in a Sumatran swamp forest. *International Journal of Primatology*, 22, 877–911.

Taberlet P, Griffin S, Goossens B *et al.* (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*, 14, 323–327.

Travis JMJ, Dytham C (1999) Habitat persistence, habitat availability and the evolution of dispersal. *Proceedings of the Royal Society London*, 266, 723–728.

Utami SS, Goossens B, Bruford MW, de Ruiter J, van Hooff JARAM (2002) Male bimaturism and reproductive success in Sumatran orang-utans. *Behavioral Ecology*, 13, 643–652.

Vigilant L (1999) An evaluation of techniques for the extraction and amplification of DNA from naturally shed hairs. *Biological Chemistry*, 380, 1329–1331.

Walsh PD, Abernethy KA, Bermejo M *et al.* (2003) Catastrophic ape decline in western equatorial Africa. *Nature*, 422, 611–614.

Warren KS, Nijman IJ, Lenstra JA *et al.* (2000) Microsatellite DNA variation in Bornean orang-utans (*Pongo pygmaeus*). *Journal of Medical Primatology*, 29, 57–62.

Warren KS, Verschoor EJ, Langenhuizen S *et al.* (2001) Speciation and intrasubspecific variation of Bornean orang-utans, *Pongo pygmaeus pygmaeus*. *Molecular Biology and Evolution*, 18, 472–480.

Webb JK, Brook BW, Shine R (2002) What makes a species vulnerable to extinction? Comparative life-history traits of two sympatric snakes. *Ecological Research*, 17, 59–67.

Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics*, 35, 235–254.

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

Wich SA, Singleton I, Utami-Atmoko SS *et al.* (2003) The status of the Sumatran orang-utan *Pongo abelii*: an update. *Oryx*, 37, 49–54.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163, 1177–1191.

Zhang YW, Morin PA, Ryder OA, Zhang YP (2001) A set of human tri- and tetra-nucleotide microsatellite loci useful for population analyses in gorillas (*Gorilla gorilla gorilla*) and orang-utans (*Pongo pygmaeus*). *Conservation Genetics*, 2, 391–395.

Zhi L, Karesh WB, Janczewski DN *et al.* (1996) Genomic differentiation among natural populations of orang-utan (*Pongo pygmaeus*). *Current Biology*, 6, 326–336.

B Goossens is a research associate at Cardiff University working on the population genetic structure, mating system and conservation genetics in a number of endangered mammal species. L Chikhi is a researcher at the 'Evolution et Diversité Biologique' CNRS Group in Toulouse, interested in using molecular markers to infer ancient population demography, with a particular emphasis on humans and primates. MF Jalil is a Malaysian PhD student at Cardiff University working on a comparative phylogeography of three primate species in Sabah. M Ancrenaz and I Lackman-Ancrenaz are the Directors of the KOCP based in the Kinabatangan flood plain. M Mohamed is the Director of the Institute for Tropical Biology and Conservation at Universiti Malaysia Sabah and is interested in entomological biodiversity and ecology. P Andau is the Director of the Sabah Wildlife Department. MW Bruford is interested in the evolutionary ecology and population genetics of threatened species.

Appendix 1

Tests of linkage disequilibrium (LD). LD was measured using the correlation coefficient (Weir 1979). We represent the proportion of randomised values greater or equal to the observed correlation coefficient

Locus 1	Locus 2	S1	S2	S3	S4	S5	S6	S7	S8	S9
D5S1457	D5S1470	0.9	76.6	92.9	39.5	98.1	19.4	3.1	24.2	27.1
D5S1457	D1S550	43.5	58.4	7.5	73.7	38.7	29.1	14.1	87.1	65.4
D5S1457	D2S1326	47	4.6	35.4	88	85.3	78.6	100	39.5	5
D5S1457	D3S2459	49.4	16.9	0	53.8	54.6	66.9	61.3	72.6	0
D5S1457	D4S1627	35.9	31	22.5	40.1	72.6	53.8	71	32.3	84.8
D5S1457	D4S2408	6	60.1	4.4	51.2	18.3	22.8	100	0.3	16.2
D5S1457	D5S1505	84.6	89.9	99.8	4.9	81.8	63.6	20.9	2.6	45.5
D5S1457	D6S501	4.8	58.4	9.6	49.8	4.4	13.5	10.2	88.8	21.8
D5S1457	D13S321	11.5	1.8	5.6	75.1	54.9	60.5	22.1	65	30.7
D5S1457	D13S765	77.1	31.2	16.8	78.9	9.4	32.9	37.9	19.7	11.2
D5S1457	D12S375	3	51.5	44.7	0.5	59.5	71.8	63.9	64	73.6
D5S1457	D2S141	4.6	76.4	33.1	20.7	51.1	28.5	62.5	39.2	92.3
D5S1457	D16S420	29.4	22.8	11.7	73.4	78.5	29.4	12.6	45.4	35.8
D5S1470	D1S550	37.9	73	30.1	1.7	32.6	83	6.7	42.5	74
D5S1470	D2S1326	28	36.1	4	28	49.6	90.5	41.2	98.5	52.3
D5S1470	D3S2459	82.3	0.6	51.9	12.2	52.5	85.4	57.7	58.7	22
D5S1470	D4S1627	0.1	63.6	18	76	32.9	86.2	20	93	87.4
D5S1470	D4S2408	0.2	42.6	79.2	2.7	96.6	8.2	100	4.8	28.4
D5S1470	D5S1505	8.2	96.7	17.3	88.2	3.8	68.1	19.1	1.5	10
D5S1470	D6S501	16.9	3.9	22.4	74.9	17.1	12.6	52.9	98.2	3
D5S1470	D13S321	62.7	30.5	24.1	68.3	51.6	16.3	20	72.1	87.9
D5S1470	D13S765	57	20.6	76.6	64.2	40.2	55.2	79.5	88.8	10.5
D5S1470	D12S375	17.8	89	17.1	56.6	71.3	99.1	6.9	11.6	64.9
D5S1470	D2S141	3.3	3.4	50.2	46.5	13.2	95.2	22.5	33.9	90.9
D5S1470	D16S420	30.7	80.9	63.3	0.2	76.5	52.4	2.7	85.7	33.9
D1S550	D2S1326	85.7	91.2	17.4	57.9	36.3	56.3	89.8	40.8	63.8
D1S550	D3S2459	81.6	6.2	35.9	87.1	71.5	42.1	49.4	42.1	55.4
D1S550	D4S1627	18.2	54.7	17.1	91.8	74.2	80.7	38.4	74.2	75.9
D1S550	D4S2408	46.3	49.1	32.3	0.6	43	5.1	79.6	52.2	60
D1S550	D5S1505	38.3	7	1.6	88.6	30.8	25.4	32	4.8	50.5
D1S550	D6S501	92.5	34.1	16.6	8	2.4	1.8	43.9	46.9	48.6
D1S550	D13S321	53.4	14.6	1.5	12.4	55.9	7.9	61.2	70.7	62.7
D1S550	D13S765	3.4	86.6	2.3	67.7	95.8	77.7	60.3	79.4	36.9
D1S550	D12S375	39.2	52.8	86.9	22.6	15.7	41.3	12.3	71.2	46.1
D1S550	D2S141	42.3	8.9	58	37.3	7.7	6.7	28.8	82.8	22.4
D1S550	D16S420	0.2	73.1	68.5	54.6	28.6	0	8.9	94.9	84.6
D2S1326	D3S2459	21.6	21	0	36.6	15.3	36.6	10.2	64.1	14.5
D2S1326	D4S1627	41.5	51.5	2.8	35.2	1.3	61	19	26.6	83.9
D2S1326	D4S2408	3.7	80.5	28.7	51.9	2.4	65.7	40	59.3	91.9
D2S1326	D5S1505	18.3	50.9	2.8	9.2	71.6	2.4	37.5	47.7	57.5
D2S1326	D6S501	50.4	83.9	1.9	20.3	12.3	5.5	9.4	1	3.2
D2S1326	D13S321	84.2	23.5	4.8	15.1	86.1	72.7	25.1	39.6	9.7
D2S1326	D13S765	13.5	9.4	8.2	75.5	43.9	40.5	10.3	51.1	20.1
D2S1326	D12S375	10.9	72	0.3	28.9	99.3	75.6	60.6	18.2	9.9
D2S1326	D2S141	3.4	56	42.1	73.7	40.9	67.1	27.5	36.5	96.2
D2S1326	D16S420	59.4	0.3	0.9	5.6	65.7	7.6	9.8	64.5	16.2
D3S2459	D4S1627	51.2	72.7	4.9	13.6	0.4	45.1	90.4	2.5	9.2
D3S2459	D4S2408	93	47.9	0.4	71.2	0.1	61.3	100	54.9	0.2
D3S2459	D5S1505	69	71.6	23.2	18.3	27.4	65.9	50.1	56.8	19.5
D3S2459	D6S501	0.9	54.5	6.6	3	8	81.7	39.5	12.9	22.3
D3S2459	D13S321	21.9	69.3	1.9	55.5	80.4	50.5	42.1	57.7	39.9
D3S2459	D13S765	77.2	33.3	5.9	58.9	18.8	30.6	27.8	75.3	37.4
D3S2459	D12S375	55.8	69.6	3.8	27.4	14	83.5	27.3	76.3	34.7
D3S2459	D2S141	35.8	17.2	27.9	32	58.4	0.8	71.1	88.1	78.1
D3S2459	D16S420	61.3	78.1	3.7	13.2	74.7	52.6	5.7	87.9	36.2

Appendix 1 *Continued*

Locus 1	Locus 2	S1	S2	S3	S4	S5	S6	S7	S8	S9
D4S1627	D4S2408	0	19.3	34.8	26.7	25.3	13.7	39.2	29.5	54.3
D4S1627	D5S1505	87.2	2.1	44.8	78.5	16.2	67.6	14.6	10.4	48.8
D4S1627	D6S501	7.4	36.8	52	30.5	31.4	80.7	46.5	3.5	20.2
D4S1627	D13S321	29.6	12.4	8	36.4	24	39.5	5.4	4.5	92
D4S1627	D13S765	0.2	14.3	64.4	45.9	75.1	49.2	100	73.2	88.9
D4S1627	D12S375	21.4	76	0.7	7.8	56.5	9.3	49.9	97.6	21
D4S1627	D2S141	35.5	42.5	11.1	0.8	92	18.9	7.6	5.2	12.5
D4S1627	D16S420	30.7	94.8	35.6	95.7	56.4	51	11.3	15.8	6
D4S2408	D5S1505	33.9	27.9	80.8	24	91.2	71	100	1.9	19.6
D4S2408	D6S501	16.6	91.6	16.3	7.4	6.8	40.5	100	51.9	58.7
D4S2408	D13S321	3.8	21.6	2.8	89.4	68.7	13.8	100	21.7	47.3
D4S2408	D13S765	11.9	42.8	12.3	80.6	12.9	14.7	39.9	18.6	36.8
D4S2408	D12S375	65.1	64.2	5.7	1.7	52.1	23.1	42.3	76.6	54.7
D4S2408	D2S141	78.3	19	21.1	38.5	39.2	37.9	79.1	6.1	69.6
D4S2408	D16S420	7.4	6	41.8	55.8	85.5	10	60.4	23.8	42.9
D5S1505	D6S501	1.7	81.8	12.2	23.1	14.5	93.2	21.4	78.4	74.3
D5S1505	D13S321	22.4	45.2	2.5	83.4	85.5	82.2	9.4	10.7	52.6
D5S1505	D13S765	84.6	29.4	6.2	70	58.7	25.4	72	65.2	91.5
D5S1505	D12S375	34.8	23.4	86.6	14.2	95.3	64.2	100	21.8	45.4
D5S1505	D2S141	5	4.9	56.6	80.9	57.9	99.5	15	18.8	46.9
D5S1505	D16S420	54.1	16	9.7	3.5	30.5	72.1	30.5	7.5	41.5
D6S501	D13S321	7.8	47	6.2	33.8	78.6	59.7	20.6	20.7	75.5
D6S501	D13S765	14.9	39.1	1.3	79.8	0.1	4	31.4	61.4	15.8
D6S501	D12S375	47.1	72.3	8.4	75.6	57.7	36.3	63.3	87.5	51.1
D6S501	D2S141	27.1	81.6	6.4	90.4	36.9	50.2	20.4	82	56.4
D6S501	D16S420	90.8	60.7	0.4	33.7	68.6	9.4	31.7	60.7	77.3
D13S321	D13S765	66.2	78.4	14	18.2	13.5	61.2	100	20.5	58.7
D13S321	D12S375	26.3	78.8	18.7	35	28.1	26.4	100	22.3	3.8
D13S321	D2S141	14.6	75.8	6.5	36	49.2	47.2	18.4	68.2	0.7
D13S321	D16S420	23.3	18.6	39.1	32.1	33.7	57.5	28.9	6.5	41.6
D13S765	D12S375	76	55.4	65.8	67.8	51.8	19.1	18.7	26.4	34
D13S765	D2S141	50	4.6	98.4	62.7	68.2	50.8	78.5	35.5	95.8
D13S765	D16S420	29.7	5.7	31.2	85.1	84.6	69.9	19.9	56.9	44.1
D12S375	D2S141	1.7	32.4	4.6	11.3	12.1	44.5	79.6	24.6	47.4
D12S375	D16S420	35	12	70.6	1.6	1.9	88.1	16	75	21.1
D2S141	D16S420	6.5	0.8	22.7	41.6	0.4	4.8	30.7	10.3	55

Appendix III

Long-tailed macaque and proboscis monkey samples (arranged by extract number followed by GPS code, species name and box number).

I. Long-tailed macaque.

Extract	GPS Code	Species	BOX	Remark	
1	192	Macaques	1	North	
2	192	Macaques	1	North	
3	192	Macaques	1	North	
4	192	Macaques	1	North	
5	192	Macaques	1	North	
6	192	Macaques	1	North	
7	192	Macaques	1	North	
8	192	Macaques	1	North	
9	167	Macaques	1	North	
10	167	Macaques	1	North	
11	167	Macaques	1	North	
12	167	Macaques	1	North	
13	92	Macaques	1	North	
14	92	Macaques	1	North	
15	92	Macaques	1	North	
16	92	Macaques	1	North	
17	92	Macaques	1	North	
18	92	Macaques	1	North	
19	92	Macaques	1	North	
20	92	Macaques	1	North	
21	375	Macaques	1	South: Outside of Lot 9	
22	375	Macaques	1	South: Outside of Lot 9	
23	375	Macaques	1	South: Outside of Lot 9	
24	375	Macaques	1	South: Outside of Lot 9	
25	94	Macaques	1	North: Outside of Lot 7	
26	94	Macaques	1	North: Outside of Lot 7	
27	94	Macaques	1	North: Outside of Lot 7	
28	94	Macaques	1	North: Outside of Lot 7	
29	94	Macaques	1	North: Outside of Lot 7	
30	94	Macaques	1	North: Outside of Lot 7	
31	94	Macaques	1	North: Outside of Lot 7	
32	94	Macaques	1	North: Outside of Lot 7	
33	57	Macaques	1	South	
34	57	Macaques	1	South	
35	57	Macaques	1	South	
36	57	Macaques	1	South	
37	270	Macaques	1	North: Outside of Lot 7	
38	270	Macaques	1	North: Outside of Lot 7	
39	270	Macaques	1	North: Outside of Lot 7	
40	270	Macaques	1	North: Outside of Lot 7	
41	270	Macaques	1	North: Outside of Lot 7	
42	270	Macaques	1	North: Outside of Lot 7	
43	221	Macaques	1	North	
44	221	Macaques	1	North	
45	221	Macaques	1	North	
46	221	Macaques	1	North	
47	397	Macaques	1	South: Between Lot 7 - 8	
48	397	Macaques	1	South: Between Lot 7 - 8	
49	259	Macaques	1	North	
50	259	Macaques	1	North	
51	259	Macaques	1	North	
52	259	Macaques	1	North	

53	259	Macaques	1	North	
54	259	Macaques	1	North	
55	259	Macaques	1	North	
56	259	Macaques	1	North	
57	96	Macaques	1	North	
58	96	Macaques	1	North	
59	96	Macaques	1	North	
60	96	Macaques	1	North	
61	99	Macaques	1	South	
62	99	Macaques	1	South	
63	99	Macaques	1	South	
64	99	Macaques	1	South	
65	99	Macaques	1	South	
66	99	Macaques	1	South	
67	156	Macaques	1	South	
68	156	Macaques	1	South	
69	156	Macaques	1	South	
70	156	Macaques	1	South	
71	156	Macaques	1	South	
72	156	Macaques	1	South	
73	156	Macaques	1	South	
74	156	Macaques	1	South	
75	156	Macaques	1	South	
76	156	Macaques	1	South	
77	156	Macaques	1	South	
78	156	Macaques	1	South	
79	156	Macaques	1	South	
80	156	Macaques	1	South	
81	156	Macaques	2	South	
82	156	Macaques	2	South	
83	156	Macaques	2	South	
84	156	Macaques	2	South	
85	LM04	Macaques	2	Pangi Forest Reserve	
86	LM04	Macaques	2	Pangi Forest Reserve	
87	LM04	Macaques	2	Pangi Forest Reserve	
88	LM04	Macaques	2	Pangi Forest Reserve	
89	LM04	Macaques	2	Pangi Forest Reserve	
90	LM04	Macaques	2	Pangi Forest Reserve	
91	LM04	Macaques	2	Pangi Forest Reserve	
92	LM04	Macaques	2	Pangi Forest Reserve	
93	LM04	Macaques	2	Pangi Forest Reserve	
94	LM04	Macaques	2	Pangi Forest Reserve	
95	LM04	Macaques	2	Pangi Forest Reserve	
96	LM04	Macaques	2	Pangi Forest Reserve	
97	LM04	Macaques	2	Pangi Forest Reserve	
98	LM04	Macaques	2	Pangi Forest Reserve	
99	226	Macaques	2	North	
100	226	Macaques	2	North	
101	226	Macaques	2	North	
102	226	Macaques	2	North	
103	225	Macaques	2	North	
104	225	Macaques	2	North	
105	225	Macaques	2	North	
106	225	Macaques	2	North	
107	170	Macaques	2	North: Between Lot 4 - 5	
108	170	Macaques	2	North: Between Lot 4 - 5	
109	LM06	Macaques	2	Pangi Forest Reserve	
110	LM06	Macaques	2	Pangi Forest Reserve	
111	LM06	Macaques	2	Pangi Forest Reserve	
112	LM06	Macaques	2	Pangi Forest Reserve	
113	LM06	Macaques	2	Pangi Forest Reserve	

114	LM06	Macaques	2	Pangi Forest Reserve	
115	LM06	Macaques	2	Pangi Forest Reserve	
116	LM06	Macaques	2	Pangi Forest Reserve	
117	LM06	Macaques	2	Pangi Forest Reserve	
118	LM06	Macaques	2	Pangi Forest Reserve	
119	LM06	Macaques	2	Pangi Forest Reserve	
120	LM06	Macaques	2	Pangi Forest Reserve	
121	LM06	Macaques	2	Pangi Forest Reserve	
122	LM06	Macaques	2	Pangi Forest Reserve	
123	LM06	Macaques	2	Pangi Forest Reserve	
124	LM06	Macaques	2	Pangi Forest Reserve	
125	LM06	Macaques	2	Pangi Forest Reserve	
126	LM06	Macaques	2	Pangi Forest Reserve	
127	LM06	Macaques	2	Pangi Forest Reserve	
128	LM06	Macaques	2	Pangi Forest Reserve	
129	LM06	Macaques	2	Pangi Forest Reserve	
130	LM06	Macaques	2	Pangi Forest Reserve	
131	LM06	Macaques	2	Pangi Forest Reserve	
132	LM06	Macaques	2	Pangi Forest Reserve	
133	LM06	Macaques	2	Pangi Forest Reserve	
134	LM06	Macaques	2	Pangi Forest Reserve	
135	LM06	Macaques	2	Pangi Forest Reserve	
136	LM06	Macaques	2	Pangi Forest Reserve	
137	LM06	Macaques	2	Pangi Forest Reserve	
138	LM06	Macaques	2	Pangi Forest Reserve	
139	LM06	Macaques	2	Pangi Forest Reserve	
140	LM06	Macaques	2	Pangi Forest Reserve	
141	LM06	Macaques	2	Pangi Forest Reserve	
142	LM06	Macaques	2	Pangi Forest Reserve	
143	LM06	Macaques	2	Pangi Forest Reserve	
144	LM06	Macaques	2	Pangi Forest Reserve	
145	216	Macaques	2	South	
146	216	Macaques	2	South	
147	LM05	Macaques	2	South: Outside Lot 3	
148	LM05	Macaques	2	South: Outside Lot 3	
149	LM05	Macaques	2	South: Outside Lot 3	
150	LM05	Macaques	2	South: Outside Lot 3	
151	LM05	Macaques	2	South: Outside Lot 3	
152	LM05	Macaques	2	South: Outside Lot 3	
153	LM05	Macaques	2	South: Outside Lot 3	
154	LM05	Macaques	2	South: Outside Lot 3	
155	LM05	Macaques	2	South: Outside Lot 3	
156	LM05	Macaques	2	South: Outside Lot 3	
157	LM05	Macaques	2	South: Outside Lot 3	
158	LM05	Macaques	2	South: Outside Lot 3	
159	LM05	Macaques	2	South: Outside Lot 3	
160	LM05	Macaques	2	South: Outside Lot 3	
161	LM05	Macaques	3	South: Outside Lot 3	
162	LM05	Macaques	3	South: Outside Lot 3	
163	LM05	Macaques	3	South: Outside Lot 3	
164	LM05	Macaques	3	South: Outside Lot 3	
165	LM05	Macaques	3	South: Outside Lot 3	
166	LM05	Macaques	3	South: Outside Lot 3	
167	LM05	Macaques	3	South: Outside Lot 3	
168	LM05	Macaques	3	South: Outside Lot 3	
169	LM05	Macaques	3	South: Outside Lot 3	
170	LM05	Macaques	3	South: Outside Lot 3	
171	LM05	Macaques	3	South: Outside Lot 3	
172	LM05	Macaques	3	South: Outside Lot 3	
173	LM05	Macaques	3	South: Outside Lot 3	
174	LM05	Macaques	3	South: Outside Lot 3	

175	LM05	Macaques	3	South: Outside Lot 3	
176	LM05	Macaques	3	South: Outside Lot 3	
177	LM05	Macaques	3	South: Outside Lot 3	
178	LM05	Macaques	3	South: Outside Lot 3	
179	LM05	Macaques	3	South: Outside Lot 3	
180	LM05	Macaques	3	South: Outside Lot 3	
181	LM05	Macaques	3	South: Outside Lot 3	
182	LM05	Macaques	3	South: Outside Lot 3	
183	LM05	Macaques	3	South: Outside Lot 3	
184	LM05	Macaques	3	South: Outside Lot 3	
185	269	Macaques	3	North: Outside of Lot 7	
186	269	Macaques	3	North: Outside of Lot 7	
187	218	Macaques	3	South	
188	218	Macaques	3	South	
Extract	**KOD GPS**	**Species**	**Box**	**Remark**	**Xtract**
1	269	Macaques		North: Outside of Lot 7	New Extract 185-186
2	269	Macaques		North: Outside of Lot 7	New Extract 185-186
3	226	Macaques		North	New Extract 99-102
4	226	Macaques		North	New Extract 99-102
5	**281**	**Macaques**		**South: Between Lot 4 - 5**	**Unique Extract**
6	**281**	**Macaques**		**South: Between Lot 4 - 5**	**Unique Extract**
7	96	Macaques		North	New Extract 57 - 60
8	96	Macaques		North	New Extract 57 - 60
9	**216**	**Macaques**		**South**	**Unique Extract**
10	**216**	**Macaques**		**South**	**Unique Extract**
11	225	Macaques		North	New Extract 103-106
12	225	Macaques		North	New Extract 103-106
13	**LM03**	**Macaques**		**North**	**Unique Extract**
14	**LM03**	**Macaques**		**North**	**Unique Extract**
15	**LM01**	**Macaques**		**South: Outside of Lot 1**	**Unique Extract**
16	**LM01**	**Macaques**		**South: Outside of Lot 1**	**Unique Extract**
17	221	Macaques		North	New Extract 43-46
18	221	Macaques		North	New Extract 43-46
19	375	Macaques		South: Outside of Lot 9	New Extract 21-24
20	375	Macaques		South: Outside of Lot 9	New Extract 21-24
21	LM04	Macaques		Pangi Forest Reserve	New Extract 85-98
22	LM04	Macaques		Pangi Forest Reserve	New Extract 85-98
23	**381**	**Macaques**		**South: Outside of Lot 9**	**Unique Extract**
24	**381**	**Macaques**		**South: Outside of Lot 9**	**Unique Extract**
25	**285**	**Macaques**		**North**	**Unique Extract**
26	**285**	**Macaques**		**North**	**Unique Extract**
27	**61**	**Macaques**		**South**	**Unique Extract**
28	**61**	**Macaques**		**South**	**Unique Extract**
29	192	Macaques		North	New Extract 1-8
30	192	Macaques		North	New Extract 1-8
31	167	Macaques		North	New Extract 9-12
32	167	Macaques		North	New Extract 9-12
33	270	Macaques		North: Outside of Lot 7	New Extract 37-42
34	270	Macaques		North: Outside of Lot 7	New Extract 37-42
35	156	Macaques		South	New Extract 67-84
36	156	Macaques		South	New Extract 67-84
37	**279**	**Macaques**		**South: Outside of Lot 3**	**Unique Extract**
38	**279**	**Macaques**		**South: Outside of Lot 3**	**Unique Extract**
39	**93**	**Macaques**		**North**	**Unique Extract**
40	**93**	**Macaques**		**North**	**Unique Extract**
41	259	Macaques		North	New Extract 49-56
42	259	Macaques		North	New Extract 49-56
43	92	Macaques		North	Nex Extract 13-20
44	92	Macaques		North	Nex Extract 13-20
45	**217**	**Macaques**		**South**	**Unique Extract**

46	217	Macaques		South	Unique Extract
47	223	Macaques		South	Unique Extract
48	223	Macaques		South	Unique Extract
49	170	Macaques		North: Between Lot 4 - 5	New Extract 107-108
50	170	Macaques		North: Between Lot 4 - 5	New Extract 107-108
51	94	Macaques		North: Outside of Lot 7	New Extract 25-32
52	94	Macaques		North: Outside of Lot 7	New Extract 25-32
53	LM02 (57)	Macaques		South	New Extract 33-36
54	LM02 (57)	Macaques		South	New Extract 33-36
55	LM06	Macaques		Pangi Forest Reserve	New Extract 109-144
56	LM06	Macaques		Pangi Forest Reserve	New Extract 109-144
57	397	Macaques		South: Between Lot 7 - 8	New Extract 47 - 48
58	397	Macaques		South: Between Lot 7 - 8	New Extract 47 - 48
59	218	Macaques		South	New Extract 187-188
60	218	Macaques		South	New Extract 187-188
Extract	Kod GPS	Species		Remark	
Box 7: 1	Labuk Bay	Macaques			
Box 7: 2	Labuk Bay	Macaques			
Box 7:9	Tawau Hill	Macaques			
Box 7:10	Tawau Hill	Macaques			
Box 7:11	Tawau Hill	Macaques			
Box 7:12	Tawau Hill	Macaques			

II. Proboscis monkey

Extract	Kod GPS	Species	BOX	Remark	
189	PM03	Proboscis	3	North: Outside of Lot 2	
190	PM03	Proboscis	3	North: Outside of Lot 2	
191	133	Proboscis	3	North	
192	133	Proboscis	3	North	
193	PM03	Proboscis	3	North: Outside of Lot 2	
194	PM03	Proboscis	3	North: Outside of Lot 2	
195	PM03	Proboscis	3	North: Outside of Lot 2	
196	PM03	Proboscis	3	North: Outside of Lot 2	
197	PM03	Proboscis	3	North: Outside of Lot 2	
198	PM03	Proboscis	3	North: Outside of Lot 2	
199	PM03	Proboscis	3	North: Outside of Lot 2	
200	PM03	Proboscis	3	North: Outside of Lot 2	
201	PM03	Proboscis	3	North: Outside of Lot 2	
202	PM03	Proboscis	3	North: Outside of Lot 2	
203	227	Proboscis	3	North	
204	227	Proboscis	3	North	
205	47	Proboscis	3	South: Margin Lot 1	
206	47	Proboscis	3	South: Margin Lot 1	
207	47	Proboscis	3	South: Margin Lot 1	
208	47	Proboscis	3	South: Margin Lot 1	
209	47	Proboscis	3	South: Margin Lot 1	
210	47	Proboscis	3	South: Margin Lot 1	
211	47	Proboscis	3	South: Margin Lot 1	
212	47	Proboscis	3	South: Margin Lot 1	
213	47	Proboscis	3	South: Margin Lot 1	
214	47	Proboscis	3	South: Margin Lot 1	
215	47	Proboscis	3	South: Margin Lot 1	
216	47	Proboscis	3	South: Margin Lot 1	
217	323	Proboscis	3	North	
218	323	Proboscis	3	North	
219	323	Proboscis	3	North	
220	323	Proboscis	3	North	

221	320	Proboscis	3	North: Outside of Lot 4	
222	320	Proboscis	3	North: Outside of Lot 4	
223	320	Proboscis	3	North: Outside of Lot 4	
224	320	Proboscis	3	North: Outside of Lot 4	
225	320	Proboscis	3	North: Outside of Lot 4	
226	320	Proboscis	3	North: Outside of Lot 4	
227	320	Proboscis	3	North: Outside of Lot 4	
228	320	Proboscis	3	North: Outside of Lot 4	
229	313	Proboscis	3	South: Outside of Lot 3	
230	313	Proboscis	3	South: Outside of Lot 3	
231	313	Proboscis	3	South: Outside of Lot 3	
232	313	Proboscis	3	South: Outside of Lot 3	
233	313	Proboscis	3	South: Outside of Lot 3	
234	313	Proboscis	3	South: Outside of Lot 3	
235	292	Proboscis	3	North	
236	292	Proboscis	3	North	
237	292	Proboscis	3	North	
238	292	Proboscis	3	North	
239	292	Proboscis	3	North	
240	292	Proboscis	3	North	
241	321	Proboscis	4	North: Outside of Lot 4	
242	321	Proboscis	4	North: Outside of Lot 4	
243	321	Proboscis	4	North: Outside of Lot 4	
244	321	Proboscis	4	North: Outside of Lot 4	
245	321	Proboscis	4	North: Outside of Lot 4	
246	321	Proboscis	4	North: Outside of Lot 4	
247	321	Proboscis	4	North: Outside of Lot 4	
248	321	Proboscis	4	North: Outside of Lot 4	
249	321	Proboscis	4	North: Outside of Lot 4	
250	321	Proboscis	4	North: Outside of Lot 4	
251	321	Proboscis	4	North: Outside of Lot 4	
252	321	Proboscis	4	North: Outside of Lot 4	
253	325	Proboscis	4	South: Outside of Lot 6	
254	325	Proboscis	4	South: Outside of Lot 6	
255	325	Proboscis	4	South: Outside of Lot 6	
256	325	Proboscis	4	South: Outside of Lot 6	
257	325	Proboscis	4	South: Outside of Lot 6	
258	325	Proboscis	4	South: Outside of Lot 6	
259	325	Proboscis	4	South: Outside of Lot 6	
260	325	Proboscis	4	South: Outside of Lot 6	
261	325	Proboscis	4	South: Outside of Lot 6	
262	325	Proboscis	4	South: Outside of Lot 6	
263	325	Proboscis	4	South: Outside of Lot 6	
264	325	Proboscis	4	South: Outside of Lot 6	
265	15	Proboscis	4	North: Balad Dami	
266	15	Proboscis	4	North: Balad Dami	
267	15	Proboscis	4	North: Balad Dami	
268	15	Proboscis	4	North: Balad Dami	
269	357	Proboscis	4	South: Outside Lot 10	
270	357	Proboscis	4	South: Outside Lot 10	
271	357	Proboscis	4	South: Outside Lot 10	
272	357	Proboscis	4	South: Outside Lot 10	
273	357	Proboscis	4	South: Outside Lot 10	
274	357	Proboscis	4	South: Outside Lot 10	
275	357	Proboscis	4	South: Outside Lot 10	
276	357	Proboscis	4	South: Outside Lot 10	
277	311	Proboscis	4	South: Outside Lot 3	
278	311	Proboscis	4	South: Outside Lot 3	
279	311	Proboscis	4	South: Outside Lot 3	
280	311	Proboscis	4	South: Outside Lot 3	
281	311	Proboscis	4	South: Outside Lot 3	

282	311	Proboscis	4	South: Outside Lot 3	
283	293	Proboscis	4	South: Outside Lot 3	
284	293	Proboscis	4	South: Outside Lot 3	
285	293	Proboscis	4	South: Outside Lot 3	
286	293	Proboscis	4	South: Outside Lot 3	
287	293	Proboscis	4	South: Outside Lot 3	
288	293	Proboscis	4	South: Outside Lot 3	
289	55	Proboscis	4	South: Margin Lot 1	
290	55	Proboscis	4	South: Margin Lot 1	
291	55	Proboscis	4	South: Margin Lot 1	
292	55	Proboscis	4	South: Margin Lot 1	
293	55	Proboscis	4	South: Margin Lot 1	
294	55	Proboscis	4	South: Margin Lot 1	
295	296	Proboscis	4	North	
296	296	Proboscis	4	North	
297	296	Proboscis	4	North	
298	296	Proboscis	4	North	
299	296	Proboscis	4	North	
300	296	Proboscis	4	North	
301	398	Proboscis	4	North: Between Lot 7 & 8	
302	398	Proboscis	4	North: Between Lot 7 & 8	
303	398	Proboscis	4	North: Between Lot 7 & 8	
304	398	Proboscis	4	North: Between Lot 7 & 8	
305	398	Proboscis	4	North: Between Lot 7 & 8	
306	398	Proboscis	4	North: Between Lot 7 & 8	
307	398	Proboscis	4	North: Between Lot 7 & 8	
308	398	Proboscis	4	North: Between Lot 7 & 8	
309	398	Proboscis	4	North: Between Lot 7 & 8	
310	398	Proboscis	4	North: Between Lot 7 & 8	
311	398	Proboscis	4	North: Between Lot 7 & 8	
312	398	Proboscis	4	North: Between Lot 7 & 8	
313	303	Proboscis	4	South: Outside of Lot 4	
314	303	Proboscis	4	South: Outside of Lot 4	
315	303	Proboscis	4	South: Outside of Lot 4	
316	303	Proboscis	4	South: Outside of Lot 4	
317	294	Proboscis	4	South: Outside of Lot 4	
318	294	Proboscis	4	South: Outside of Lot 4	
319	294	Proboscis	4	South: Outside of Lot 4	
320	294	Proboscis	4	South: Outside of Lot 4	
321	295	Proboscis	5	South: Outside of Lot 4	
322	295	Proboscis	5	South: Outside of Lot 4	
323	295	Proboscis	5	South: Outside of Lot 4	
324	295	Proboscis	5	South: Outside of Lot 4	
325	385	Proboscis	5	North: Outside Lot 8	
326	385	Proboscis	5	North: Outside Lot 8	
327	385	Proboscis	5	North: Outside Lot 8	
328	385	Proboscis	5	North: Outside Lot 8	
329	291	Proboscis	5	North	
330	291	Proboscis	5	North	
331	291	Proboscis	5	North	
332	291	Proboscis	5	North	
333	291	Proboscis	5	North	
334	291	Proboscis	5	North	
335	291	Proboscis	5	North	
336	291	Proboscis	5	North	
337	366	Proboscis	5	South: Outside of Lot 10	
338	366	Proboscis	5	South: Outside of Lot 10	
339	317	Proboscis	5	South: Outside of Lot 3	
340	317	Proboscis	5	South: Outside of Lot 3	
341	317	Proboscis	5	South: Outside of Lot 3	
342	317	Proboscis	5	South: Outside of Lot 3	

343	317	Proboscis	5	South: Outside of Lot 3	
344	317	Proboscis	5	South: Outside of Lot 3	
345	317	Proboscis	5	South: Outside of Lot 3	
346	317	Proboscis	5	South: Outside of Lot 3	
347	317	Proboscis	5	South: Outside of Lot 3	
348	317	Proboscis	5	South: Outside of Lot 3	
349	PM01	Proboscis	5	North: Outside of Lot 2	
350	PM01	Proboscis	5	North: Outside of Lot 2	
351	PM01	Proboscis	5	North: Outside of Lot 2	
352	PM01	Proboscis	5	North: Outside of Lot 2	
353	373	Proboscis	5	South: Outside Lot 9	
354	373	Proboscis	5	South: Outside Lot 9	
355	316	Proboscis	5	South: Outside of Lot 3	
356	316	Proboscis	5	South: Outside of Lot 3	
357	316	Proboscis	5	South: Outside of Lot 3	
358	316	Proboscis	5	South: Outside of Lot 3	
359	316	Proboscis	5	South: Outside of Lot 3	
360	316	Proboscis	5	South: Outside of Lot 3	
361	355	Proboscis	5	South: Small tributaries	
362	355	Proboscis	5	South: Small tributaries	
363	48	Proboscis	5	South	
364	48	Proboscis	5	South	
365	48	Proboscis	5	South	
366	48	Proboscis	5	South	
367	48	Proboscis	5	South	
368	48	Proboscis	5	South	
369	48	Proboscis	5	South	
370	48	Proboscis	5	South	
371	48	Proboscis	5	South	
372	48	Proboscis	5	South	
373	121	Proboscis	5	North	
374	121	Proboscis	5	North	
375	121	Proboscis	5	North	
376	121	Proboscis	5	North	
377	121	Proboscis	5	North	
378	121	Proboscis	5	North	
379	361	Proboscis	5	South: Outside of Lot 10	
380	361	Proboscis	5	South: Outside of Lot 10	
381	361	Proboscis	5	South: Outside of Lot 10	
382	361	Proboscis	5	South: Outside of Lot 10	
383	361	Proboscis	5	South: Outside of Lot 10	
384	361	Proboscis	5	South: Outside of Lot 10	
385	134	Proboscis	5	North	
386	134	Proboscis	5	North	
387	314	Proboscis	5	South: Outside of Lot 3	
388	314	Proboscis	5	South: Outside of Lot 3	
389	314	Proboscis	5	South: Outside of Lot 3	
390	314	Proboscis	5	South: Outside of Lot 3	
391	314	Proboscis	5	South: Outside of Lot 3	
392	314	Proboscis	5	South: Outside of Lot 3	
393	314	Proboscis	5	South: Outside of Lot 3	
394	314	Proboscis	5	South: Outside of Lot 3	
395	314	Proboscis	5	South: Outside of Lot 3	
396	314	Proboscis	5	South: Outside of Lot 3	
397	99	Proboscis	5	South	
398	99	Proboscis	5	South	
399	99	Proboscis	5	South	
400	99	Proboscis	5	South	
401	322	Proboscis	6	North: Outside of Lot 4	
402	322	Proboscis	6	North: Outside of Lot 4	
403	322	Proboscis	6	North: Outside of Lot 4	

404	322	Proboscis	6	North: Outside of Lot 4	
405	322	Proboscis	6	North: Outside of Lot 4	
406	322	Proboscis	6	North: Outside of Lot 4	
407	322	Proboscis	6	North: Outside of Lot 4	
408	322	Proboscis	6	North: Outside of Lot 4	
409	101	Proboscis	6	North	
410	101	Proboscis	6	North	
411	101	Proboscis	6	North	
412	.101	Proboscis	6	North	
413	220	Proboscis	6	North	
414	220	Proboscis	6	North	
415	220	Proboscis	6	North	
416	220	Proboscis	6	North	
417	220	Proboscis	6	North	
418	220	Proboscis	6	North	
419	220	Proboscis	6	North	
420	220	Proboscis	6	North	
421	195	Proboscis	6	North	Sambungan 433-434
422	195	Proboscis	6	North	
423	195	Proboscis	6	North	
424	195	Proboscis	6	North	
425	219	Proboscis	6	South	
426	219	Proboscis	6	South	
427	219	Proboscis	6	South	
428	219	Proboscis	6	South	
429	219	Proboscis	6	South	
430	219	Proboscis	6	South	
431	219	Proboscis	6	South	
432	219	Proboscis	6	South	
433	195	Proboscis	6	North	
434	195	Proboscis	6	North	
435	193	Proboscis	6	South: Outside of Lot 4	
436	193	Proboscis	6	South: Outside of Lot 4	
437	193	Proboscis	6	South: Outside of Lot 4	
438	193	Proboscis	6	South: Outside of Lot 4	
439	193	Proboscis	6	South: Outside of Lot 4	
440	193	Proboscis	6	South: Outside of Lot 4	
441	193	Proboscis	6	South: Outside of Lot 4	
442	193	Proboscis	6	South: Outside of Lot 4	
443	193	Proboscis	6	South: Outside of Lot 4	
444	193	Proboscis	6	South: Outside of Lot 4	
445	58	Proboscis	6	North	
446	58	Proboscis	6	North	
447	58	Proboscis	6	North	
448	58	Proboscis	6	North	
449	58	Proboscis	6	North	
450	58	Proboscis	6	North	
451	58	Proboscis	6	North	
452	58	Proboscis	6	North	
453	58	Proboscis	6	North	
454	58	Proboscis	6	North	
455	58	Proboscis	6	North	
456	58	Proboscis	6	North	
457	58	Proboscis	6	North	
458	58	Proboscis	6	North	
459	58	Proboscis	6	North	
460	58	Proboscis	6	North	
461	103	Proboscis	6	South	
462	103	Proboscis	6	South	
463	103	Proboscis	6	South	
464	103	Proboscis	6	South	

465	103	Proboscis	6	South	
466	103	Proboscis	6	South	
467	103	Proboscis	6	South	
468	103	Proboscis	6	South	
469	64	Proboscis	6	South: outside of Lot 1	
470	64	Proboscis	6	South: outside of Lot 1	
471	64	Proboscis	6	South: outside of Lot 1	
472	64	Proboscis	6	South: outside of Lot 1	
473	64	Proboscis	6	South: outside of Lot 1	
474	64	Proboscis	6	South: outside of Lot 1	
475	64	Proboscis	6	South: outside of Lot 1	
476	64	Proboscis	6	South: outside of Lot 1	
477	64	Proboscis	6	South: outside of Lot 1	
478	64	Proboscis	6	South: outside of Lot 1	
479	64	Proboscis	6	South: outside of Lot 1	
480	64	Proboscis	6	South: outside of Lot 1	
481	11	Proboscis	7	South: Margin Danau Pitas	
482	11	Proboscis	7	South: Margin Danau Pitas	
483	11	Proboscis	7	South: Margin Danau Pitas	
484	11	Proboscis	7	South: Margin Danau Pitas	
485	11	Proboscis	7	South: Margin Danau Pitas	
486	11	Proboscis	7	South: Margin Danau Pitas	
487	11	Proboscis	7	South: Margin Danau Pitas	
488	11	Proboscis	7	South: Margin Danau Pitas	
489	11	Proboscis	7	South: Margin Danau Pitas	
490	11	Proboscis	7	South: Margin Danau Pitas	
491	11	Proboscis	7	South: Margin Danau Pitas	
492	11	Proboscis	7	South: Margin Danau Pitas	
493	102	Proboscis	7	North	
494	102	Proboscis	7	North	
495	102	Proboscis	7	North	
496	102	Proboscis	7	North	
497	102	Proboscis	7	North	
498	102	Proboscis	7	North	
499	102	Proboscis	7	North	
500	102	Proboscis	7	North	
501	102	Proboscis	7	North	
502	102	Proboscis	7	North	
503	102	Proboscis	7	North	
504	102	Proboscis	7	North	
505	102	Proboscis	7	North	
506	102	Proboscis	7	North	
507	102	Proboscis	7	North	
508	102	Proboscis	7	North	
509	102	Proboscis	7	North	
510	102	Proboscis	7	North	
511	102	Proboscis	7	North	
512	102	Proboscis	7	North	
513	102	Proboscis	7	North	
514	102	Proboscis	7	North	
515	102	Proboscis	7	North	
516	102	Proboscis	7	North	
517	102	Proboscis	7	North	
518	102	Proboscis	7	North	
519	102	Proboscis	7	North	
520	102	Proboscis	7	North	
521	64	Proboscis	7	South: Outside of Lot 1	
522	64	Proboscis	7	South: Outside of Lot 1	
523	64	Proboscis	7	South: Outside of Lot 1	
524	64	Proboscis	7	South: Outside of Lot 1	
525	64	Proboscis	7	South: Outside of Lot 1	

526	64	Proboscis	7	South: Outside of Lot 1	
527	64	Proboscis	7	South: Outside of Lot 1	
528	64	Proboscis	7	South: Outside of Lot 1	

Extract	KOD GPS	Species		Remark	Xtract
1	11	Proboscis		South: Margin Danau Pitas	New Extract 481-492
2	11	Proboscis		South: Margin Danau Pitas	New Extract 481-492
3	291	Proboscis		North	New Extract 329-336
4	291	Proboscis		North	New Extract 329-336
5	317	Proboscis		South: Outside of Lot 3	New Extract 339-348
6	317	Proboscis		South: Outside of Lot 3	New Extract 339-348
7	316	Proboscis		South: Outside of Lot 3	New Extract 355-360
8	316	Proboscis		South: Outside of Lot 3	New Extract 355-360
9	48	Proboscis		South	New Extract 363-372
10	48	Proboscis		South	New Extract 363-372
11	195	Proboscis		North	New Extract 421-424
12	195	Proboscis		North	New Extract 421-424
13	102	Proboscis		North	New Extract 493-520
14	102	Proboscis		North	New Extract 493-520
15	219	Proboscis		South	New Extract 425-432
16	219	Proboscis		South	New Extract 425-432
17	PM02(40)	Proboscis		North	Unique Extract
18	PM02(40)	Proboscis		North	Unique Extract
19	373	Proboscis		South: Outside of Lot 9	New Extract 353-354
20	373	Proboscis		South: Outside of Lot 9	New Extract 353-354
21	322	Proboscis		North: Outside of Lot 4	New Extract 401-408
22	322	Proboscis		North: Outside of Lot 4	New Extract 401-408
23	361	Proboscis		South: Outside of Lot 10	New Extract 379-384
24	361	Proboscis		South: Outside of Lot 10	New Extract 379-384
25	356	Proboscis		North	Unique Extract
26	356	Proboscis		North	Unique Extract
27	365	Proboscis		South: Outside of Lot 10	Unique Extract
28	365	Proboscis		South: Outside of Lot 10	Unique Extract
29	318	Proboscis		North	Unique Extract
30	318	Proboscis		North	Unique Extract
31	134	Proboscis		North	New Extract 385-386
32	134	Proboscis		North	New Extract 385-386
33	101	Proboscis		North	New Extract 409-412
34	101	Proboscis		North	New Extract 409-412
35	PM01(39)	Proboscis		North: Outside of Lot 2	New Extract 349-352
36	PM01(39)	Proboscis		North: Outside of Lot 2	New Extract 349-352
37	133	Proboscis		North	New Extract 191-192
38	133	Proboscis		North	New Extract 191-192
39	220	Proboscis		North	New Extract 413-420
40	220	Proboscis		North	New Extract 413-420
41	227	Proboscis		North	New Extract 203-204
42	227	Proboscis		North	New Extract 203-204
43	193	Proboscis		South: Outside of Lot 4	New Extract 435-444
44	193	Proboscis		South: Outside of Lot 4	New Extract 435-444
45	103	Proboscis		South	New Extract 461-468
46	103	Proboscis		South	New Extract 461-468
47	99	Proboscis		South	New Extract 397-400
48	99	Proboscis		South	New Extract 397-400
49	398	Proboscis		North: Between Lot 7 - 8	New Extract 301-312
50	398	Proboscis		North: Between Lot 7 - 8	New Extract 301-312
51	121	Proboscis		North: Between Lot 4 - 5	New Extract 373-378
52	121	Proboscis		North: Between Lot 4 - 5	New Extract 373-378
53	385	Proboscis		North: Outside of Lot 8	New Extract 325-328
54	385	Proboscis		North: Outside of Lot 8	New Extract 325-328
55	355	Proboscis		South: Small tributaries	New Extract 361-362
56	355	Proboscis		South: Small tributaries	New Extract 361-362

Appendix IV

Mitochondrial DNA control region haplotypes for the long-tailed macaque (left domain, LD and combined, CR), orang-utan (PP) and proboscis monkey (PM). Haplotype number is followed by the haplotype sample(s) and the sequence.

A. Long-tailed macaque (Left domain)

LD01. = 033S_01, 047S_78
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTGAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD02. = 061S_06, 064S_06
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCGTGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGTTCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD03. = 058N_05, 060N_05, 067S_06, 070S_06, 041N_07
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTCCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD04. = 076S_06
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACCCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD05. = 092S_PP, 093S_PP, 098S_PP, 111S_PP
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD06. = 035S_01, 107N_45, 117S_PP, 128S_PP
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD07. = 131S_PP
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATTATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAAGTACTCTAATGT

AAACCTCAACAGTAGTACATAATATGGCCTTTCCAAAGCTCAATCCACCCCCTCATGAAT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD08. = 153S_03
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATATGGCCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD09. = 180S_03
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATACAGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD10. = 147S_03, 149S_03, 157S_03, 159S_03, 163S_03, 167S_03, 170S_03, 172S_03,
         177S_03, 182S_03, 145S_06
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAGCATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD11. = 187S_06
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATGCGGTCCTTCCAAAGTTCAATCCATCCCCTCATGGAT
ACCAACTAAACCAATCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTGTTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD12. = L14S_01, L16S_01, 120S_PP, 121S_PP, 130S_PP, 134S_PP, 135S_PP, 138S_PP,
         140S_PP
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD13. = L19S_09, L23S_09
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
GAACTCTAACAGTAGTCCATAGCATGGTTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD14. = L25N_04, L32N_05, 013N_05, 078S_06, 082S_06, 114S_PP, 116S_PP, 142S_PP
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAATACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGGTGT
AAACTTTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCATCCCCCCATGAGT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG

CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD15. = L38S_03, 056N_05, 185N_07
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT

TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD16. = 003N_05, 005N_05, 008N_05, 015N_05, 017N_05, 020N_05, 046N_05, 051N_05, 054N_05, 079S_06, L10S_06, L34N_07, L40N_07, 038N_07, 144S_PP
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATGCGGTCCTTCCAAAGTTCAATCCATCCCCTCATGGAT
ACCAACTAAACCAATCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD17. = 011N_05, L46S_06
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAATAACATGCTTATAAGCAAGTACTCTGATGT
GAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTGTTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD18. = 100N_05, 104N_05, 106N_05, L47S_06, L02N_07
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD19. = L06S_45, L50N_45, 025N_07, 027N_07, 032N_07, 039N_07
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAGGCTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD20. = L28S_01, L53S_01
AACTTCATAAGATGACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LD21. = L58S_78, 021S_09, 023S_09
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

LABUK = LABUKN
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGCCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTA

B. Long-tailed macaque (combined left and right domain)

CR01. = 005N_05
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATGCGGTCCTTCCAAAGTTCAATCCATCCCCTCATGGAT
ACCAACTAAACCAATCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAATATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR02. = 013N_05
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAATACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGGTGT
AAACTTTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCATCCCCCCATGAGT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTCAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR03. = 023S_09
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAATATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR04. = 056N_05
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG

CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR05. = 061S_06, 064S_06

AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCGTGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGTTCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR06. = 058N_05, 060N_05, 067S_06, 070S_06, 041N_07,

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTCCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR07. = 093S_PP_098S_PP

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR08. = 104N_05

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT

TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATTGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTGATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR09. = 111S_PP

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR10. = 114S_PP, 116S_PP

AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAATACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGGTGT
AAACTTTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCATCCCCCCATGAGT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTGATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR11. = 107N_45, 117S_PP, 128S_PP

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA

```
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT
```

CR12. = 131S_PP

```
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATTATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAAGTACTCTAATGT
AAACCTCAACAGTAGTACATAATATGGCCTTTCCAAAGCTCAATCCACCCCCTCATGAAT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT
```

CR13. = 142S_PP

```
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGGTGT
AAACTTTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCATCCCCCCATGAGT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT
```

CR14. = 016N_05, 017N_05, 020N_05, 144S_PP

```
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATGCGGTCCTTCCAAAGTTCAATCCATCCCCTCATGGAT
ACCAACTAAACCAATCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTCAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT
```

CR15. = 153S_03

```
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATATGGCCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ATCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
```

CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR16. = 180S_03
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATACAGTCCTTCCAAAGTTCAATCCACCTCCCCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR17. = 147S_03, 149S_03, 157S_03, 159S_03, 163S_03, 167S_03, 170S_03, 172S_03, 177S_03, 182S_03, 145S_06
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAGCATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR18. = 100N_05, 106N_05, L02N_07 .
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATTGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

CR19. = L14S_01, 120S_PP, 121S_PP, 130S_PP, 134S_PP, 135S_PP, 138S_PP, 140S_PP
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR20. = L16S_01
AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTGATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR21. = L28S_01
AACTTCATAAGATGACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ACCATTACATATCAACATAACATTCCTGAATAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATTGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR22. = L38N_04, 185N_07
AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAAGTACTCTGATGT
AAACTCTAACAGTAGTCCATAACACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC

GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAATATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAGAAACTACCTCACAACTACACTAACACCCCT

CR23. = 003N_05, 008N_05, 046N_05, 052N_05, 054N_05, 079S_06, 038N_07, L34N_07, L40N_07

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATGCGGTCCTTCCAAAGTTCAATCCATCCCTCATGGAT
ACCAACTAAACCAATCCATGCCAGTCGTCCATAGTACATTAAGTCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATCTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAGAAACTACCTCACAACTACACTAACACCCCT

CR24. = 011N_05, L46S_06

AACTTCATAAGATAACCTTGATATCAACCTACCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGATTGTACATAACACAT
ATCATTACATATCAACATAACATTCCTGAATAACATGCTTATAAGCAAGTACTCTGATGT
GAACTCTAACAGTAGTCCATAGCATGGCTCTTCCAAAGTTCAACCCACCTCCCCATGAAT
ATCAACTAAACCAGCTCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTGTTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTCATTAATCGCACCTACGTTCAATATTCTAG
CTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
ATCTTAACCAAGTACCATTCTCACCACGCCAATAAACCACAACCATACCTCGTCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAGCTACCTCACAACTACACTAACACCCCT

CR25. = L06S_45, L50N_45, 025N_07, 027N_07, 032N_07, 039N_07

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTAATGT
AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAGGCTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATTGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAGAAACTACCTCACAACTACACTAACACCCCT

CR26. = L58S_78, 021S_09

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT

AAACTCTAACAGTAGTCCATAATACGGTCCTTCCAAAGTTCAATCCACCCCCTCATGGAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CGGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAATATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCAAAAGAAACTACCTCACAACTACACTAACACCCCT

LABUK = LABUKN

AACTTCATAAGATAACCTTGATATCAACCTATCCACAATATTACTATGTAATTCGTGCAT
TACTGCTAGCCAACATGTATAATATATAGTACTATATATGCTTGACTGTACATAACACAT
ATCATTACATATCAACATAACATTCCCGAACAACATGCTTACAAGCAGGTACTCTGATGT
AAACTCTAACAGTAGTCCATAATACGGCCCTTCCAAAGTTCAATCCACCTCCTCATGAAT
ACCAACTAAACCAGTCCATGCCAGTCGTCCATAGTACATTAAATCGTTCATCGGACATAG
CACATATCTATTAAATAATCCTCCTCACCACGGATGCCCCCCCTCACTTACGATGGATCA
CAGGTCTATCACCCTATTAACCAGTCACGGGAGATTTCCATGCATTTGGTATCTTTTATC
TCTGGTCCGCACGCAACCCCATCGCAGAATGCTGACTCCCACCACATCCCGTCCTGAATG
CGCCTGTCTTTGATTCCTAGTACATGCAGTTATTAATCGCACCTACGTTCAATATTCTAG
TTCCACGCAAACCTTAGCAGGGTGTTATTTAATCCATGCTTGTAGGACATATCAATAATC
GTCTTAACCAAATACCATTCTCACCACGCCAATAAACCACAACCATACCTCATCAAACCC
CCCCACCCCCATCTCTGACCTTCATCCAAAACCCACTCTTGCCAAACCCCAAAAACAAAA
GTCTTAACATATCCGATCAGAGCTCGCGTTTTTATCTTTTAGGTGTGCACAACTCCAACT
GCCATTCTCTCAACTAACAAACATTTACCTCACTAAACGCCCCTCACACCAACCCATAAT
AAACCCTTCTCACACAACCCGAAAGAAACTACCTCACAACTACACTAACACCCCT

C. Orang-utan (left domain)

OU1. = 143N_05
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCTACGGATGCCCCCCCT
```

OU2. = 153N_02
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAAATCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU3. = 064S_06
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAGATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACTAACCCAACCAAGCT
TTAAAGTACATAGCGCATAACACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU4. = 003S_01, 109S_01
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAATACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTGAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU5. = 134S_01, 057S_06, 088S_09
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAATACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU6. = 135S_01
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAAATCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU7. = 170N_05
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAGATACA
TACATCAATCCCCCCACATAACCCCTTCTCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT
```

OU8. = 185N_04
```
CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
```

AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAGATACA
TACATCAATCCCCCCACATAACCCCTTCTCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCTACGGATGCCCCCCCT

OU9. = 140S_03, 201S_03, 205S_03

CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACTAACCCAACCAAGCT
TTAAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT

OU10. = 145N_02, 150N_02, 197N_02, 218N_02, 028S_09

CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAGATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCTACGGATGCCCCCCCT

OU11. = 228N_02, 112N_02, 161N_02, 148S_03, 106S_03, 203S_03, 035N_04, 168N_04, 082N_04, 085N_04, 172N_04, 182S_04, 210N_04, 002N_04, 067N_04, 032N_04, 124N_05, 132N_05, 046N_05, 052N_05, 070S_06, 038S_06, 071S_06, 054N_07, 074N_08, 090N_08, 212S_09, 060N_10

CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAGATACA
TACATCAATCCCCCCACATAACCCCTTCTCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCTACGGATGCCCCCCCT

OU12. = 014S_01, 077S_01, 128S_01, 130S_01, 122S_01, 192S_01, 199S_01, 208S_01, 113S_03, 030S_03, 138S_03, 190S_03, 039AS_06, 042S_06, 101S_06, 158S_06, 098S_06, 048N_08, 008S_09, 039BS_09, 091S_09, 010N_10, 012N_10, 022N_10, 023N_10

CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAACACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCCCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCACATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCCACGGATGCCCCCCCT

OU13. = 100N_05

CCCAGTACCGACCCATTTCCAGCGGCCTATGTATTTCGTACATTCCTGCCAGCCAACATG
AATATCACCCAACACAACAATCGCTTAACCACCTATAATACATACAAAGCCCAATCCACA
CCCAACCTCTACCCCCCGCTTACAAGCAAGTACCCCCCCATGCCCCCCCACCCAAATACA
TACATCAATCCCCCCACATAACCCCTTCTCCCCCCGCATACCAACCAACCCAACCAAGCT
TTAAAGTACATAGCGCATAATACCCCTACCGTACATAGCACATTTCTACTAACTCCCTGC
TTAACCCTACGGATGCCCCCCCT

D. Proboscis monkey (left domain)

PM01. = 221N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGATGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM02. = 225N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGATGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGACATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM03. = 241N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGATACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGAAAATATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACACAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM04. = 259S_06
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACTAACCGATACTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM05. = 265N_BD, 267N_BD
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM06. = 269N_10
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGATGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATCACTATTGAGC
ATCCCTAAAACAAT

PM07. = 315N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAGCCCACCGGAATACCAACCGATACTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM08. = 219N_04, 245N_04, 323N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGACATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM09. = 327N_08
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGCTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCCTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM10. = 331N_04
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTACATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGACATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM11. = 337N_10
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCCTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATCCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM12. = 360S_03
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGTCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM13. = 370S_01, 229S_03, 283S_03, 343S_03, 217N_04, 317N_04

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM14. = 379N_10, 383N_10

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTAACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM15. = 281S_03_387S_03

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGTACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM16. = 401N_04

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATTTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM17. = 407N_04

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM18. = 279S_03, 417N_05

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATTTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATATCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM19. = 431S_06

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATACTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM20. = 469S_01, 453N_02
```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCTAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATTATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM21. = 488S_01
```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGTACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM22. = 497N_05
```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATACTATATA
TTCATTGATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM23. = 363S_01, 473S_01, 479S_01, 523S_01, 445N_02, 199N_02
```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAGACAAT
```

PM24. = 285S_03, 330N_04, 385N_04, L03N_04
```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAACATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATTTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAATAAT
```

PM25. = L10S_01
```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
```

```
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM26. = 421N_05, 423N_05, 433N_05, L11N_05

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAATATCTGACTATAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGCTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAGT
```

PM27. = L13N_05

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCTAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM28. = 205S_01, 209S_01, 289S_01, 339S_03, L05S_03, 382N_10, L25N_10

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM29. = L27N_10

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TCTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM30. = 521S_01, 449N_02, 459N_02, 297N_04, 313N_04, L29N_04

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM31. = L31N_04, 411N_05, 253S_06, 255S_06, 301N_78

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
```

TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATACTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM32. = 291S_01, 351N_02, 193N_02, 249N_04, L33N_05, L15S_06
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGATGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGACATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM33. = 481S_01, 485S_01, 490S_01, L01S_01, 190N_02, 195N_02, L35N_02, 437N_04, 425S_06, 309N_78, L19S_09
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATACTATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM34. = 236N_04, 321N_04, 191N_05, L37N_05
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACTGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM35. = L39N_05
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATTTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATATCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGACATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM36. = 203N_04, L41N_04, 461S_06, 465S_06
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCACTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM37. = L43N_04, 273N_10
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG

```
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM38. = L22N_04, L45S_06

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTCCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM39. = 233S_03, 389S_03, 237N_04, 295N_04, 403N_04, 435N_04, 501N_05, 505N_05, 511N_05, 397S_06, 399S_06, L47S_06

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM40. = 305N_78, L49N_78

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TCTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM41. = 373N_45, 375N_45, L51N_45

```
ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATAGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATACA
TTCATTAATCGTACATAGTACATTAGATTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM42. = 207S_01, 212S_01, L07S_03, 439N_04, 441N_04, 443N_04, L18N_04, L53N_08, 325N_08, L24N_10

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCGTCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAATATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT
```

PM43. = L56N_10

```
ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
```

GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM44. = 367S_01, 394S_03, 357S_03, 361N_10, L58N_10

ATCTTTCCCCAGGGCAACTCAGAAAGAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAGCATGAATA
TTATATAGTACTATAAATGTTTTATCGTCCATAAGACATAAAATTACATATTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGGACCTTCATAGAAGTATCTGACTACAATACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT

PM45. = LabukN

ATCTTTCCCCAGGGCAACTCAGAAAAAGAGCACTCAACTCCACCACCAACACCCAAAATT
GGCATTCTATTTAAACTACTTTCTGTATTCTAGGGGGTACAACTCAAAAAATAACTCAAG
TACAATCTAGCTTTATATGCCCCTATGTAATTCGTGCATTACTGCTAGCCAACATGAATA
TTATATAGTACTATAAATGTTTTACCGTCCATAGGACATAAAATTACATACTTACTAACA
TTTCATCCAGAACATGCTTACAAGCAAGAACCTTCATGGAAGTATCTGACTACAACACAT
ATCATTCAAGCCTTCAAATACCATGGTGTAACCCACCGGAATACCAACCGATGTCATATA
TTCATTAATCGTACATAGTACATTAGGTTCTTTATCGGGCATAGCACATTACTATTGAGC
ATCCCTAAAACAAT