Molecular evolution of primates – featuring mobile elements

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von

Dipl. Biol.
Martin Osterholz

aus
Brunsbüttel

Göttingen, September 2008

Referent: Prof. Dr. Peter Kappeler
Korreferent: Prof. Dr. Ernst Wimmer
Tag der mündlichen Prüfung: 22.10.2008

Contents

1 General Introduction

1.1 Who with Whom – An Introduction into Primate Systematics

The order Primates is one of the most diverse mammalian orders. With the exception of the genus *Homo*, primates are mainly distributed in tropical and subtropical regions of Africa, Asia and Central and South America. The order was first recognized in the 18th century by Linnaeus and was divided into four different groups with 35 species. Today, the number of species varies from 230 (Rowe, 1996) to more than 350 (Groves, 2001; Geissmann, 2003) and is solely exceeded in number by Rodentia and Chiroptera. However, the identification of new species is still in progress. In the last few years, several new subspecies, species and even genera were described (Jones et al., 2005; Kappeler et al., 2005; Sinha et al., 2005; Andriaholinirina et al., 2006; Davenport et al., 2006; Wallace et al., 2006; Boubli et al., 2008; Roos et al., 2008). This phenomenon is not only based on the application of molecular methods, which allow to determine genetic diversity, but also on the increasing numbers of behavioral and ecological studies and the discovery of taxa in former unexplored areas.

The definition of the order Primates and its classification in its own order beyond other mammals is a complex issue, at least based on morphological characteristics. Mivart (1873) removed Chiroptera and Dermoptera from Primates and proposed different significant features leading to their uniqueness among other mammals. Examples are the retention of the clavicle in the pectoral girdle, a reduced number and three different types of teeth, a caecum, a penis pendulous and testes scrotal. Opposing thumbs combined with fingernails rather than claws afford gripping branches. However, not all these characteristics are synapomorphic and can also be found to some degree in non-primate orders (Geissmann, 2003). For a more adequat characterization of primates Martin (1990) provides further features. Primates have a trend towards reduced snout and olfactory senses. In contrast, the visual system is complex with high acuity and color vision and the brain is large compared to body size. Most of these characteristics can not be detected in fossils and therefore it is difficult to determine affiliations of fossil records to extant primates or to the origin of all primates respectively. It is actually unlikely to determine the origin and age of the order, because only few fossil records are available. It is assumed that only 7 % of all extinct primates are known from fossils (Tavaré et al., 2002). However, based on these few records (e.g. *Purgatorius*, North America), the most recent common ancestor most likely lived about 65 million years ago (mya) (Shoshani et al., 1996). However, by combining fossil records with statistical analyses, the origin of primates was recently dated back to 81.5 mya (Tavaré et al., 2002).

Genetic analyses form another possibility to detect primate-specific characteristics. Based on these studies, it is now widely accepted that the order Primates is a monophyletic group tightly clustering together with Scandentia and Dermoptera (Murphy et al., 2001a; Murphy et al., 2001b; Schmitz et al., 2005; Kriegs et al., 2006). However, mitochondrial sequence analyses indicated that dermopterans form a clade with anthropoid primates to the exclusion of other primates (Murphy et al., 2001a; Arnason et al., 2002; Schmitz et al., 2002a), leading to the assumption that primates are paraphyletic. Other studies using mobile elements, which insert into the genome and serve as molecular synapomorphic traits, clearly confirmed the monophyly of the order to the exclusion of dermopterans and other mammalian orders (Schmitz et al., 2002a; Schmitz and Zischler, 2003). Moreover, it was shown that Alu elements are a special type of short interspersed elements (SINEs), which are only present in primates, representing therefore another hint for a common origin of primates (Schmitz et al., 2002a).

Primates were originally divided into Prosimii (lemurs, lorises and tarsiers) and Anthropoidea (monkeys and apes), mainly based on morphological similarities between tarsier, lemurs and lorises. The affiliation of the tarsiers was one of the most basic questions in primate phylogeny. Even sequence data were not able to resolve this problem with significance (Andrews et al., 1998; Goodman et al., 1998; Schmitz et al., 2002b). Finally, SINE insertions provided strong molecular evidence for a sister relationship of tarsier to monkeys and apes, leading to the division into Strepsirrhini and Haplorrhini (Kuryshev et al., 2001; Schmitz et al., 2001; Schmitz et al., 2005). Morphologically, Strepsirrhini differ from Haplorhini in having a post-orbital bar, a moist nose, a tapetum lucidum (except some Lemuriformes taxa) and a dental comb. Hence, the current consensus view distinguishes two suborders, the "wet-nosed" (Strepsirrhini) and "dry-nosed" primates (Haplorrhini) (Groves, 2001; Geissmann, 2003), which diverged from each other 63-78 mya – depending on the applied calibration points (Goodman et al., 1998; Steiper and Young, 2006).

Strepsirrhini are a remarkable diverse group, comprising actually about 21% of all primate genera. They are traditionally divided into two infraorders, the Malagasy Lemuriformes (lemurs) and the Lorisiformes (lorises and galagos), which occur in Africa and Asia (Geissmann, 2003). They diverged from each other about 50-57 mya (Goodman et al., 1998; Roos et al., 2004; Steiper and Young, 2006). However, Groves (2001) separates the aye-aye from the Lemuriformes (*Daubentonia madagascariensis*) in its own infraorder Chiromyiformes, which seems to be appropriate due to several morphological autapomorphies, chromosomal and genetic studies and the early separation from the other Lemuriformes (Yoder et al., 1996; Groves, 2001; Roos et al., 2004; Warter et al., 2005; Mittermeier et al., 2006; Horvath et al., 2008). The respective monophyly of Lorisiformes and Lemuriformes is confirmed by several morphological features (Groves, 2001; Mittermeier et al., 2006) and molecular analyses (Yoder et al., 1996; Roos et al., 2004; Warter et al., 2005; Horvath et al., 2008). Besides the aye-aye, lemurs are further divided into four different families (Lemuridae, Indriidae, Lepilemuridae and Cheirogaleidae), whereas the lorises are classified into two families (Galagonidae and Lorisidae) (Groves, 2001).

Haplorrhini are divided into Tarsiiformes (tarsiers) and Anthropoidea, which include Platyrrhini (New World monkeys) and Catarrhini (Old World monkeys, apes and humans). The divergence for Tarsiiformes occurred relatively early about 58 mya (Goodman et al., 1998).

Platyrrhini are a tremendously diverse group distributed in South and Central America and diverged from the Catarrhini about 40-43 mya (Goodman et al., 1998; Steiper and Young, 2006). This diversity is manifested in the presence of up to thirteen sympatric species (Fleagle, 1988). There is no consensus about how many families should be recognized, although most recent studies arrange them into three families (Cebidae (e.g. marmosets, tamarins and squirrel monkeys), Atelidae (howler, spider and woolly monkeys) and Pitheciidae (titis, sakis and uakaris), Opazo et al., 2006; Goodman et al., 1998). Alternative classifications with two (Napier and Napier, 1967), four (Groves, 2001) or even five families (Geissmann, 2003) were also proposed. The branching order of Cebidae, Atelidae and Pitheciidae is still an intensely discussed issue, because morphological as well as molecular studies produced incongruent results (e.g. Kay, 1990; Opazo et al., 2006; Ray et al. 2005; Rosenberger, 1981; Schneider et al., 1996). It is unclear whether Pitheciidae or Atelidae form the most basal family. Moreover, it is not evident whether Cebidae, including the subfamily Callithrichinae, really form a monophyletic unity.

Within the Catarrhini two different superfamilies can be distinguished, with tail (Cercopithecoidea) and without tail (Hominoidea). They diverged from each other about 23-30 mya (Goodman et al., 1998; Raaum et al., 2005; Steiper and Young, 2006). Old World monkeys or Cercopithecidae, the single family within Cercopithecoidea, are further divided into two different subfamilies: Colobinae (leaf-eating monkeys) and Cercopithecinae (cheek-pouched monkeys) (Groves, 2001).

Both subfamilies diverged from each other about 14-16 mya (Goodman et al., 1998; Raaum et al., 2005), are highly diverse and are distributed in Africa and Asia, although Cercopithecinae are more diverse in Africa and Colobinae more in Asia. Both have different adaptation to their diet, leading to different anatomical features. Colobine monkeys have a complex multichambered stomach particularly to digest leaves and fruits. In contrast, Cercopithecinae possess cheek pouches and are omnivorous (Strasser and Delson, 1987). Colobine monkeys can further be divided into an African clade (tribe Colobini) and an Asian clade (tribe Presbytini) (Groves, 2001). Latter include a langur (e.g. hanuman langur and delacour langur) and an odd-nosed monkey group (e.g. proboscis monkey and snub-nosed monkey). Especially in the evolutionary history of colobine monkeys are a lot of debated and unanswered questions, e.g. the respective monophyly of Asian and African colobines. Unfortunately, only few molecular and morphological studies provide data for them, since material is limited due to their partial endangered status and difficulties in captive breeding. Cercopithecinae are divided into the two tribes Cercopithecini (vervet monkeys and guenons) and Papionini (macaques, baboons, mandrills, gelada and mangabeys), mainly based on chromosomal (42 chromosomes in Papioni, variable number in Cercopithecini) and behavioral data (Groves, 2001). A long standing issue

in papionine phylogeny was the paraphyly of the mangabeys (Disotell, 1994). In cercopithecine phylogeny molecular data provided evidence for an arboreal-terrestrial split (Xing et al., 2007).

The superfamily Hominoidea is divided into two families: lesser apes or gibbons (Hylobatidae) and great apes and humans (Hominidae) (Groves, 2001). Gibbons are divided into four different genera (Roos and Geissmann, 2001) and form the first split about 18 mya (Goodman et al., 1998) and are only distributed in Asia. They are well known for their characteristic brachiation movement and singing duets. Traditionally, great apes and human were divided into two families, Pongidae including all great apes and Hominidae, which comprises only human (Napier and Napier, 1967). However, molecular and also morphological studies led to a revision of this classification, including now gorilla and chimpanzee in the Homininae (Goodman et al., 1990; Groves, 2001; Salem et al., 2003) and the orang-utan in Ponginae, while latter diverged from the Homininae about 14-18 mya (Goodman et al., 1998; Steiper and Young, 2006). Among the African great apes, the gorillas were separated from the lineage leading to chimpanzee and human about 7-9 mya (Goodman et al., 1998; Steiper and Young, 2006). The latter two diverged from each other about 6 mya (Goodman et al., 1998; Raaum et al., 2005; Steiper and Young, 2006).

1.2 Molecular Systematics

Systematics is an essential biological discipline dealing with the study of biodiversity and evolutionary relationships of organisms in past and present. In 1758, Linnaeus *Systema Naturae* forms the first fundament for a hierarchic system of life. Wallace (1858) and Darwin (1859) independently expanded this fundament with the theory of natural selection as the reason for speciation. Based on the cladistic concept of Hennig (1950), particularly derived traits (apomorphies), instead of ancestral traits (plesiomorphies), are used to determine the phylogenetic relationships of organisms and to classify them. Systematics is not only the basis for all studies of organism, but also relevant for conservation interests when for instance evaluating the preservation of genetic diversity, i.e. determining evolutionary siginificant units (Ryder, 1986). Molecular systematics started in the middle of the 20th century to expand the knowledge on the evolutionary relationships of organisms, which was established by classical disciplines in terms of morphology, paleontology or behavioral studies. First attempts were done by studying structures of proteins or chromosomes. These techniques were partly replaced and expanded by DNA-DNA hybridization (Britten and Kohne, 1968), although this method showed only distances between two organisms instead of distinct traits. Another method represented RFLP (restriction fragment length polymorphism), with which it was possible to detect variations in the DNA sequence via restriction enzymes. After DNA sequencing became an important tool in the 70s and PCR in the 80s, the analysis of multiple sequences displaced many of the former methods and was further improved by numerous statistical techniques. However, all these different methods are limited by homoplasy, a character similarity, which is not derived from a common

ancestor leading to misgrouping of taxa. Future research in systematics will focus not only on marker genes, but also on comparative whole genome analyses, as DNA sequencing has been improved in efficiency (highthroughput DNA sequencing). But evolutionary analyses based on sequence information from complete genome can also be hampered by homoplasies, whereas rare genomic changes (e.g. random insertions and deletions (indels) or retrotransposons) form more reliable tools. Homoplasies were reported however for indels as well (de Jong et al., 2003), in contrast to retrotransposons.

In the 90s the application of mobile elements, in particular retroposons, as a phylogenetic marker system was proposed with almost homoplasy-free character (Ryan and Dugaiczyk, 1989; Okada, 1991; Murata et al., 1993). Shedlock and Okada (2000) even classified these elements as nearly perfect genetic apomorphies in terms of Hennig (1950). Since that time, retroposons have helped to elucidate several controversial phylogenetic relationships in mammalian, avian, reptile and fish phylogeny (Shimamura et al., 1997; Takahashi et al., 1998, 2001; Roos et al., 2004; Schmitz et al., 2005; Xing et al., 2005; Kriegs et al., 2006; Piskurek et al., 2006; Kriegs et al., 2007; Matveev et al., 2007; Xing et al., 2007).

1.2.1 Sequence Analyses

Since the establishment of PCR and sequencing methods in the 80s, the analysis of nucleotide sequences of marker genes has become one of the standard techniques in molecular systematics. These techniques have been used to compare nucleotides at homologous positions in the mitochondrial as well as in the nuclear DNA of different organisms. Every nucleotide position forms a character with four different options (adenine, thymine, guanine and cytosine). Particularly non-coding regions are part of analyses, since they do not underly a selective pressure. However, this method is limited by base composition similarities (Schmitz et al., 2002a) or when identical characters arise parallel in species without common ancestor (homoplasy). To reconstruct phylogenetic trees from nucleotide data different statistical models can be used like Maximum-Parsimony (MP), Maximum-Likelihood (ML), Neighbor-Joining (NJ) and Bayesian inference (Felsenstein, 1978, 1981; Saitou and Nei, 1987; Huelsenbeck and Ronquist, 2001). Further statistical methods like Bootstrapping (Felsenstein, 1985), Jackknifing (Farris et al., 1996) and Bayesian posterior probabilities (Huelsenbeck and Ronquist, 2001) can estimate the statistical support for the obtained tree topologies.

The choice of marker genes depends on the analysed evolutionary time window. The more variable the gene is, the more recent should be the analysed phylogenetic relationships. Ribosomal genes, which are relatively conserved, can be used for analysing splits more than 100 mya. Mitochondrial DNA has a constant mutation rate, is inherited maternally and does not recombine. Mutations accumulate 5-10 times faster than in nuclear DNA (Brown et al., 1982). Hence, mitochondrial DNA can rather be used for closer related taxa or population genetics. Analysing relative ancient splits using mitochondrial sequences could lead to unreliable results. The high copy number of mitochondrial genomes per cell (Cann et al., 1987) constitutes an advantage for analysing non-invasive samples (e.g. feces). However, the analyses of mitochondrial DNA can be hampered by

possible misamplifications of nuclear pseudogenes (Collura and Stewart, 1995; Zischler et al., 1995; Mourier et al., 2001). As nuclear marker genes several intronic regions of different genes have been proven to be useful for various phylogenetic reconstructions (Flynn and Nedbal, 1998; Chaves et al., 1999; Hellborg and Ellegren, 2003; Yoder et al., 2003). Evolutionary processes could be affected by hybridization events between different subspecies, species or even genera. To reveal hybridization events it is appropriate to analyse marker genes with maternal (mitochondrial DNA), paternal (Y-chromosomal DNA) as well as biparental (autosomal DNA) inheritance. Another general problem in these evolutionary analyses is incomplete lineage sorting (differential sorting of ancestral polymorphisms into the progeny lineages), particularly when analysing rapid radiations.

1.2.2 Mobile elements – temporal landmarks in evolution

As a result of sequencing numerous eucaryote genomes within the last years it became obvious that not more than 5 % of the genome codes for proteins. In fact, more than 40 % of the human and non-human genomes consist of a variety of repetitive sequences with apparent no global function (Smit and Riggs, 1995; Lander et al., 2001; Gibbs et al., 2007). These genome components represent remnants of elements, which have the ability to transport or duplicate themselves to other genomic locations. It is assumed that these mobile elements play a major role in shaping genomes during evolution. Mobile elements can be divided into two distinct groups, based on their mobility process. DNA transposons move to their target site by a cut and paste mechanism via a DNA intermediate, whereas retrotransposons mobilize through an RNA intermediate and a copy and paste mechanism (Shedlock and Okada, 2000; Kazazian, 2004; Shedlock et al., 2004). DNA transposons have been shown to be inactive for the last 40 million years, at least in primates (Pace and Feschotte, 2007).

Retrotransposons can further be divided into autonomous elements, which encode for their own replication machinery, and non-autonomous elements, which depend on the enzymatic machinery provided by other elements. The former group includes non-LTR (long terminal repeat) elements (LINEs, long interspersed elements) and LTR elements (retroviruses), whereas latter includes SINEs (short intersperse elements) (Shedlock and Okada, 2000; Kazazian, 2004; Shedlock et al., 2004). The retropositionary process is initiated by generating a RNA copy of the original element. This RNA intermediate is reverse transcribed into DNA by a process called target primed reverse transcription, followed by reintegration into a new genomic location. It is suggested that the retropositionary process of SINEs is mediated by a reverse transcriptase and endonuclease activity provided by LINEs (Luan et al., 1993; Feng et al., 1996; Jurka, 1997; Kazazian and Moran, 1998; Kajikawa and Okada, 2002; Kazazian, 2004). Furthermore, as a result of this process target site duplications (direct repeats) are generated.

SINEs are further divided into two different groups, either 7SL RNA or tRNA derived elements. Alu elements are primate-specific dimeric 7SL RNA-derived elements and first appear at the base of in primate evolution about 65 mya. They emerge to the elements with the highest copy number (at least one million in the genome of macaque and human) and a broad taxonomic distribution (Lander et al.,

2001; Li et al., 2001; Gibbs et al., 2007). With a length of about 300 basepairs (bp) they form the majority of SINEs and can be found mainly in untranslated and intergenic regions and introns (Deininger and Batzer, 1993). During primate evolution, Alu elements propagated in the genome in a chronological wave-like fashion and can be classified into three different temporally active subfamilies depending on specific mutations: AluJ (oldest), AluS (intermediate) and AluY (youngest) (Batzer et al., 1996). Therefore it was suggested that only a small subset of Alu elements, referred to as master-genes, are retropositional active in certain time windows (Deininger et al., 1992). However, this "master gene hypothesis" is in contrast to the "multiple source gene model" of Schmid and Maraia (1992), which assumed that also progeny Alu elements have the same ability for retroposition.

Regarding the size of a primate genome, the coincidence of an orthologous, independent and irreversible insertion of an Alu element into the same locus is insignificant. Furthermore the true orthology can be verified by their flanking direct repeats and their subfamily. Precise removal is extremely unlikely in multiple genomes and was only shown to be caused by recombination of perfect direct repeats (van de Lagemaat et al., 2005). These exact deletions were reported solely for few cases from very recent retroposed elements (Sen et al., 2006). Consequently, the lack of an Alu element at a certain locus can be considered to be the ancestral state and it is possible to determine a common ancestor by a shared insertion. However, SINE based phylogenetic analyses have some disadvantages. The major limitation is the uneven distribution of an Alu element into progeny lineages, caused by a non-fixed and unspecific insertion in a population, a phenomenon called incomplete lineage sorting. However, if there are no contradicting results, a significant support for a specific branching will be obtained by the detection of at least three independent insertions per branch (Waddell et al., 2001). Moreover, Alu elements are uninformative concerning estimation of divergence times, even though subfamilies give a hint for that, due to their activity in a certain time window (Schmitz et al., 2005).

Based on the characteristics outlined above, SINE insertions represent an effective alternative way to reconstruct phylogenies. This is well reflected in a variety of studies (Shimamura et al., 1997; Takahashi et al., 1998; Nikaido et al., 1999; Takahashi et al., 2001; Roos et al., 2004; Nishihara et al., 2005; Ray et al., 2005; Schmitz et al., 2005; Xing et al., 2005; Kriegs et al., 2006; Piskurek et al., 2006; Kriegs et al., 2007; Matveev et al., 2007; Xing et al., 2007), which for example provide decisive evidence for the position of whales within Cetartiodactyla (Nikaido et al., 1999), the phylogenetic relationships among strepsirrhine genera (Roos et al., 2004) or the monophyly of several mammalian clades (Kriegs et al., 2006).

1.3 Mobile Elements in Primate Evolution

Considering the advantages and features of mobile elements as a phylogenetic marker system outlined above, several studies, mainly using Alu elements, were already performed to resolve several important questions in primate evolution (Schmitz et al., 2001; Salem et al., 2003; Schmitz and Zischler, 2003; Roos et al., 2004; Ray et al., 2005; Schmitz et al., 2005; Xing et al., 2005; Xing et al., 2007). Aim of this thesis was to use this technique to answer the most important phylogenetic issues in primate evolution. Hence, for an overview recent progress in mobile element based primate phylogeny and remaining questions in primate evolution will be outlined in order of branching events.

As already mentioned in Chapter 1.1, several studies based on retroposons provided evidence for the position of primates as a monophyletic group within the cohort Archonta, the position of the tarsier and the major division into Strepsirrhini and Haplorhini (Kuryshev et al., 2001; Schmitz et al., 2001; Schmitz et al., 2002a; Schmitz and Zischler, 2003; Schmitz et al., 2005).

Strepsirrhine primates are divided into Lemuriformes and Lorisiformes. Ongoing questions were the origin of strepsirrhines, whether Madagascar and Asia were colonized by single immigration events and the affiliations of the aye-aye (*Daubentonia*) among strepsirrhines. Roos et al. (2004) performed a SINE based study and inferred a strong supported phylogenetic tree. They demonstrated that lorisiforms and lemuriforms form reciprocal monophyletic groups, which argued for single colonization events and for an African origin of strepsirrhines. This scenario is also supported by Yoder (1996) and Karanth et al. (2005). The results proved a sister relationship between the aye-aye and the remaining Lemuriformes. Herke et al. (2007) partially confirmed the study of Roos et al. (2004). However, both studies were not able to elucidate confidently further branching orders among Lemuriformes families and the relationships within the Galagidae.

Phylogenetic relationships among New World monkey genera are highly disputed, because morphological and molecular studies reveal contradicting results. The branching order of the families Cebidae, Atelidae and Pitheciidae has not been confidently resolved using classical techniques. Hence, two SINE based analyses were performed (Singer et al., 2003; Ray et al., 2005). The study of Singer et al. (2003) merely indicated a monophyletic origin for all New World monkeys, for the Cebidae and for the Callithrichinae. Ray et al. (2005) confirmed these results and provided strong evidence for a closer affiliation of Cebidae and Atelidae and a basal position of Pitheciidae. Furthermore the study resolved some phylogenetic relationships within Pitheciidae and Atelidae. While the first study was not comprehensive enough, the second study did not include all genera. Evidence for the phyletic unity of the Cebidae is still lacking. Moreover there are no SINE based studies regarding the phylogenetic relationships within the Callitrichinae and their affiliations to the genera *Saimiri*, *Aotus* and *Cebus* at hand.

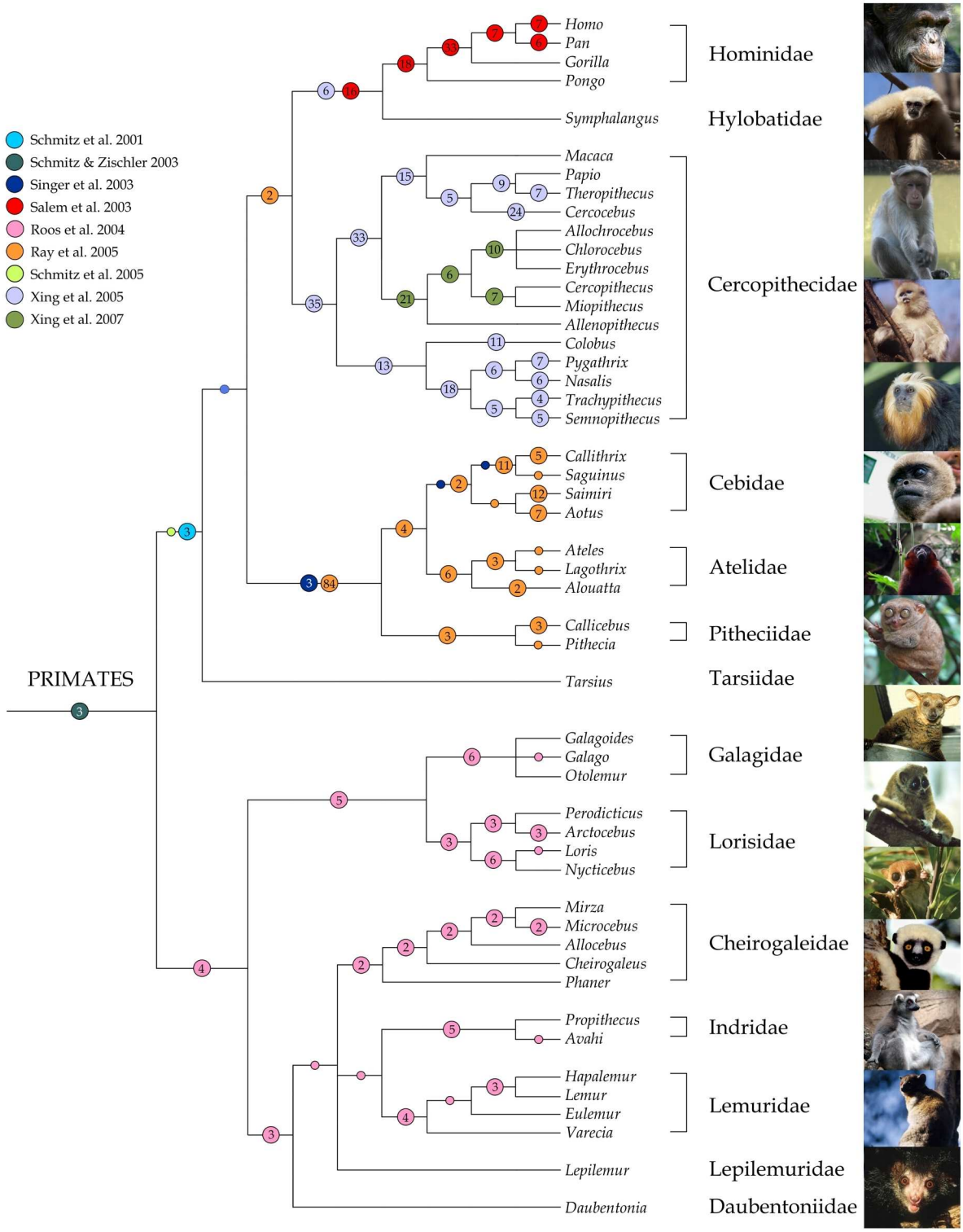

Figure 1.0: Primate phylogeny based on published mobile elements. Numbers of integrations per branch are indicated by small circles (one integration) and large circles (number of integrations inside)

Within the Old World monkeys there is little disagreement between morphological and molecular data regarding the higher-level phylogenetic affiliations within Cercopithecinae (tribes and subfamilies). Crucial disagreements in Old World monkey phylogeny can be found in colobine monkeys. This concerns especially the question of respective monophyly of Asian and African clades

and several genus affiliations. Two studies based on SINE insertions were performed to resolve questions in Old World monkey phylogeny (Xing et al., 2005; Xing et al., 2007). The first study provides strong evidence for the major clades in colobine monkeys (although solely *Colobus* was tested as African genus) and cercopithecines and was in agreement with former molecular or morphological studies (Delson, 1992; Goodman et al., 1998; Zhang and Ryder, 1998; Page et al., 1999; Page and Goodman, 2001; Sterner et al., 2006). They provided evidence for the monophyly of Papionini and Cercopithecini. Furthermore *Papio* was confirmed as sister genus to *Theropithecus*. *Cercocebus* forms the sister clade to *Papio* and *Theropithecus* and *Macaca* to all of them. However, *Lophocebus*, *Mandrillus* and the recently described *Rungwecebus* were not implemented in their study. Phylogenetic relationships within the Asian colobines were conclusively shown with *Trachypithecus* and *Semnopithecus* (*Trachypithecus vetulus*) as well as *Pygathrix* and *Nasalis* as sister clades. However, most exciting and interesting questions in colobine evolution, e.g. monophyly of African and Asian colobines, could not be examined, since relevant genera were not included. The second study of Xing et al. (2007) focused solely on the Cercopithecini and supported the theory that all terrestrial taxa (*Cercopithecus lhoesti*, *Erythrocebus patas*, *Chlorocebus aethiops*) form a monophyletic clade resulting in a paraphyly of the genus *Cercopithecus*. Furthermore the study documented the basal position of *Allenopithecus* within this tribe. Questionable is, whether further SINE analyses would be the right way to resolve the few remaining questions in Cercopithecinae, since most splitting events are assumed to be radiation-like or too recent (Schmitz et al., 2005). As mentioned in Chapter 1.1 the phylogenetic relationships within Hominidae were additionally resolved by a study based on mobile elements as well (Salem et al., 2003). However, there is no mobile element based study concerning relationships among gibbons at hand, because recent molecular and chromosomal analyses indicate that gibbons could be divided into four different genera (Roos and Geissmann, 2001).

Although many issues were addressed and solved by several SINE based studies, crucial questions in primate phylogeny remain. This is due to the fact that nearly all studies lack essential genera. Hence, the first part of this thesis addresses the remaining questions in New World monkey phylogeny and is presented and discussed in Chapter 2. In the frame of this study a SINE insertion distinguishing *Saimiri boliviensis boliviensis* from *Saimiri sciureus* was detected, two species, which are both common in zoos and biomedical research. The relevance of this finding is presented and discussed in Chapter 3. As outlined above there are still crucial disagreements in the evolutionary history of colobine monkeys. Chapter 4 and 5 are dedicated to these fundamental issues.

2 Retropositional events consolidate the branching order among New World monkey genera

Martin Osterholz[1*], Lutz Walter[1,2], and Christian Roos[1,2]

[1]Primate Genetics and [2]Gene Bank of Primates, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany

*Corresponding author

submitted to Molecular Phylogenetics and Evolution

Abstract

Due to contradicting relationships as obtained from various morphological and genetic studies, phylogenetic relationships among New World monkey genera are highly disputed. In the present study, we analyzed the presence/absence pattern of 128 SINE integrations in all New World monkey genera. Among them, 70 are specific for only a single genus, whereas another 18 are present in all New World monkey genera. The 40 remaining insertions are informative to elucidate phylogenetic relationships among genera. Several of them confirm the monophyly of the three families Cebidae, Atelidae and Pitheciidae as well as of the subfamily Callithrichinae. Further markers provide evidence for a sister grouping of Cebidae and Atelidae to the exclusion of Pitheciidae as well as for relationships among genera belonging to Callithrichinae and Atelidae. Although a close affiliation of *Saimiri*, *Aotus* and *Cebus* to Callithrichinae is shown, the relationships among these three genera are not resolved due to contradicting relationships as obtained from three independent insertions.

2.1 Introduction

New World monkeys, infraorder Platyrrhini, diverged from Old World monkeys about 40 million years ago (Goodman et al., 1998), and are distributed from northern Argentina to Mexico. Their common origin is widely accepted and confirmed by morphological features like dental, cranial or placental attributes as well as by molecular data (Hofer, 1976; Rosenberger, 1977; Luckett, 1980; Schneider et al., 1993; Horovitz and Meyer, 1995; Porter et al., 1995; Porter et al., 1997; Horovitz et al., 1998; Singer et al., 2003; Ray et al., 2005). New World monkeys represent a highly diverse group with an extensive morphological and ecological adaptive diversity, manifested in the presence of up to thirteen sympatric species (Fleagle, 1988). Accordingly, reconstructing phylogenetic relationships among platyrrhine genera is difficult.

Most recent classifications arrange New World monkeys into three families (Cebidae, Atelidae and Pitheciidae) with 14 to 15 genera (Goodman et al., 1998; Opazo et al., 2006), but alternative classifications with only two (Napier and Napier, 1967) or even four families (Groves, 2001) were also proposed. The branching order among the three families is still disputed and even recent molecular analyses were not able to confidently resolve it, mainly due to incongruences between different marker systems (Schneider et al., 1993; Harada et al., 1995; Schneider et al., 1996; Goodman et al., 1998; Canavez et al., 1999; Steiper and Ruvolo, 2003; Ray et al., 2005; Opazo et al., 2006). Questioned are also the phylogenetic relationships of genera within the Cebidae subfamily Callithrichinae and the familiy Atelidae as well as the positions of the genera *Saimiri*, *Cebus*, *Aotus* and *Callicebus*.

Within the Callithrichinae, most morphological studies agree in placing *Callimico* as the basal genus, whereas they differ in placing *Saguinus* either as a sister genus to *Callithrix,* or to a *Callithrix* + *Leontopithecus* clade (Rosenberger, 1981; Ford, 1986; Kay, 1990). However, molecular studies provide evidence for a close affiliation of *Callithrix* and *Callimico* and a basal position of *Saguinus* (Pastorini et al., 1998; Canavez et al., 1999; Chaves et al., 1999; von Dornum and Ruvolo, 1999; Singer et al., 2003). Other molecular studies suggest two clades with one comprising *Callithrix* and *Callimico* and the other

Leontopithecus and *Saguinus* (Schneider et al., 1993; Harada et al., 1995; Porter et al., 1997; Goodman et al., 1998).

Within Atelidae, various relationships among its four genera were proposed. According to Rosenberger (1981), *Ateles* and *Brachyteles* form a clade with *Lagothrix* as sister and *Alouatta* as basal genus. Kay (1990) defined two monophyletic clades (*Alouatta* + *Brachyteles* and *Lagothrix* + *Ateles*), while Ford (1986) proposed an unresolved trichotomy between *Lagothrix*, *Ateles* and *Brachyteles*, with *Alouatta* representing their sister genus. Most molecular studies agree in a close relationship between *Lagothrix* and *Brachyteles*, followed by *Ateles* and *Alouatta* (Schneider et al., 1993; Harada et al., 1995; Porter et al., 1997; Goodman et al., 1998; Horovitz et al., 1998; Canavez et al., 1999; von Dornum and Ruvolo, 1999), however, an unresolved relationship between *Brachyteles*, *Lagothrix* and *Ateles* is also proposed (Collins, 2004).

Highly disputed are the phylogenetic positions of the genera *Saimiri*, *Cebus*, *Aotus* and *Callicebus* among New World monkeys. Rosenberger (1981) and Ford (1986) proposed a sister grouping of *Callicebus* and *Aotus* as well as of *Saimiri* and *Cebus*, but they differ in positioning them among platyrrhines. *Saimiri* + *Cebus* is either the most basal clade (Ford, 1986) or forms a sister clade to Callithrichinae (Rosenberger, 1981). *Aotus* + *Callicebus* is either a sister clade to Callithrichinae, Atelidae and Pitheciidae (Ford, 1986) or solely to the latter (Rosenberger, 1981). According to Kay (1990), *Callicebus* is the most basal genus among platyrrhines followed by *Cebus*, whereas *Saimiri* and *Aotus* form a sister clade to Callithrichinae. Molecular studies support a grouping of *Callicebus* with Pitheciidae and of *Cebus, Saimiri* and *Aotus* with Callithrichinae in the family Cebidae (Harada et al., 1995; Porter et al., 1997; Horovitz et al., 1998; Canavez et al., 1999; von Dornum and Ruvolo, 1999; Steiper and Ruvolo, 2003). In all these studies, a *Cebus* + *Saimiri* clade is depicted, while the position of *Aotus* remains disputed. Some suggest an unresolved position of *Aotus* in a Cebidae polytomy (Schneider et al., 1993; von Dornum and Ruvolo, 1999) or a basal position among Cebidae (Horovitz et al., 1998). Alternatively, *Aotus* is placed as as sister taxon to *Cebus* + *Saimiri* (Prychitko et al., 2005; Opazo et al., 2006).

As indicated above, the phylogenetic relationships among New World monkey families and genera are far away from being resolved. Since morphological and molecular studies seem to be not appropriate to tackle all these issues, Alu integrations were applied as cladistic markers (Singer et al., 2003; Ray et al., 2005). Alus as primate-specific SINE (Short INterspersed Elements) elements are a class of retrotransposons integrating via an RNA intermediate into the genome (Okada, 1991). The integration of a SINE at a new locus is irreversible and precise excision is highly unlikely (Shedlock and Okada, 2000; van de Lagemaat et al., 2005). Therefore, absence of an insertion at a certain locus can supposed to be the ancestral state and due to its fixation in the genome, a common ancestor can be determined by a shared derived insertion. Furthermore, a true orthology can be verified and homoplasy can be excluded by analysing their flanking direct repeats (Schmitz et al., 2005). Consequently, SINE insertions represent a powerful molecular cladistic tool for reconstructing

phylogenetic relationships (Shimamura et al., 1997; Takahashi et al., 2001; Roos et al., 2004; Ray et al., 2005; Xing et al., 2005; Kriegs et al., 2006; Osterholz et al., 2008).

In the present study, we analyzed the presence/absence pattern of SINE insertions in all platyrrhine genera. To detect new loci, a computational as well as a subtractive hybridization approach was applied. Moreover, we tested genera, which have not been analyzed in an earlier study (Ray et al., 2005), but are crucial to depict the exact location of respective SINE insertions in the platyrrhine tree.

2.2 Methods

2.2.1 Computational approach

Sequences from Bacterial Artificial Chromosome (BAC) clones from five New World monkeys (*Callithrix jacchus*, *Saimiri boliviensis*, *Aotus nancymaae*, *Ateles geoffroyi* and *Callicebus moloch*) were obtained from GenBank database (Supplementary table 2.1). To identify SINEs of interest, we scanned the GenBank database for orthologous BAC clones from different genera. These were aligned using MAFFT software (Katoh et al., 2005) and further scanned for SINEs using RepeatMasker (www.girinst.org) (Jurka et al., 2005). For further analyses, only insertion loci specific for at least one New World monkey species were selected. The orthology of lineage specific insertions were confirmed by checking direct repeats. At least 400 bp of flanking sequences were then compared with the genomes of human, chimpanzee and rhesus macaque using BLAT search of the UCSC Genome Browser (Kent et al., 2002). Based on this information, locus-specific oligonucleotide primers in conserved flanking regions of the insertion were designed for further PCR analyses.

2.2.2 Subtractive hybridization approach

As second approach, a subtractive hybridization was applied, following methods as described in Osterholz et al. (2008), Mamedov et al. (2005) and Diatchenko et al. (1996). We conducted two subtractive hybridizations. In the first, *Lagothrix* was selected as tracer and *Saimiri* as driver, whereas in the second, *Pithecia* and *Chiropotes* were applied as tracer and *Ateles* as driver. Genomic DNA of tracer and driver was digested with RSAI (Fermentas), and subsequently, the adapters AdapA1/AdapAA1 (5'- TGTAGCGTGAAGACGACAGAAAGGGCGTGGTGCGGAGGGCGGT-3' / 5'- ACCGCCCTCCG-3') and AdapA2/AdapAA2 (5'- TGTAGCGTGAAGACCTGTCTTAGGGCGTGGTGGCCAGGGCCGT-3' / 5'- ACGGCCCTGGC-3') were ligated to the tracer fragments. Each of ~15 ng tracerA1 and tracerA2 were hybridized with ~1,500 ng driver for 20 h at 60°C. 2 µl of the hybridization result was used as template to amplify solely tracer fragments using the adapter primers A1 (5'-TGTAGCGTGAAGACGACAGAA-3') and A2 (5'-TGTAGCGTGAAGACCTGTCTT-3'). The PCR program consisted of a pre-extension step at 72°C

for 6 min to fill in adaptor ends, followed by 25 cycles, each with a denaturation step at 95°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 2 min. To enrich fragments with Alu insertions, a semi-nested PCR was added using either primer A1 or A2 and the Alu-specific AluS (5'-GGAGAATCGCTTGAACCCGGGA-3') oligonucleotide. The PCR products were separated on agarose gels and fragments over 500 bp were excised from the gel. After purification, the fragments were cloned into the pGEMTeasy vector (Promega) and transformed into electro-competent TOP10 cells (Invitrogen). Bacterial clones were collected in 96-well microtiter plates and re-screened via PCR with the primers A1 or A2 and AluS. Positive clones were sequenced and analyzed with RepeatMasker to identify repeat structures. To identify orthologous sequences in human, chimpanzee or rhesus macaque, sequences were subjected either to a BLAST (NCBI) or a BLAT (UCSC) search. Based on generated alignments, conserved locus-specific primers were designed, with the forward primer occupying a region 5'-end upstream of the insertion site. Due to the absence of the 3'-end downstream sequence in the tested new world monkey species, reverse primers were constructed solely on the basis of human, chimpanzee and rhesus macaque sequences.

2.2.3 Testing the presence/absence of insertions in New World monkeys

Samples from various New World monkeys were collected in European institutions: *Alouatta caraya, Cacajo calvus, Callimico goeldii, Callithrix pygmaea, Chiropotes albinasus x satanas, Leontopithecus chrysomelas* (Cologne zoo), *Aotus azarae* (Gettorf zoo), *Saguinus imperator, Samiri sciureus* (German Primate Center), *Cebus apella* (Romagne zoo), *Lagothrix lagothricha* (Basel zoo), *Ateles fusciceps* (Landau zoo) and *Pithecia pithecia* (Dortmund zoo). Total genomic DNA from blood or tissue material was extracted with the Qiagen DNeasy Blood & Tissue Kit. DNA from *Brachyteles arachnoides* was provided by Stefan Müller (University of Munich).

To examine the presence/absence pattern of SINE insertions, we first tested only one representative of each family using locus-specific primers as listed in supplementary table 2.2. Based on this information, further relevant genera were examined. PCR conditions for all reactions were identical and included an initial denaturation step at 94°C for 2 min., followed by 40 cycles each with of one minute of denaturation at 94°C, annealing at 58°C and extension at 72°C. At the end, a final extension step at 72°C for 5 min. was added. Besides newly detected integration loci, we also tested 29 insertions described by Ray et al. (2005) (Supplementary table 2.2), which have not been tested in all relevant genera. PCR conditions for these amplifications were conducted as described in Ray et al. (2005). Results of the amplifications were checked on 1% agarose gels and further processed with the Wizard gel purification kit (Promega). To verify the orthology of loci and insertions, PCR products were sequenced on an ABI3100-Avant sequencer using the Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems). Sequences were edited using BioEdit software (Hall, 1999) and aligned manually. Sequences generated in the present study are available in Genbank (Accession nos. xxxxxx-xxxxxx).

2.3 Results

In total, 128 integrations were studied (Fig. 2.1, supplementary table 2.2). 21 of them were obtained from the subtractive hybridization approach, whereas 78 were detected by BAC clone screening. 29 integrations were derived from a previous study (Ray et al., 2005). With the exception of two Atelinae-specific insertions described by Ray et al. (2005), which provided ambiguous results for *Brachyteles*, all others showed a clear presence/absence pattern of integrations. 70 of them represent autapomorphies, which are only present in a single genus and another 18 insertions are present in all New World monkeys. Since these markers provide no information on phylogenetic relationships among New World monkey genera, they were examined only on PCR level without confirming the integration site by sequencing. Finally, 38 insertions were identified which can be applied to reconstruct phylogenetic relationships among platyrrhine genera.

The monophyly of the family Atelidae is confirmed by five insertions. Relationships within the family are not well resolved, although at least three insertions were detected which are present in *Lagothrix, Brachyteles* and *Ateles* and absent in *Alouatta*. A common origin of the four Pitheciidae genera *Pithecia, Cacajo, Chiropotes* and *Callicebus* is confirmed by four integrations, but no loci are available which resolve relationships among them. Seven integrations are present in all tested Cebidae representatives. To resolve relationships among Cebidae genera, several informative insertions are available. These support the monophyly of the subfamily Callithrichinae (five integrations) and a *Callithrix* + *Callimico* clade (seven insertions) as well as of a clade consisting of *Callithrix, Callimico* and *Leontopithecus* to the exclusion of *Saguinus* (three insertions). Contradicting relationships were obtained for the three Cebidae genera *Saimiri*, *Cebus* and *Aotus*. Whereas one insertion is present in *Saimiri* and *Aotus*, another two integrations support a *Saimiri* + *Cebus* and a *Cebus* + *Aotus* clade. Relationships among the three platyrrhine families are not well resolved by our study, but at least one insertion was detected which is present in Cebidae and Atelidae and absent in Pitheciidae.

2.4 Discussion

In general, all phylogenetic relationships proposed by Ray et al. (2005) and Singer et al. (2003) were confirmed in the present study, however, both previous studies did not include all genera and/or described only a few integrations. In contrast, we included all New World monkey genera and analyzed the presence/absence pattern of a relatively large number of SINE integrations.

Among 128 insertions examined, 18 are present in all New World monkey genera. In combination with the 84 loci detected by Ray et al. (2005) and three by Singer et al. (2003), now 105 independent SINE insertions provide evidence for a common origin of the infraorder.

The proposed diversification of New World monkeys into the three families Cebidae, Atelidae and Pitheciidae (Goodman et al., 1998; Opazo et al., 2006) is supported by several insertions. We detected four integrations on the branch leading to all Pitheciidae genera. Together with three

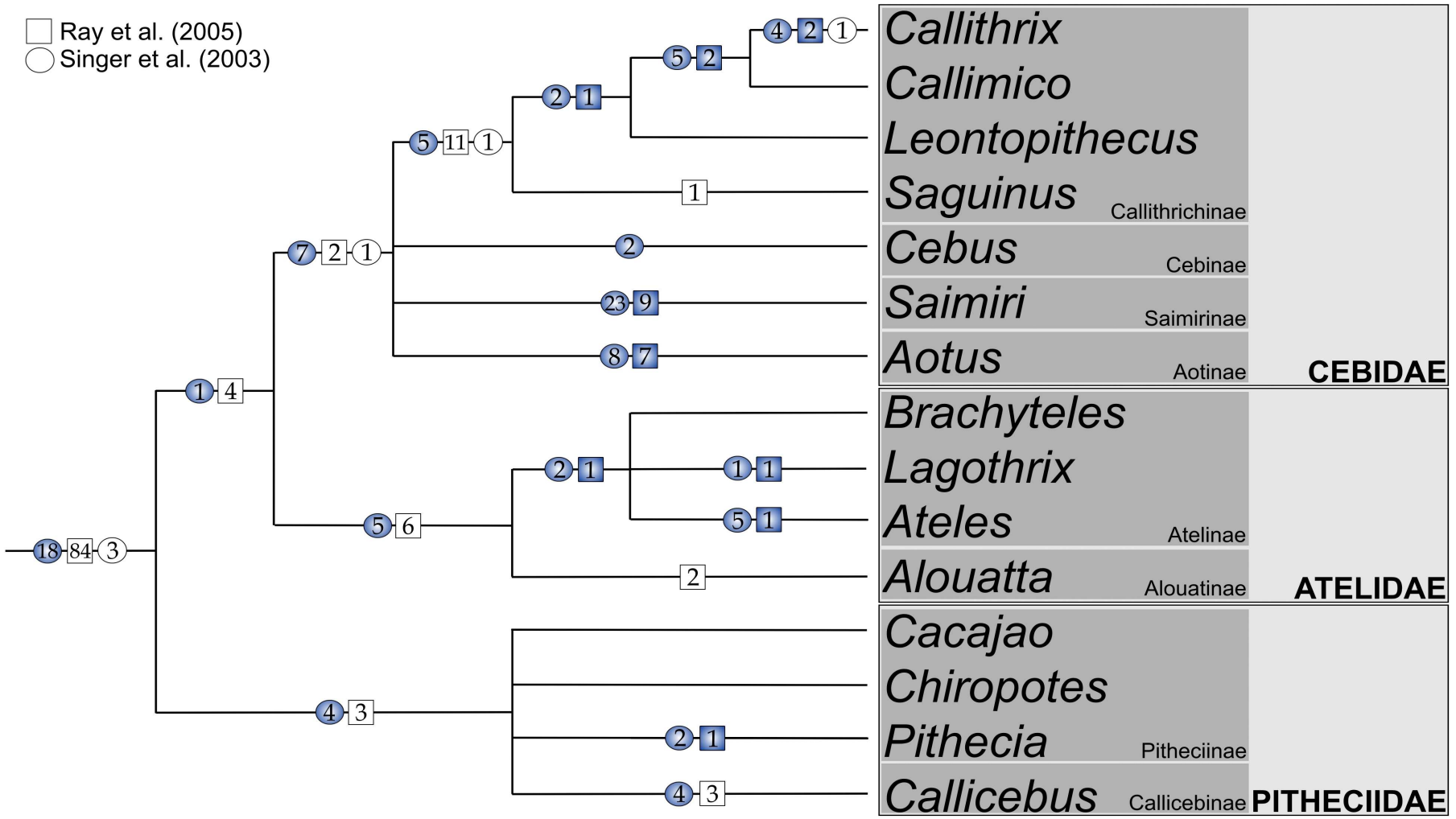

Figure 2.1 : Phylogenetic relationships among New World monkey genera as obtained from SINE insertion analyses. Numbers in filled circles indicate the number of retropositional events detected in the present study. Open squares and circles indicate SINE insertions described in previous studies (Ray et al., 2005; Singer et al., 2003) and filled squares indicate insertions published by Ray et al. (2005), but which have been tested in further genera in the present study. Ambiguous or contradicting insertions are not shown.

retropositional events described by Ray et al. (2005), a monophyletic Pitheciidae clade is confirmed. Five and seven integrations are present in all Atelidae and Cebidae genera, respectively. Combining them with earlier data (Ray et al. 2005; Singer et al. 2003), eleven and ten integration sites support the monophyly of both families.

The phylogenetic relationships among these three families are highly disputed. Morphological studies support a common origin of Atelidae and Pitheciidae (Rosenberger, 1981, 1984; Ford, 1986). Alternatively, Cebidae group with Atelidae to the exclusion of Pitheciidae (Kay, 1990). Molecular analyses based on mitochondrial sequence data (Horovitz and Meyer, 1995) or nuclear data (Steiper and Ruvolo, 2003) were not able to solve the relationships among them, although a grouping of Atelidae and Cebidae is assumed (Opazo et al., 2006). In the present study, we detected one insertion, which indicates an Atelidae + Cebidae clade. Together with four insertions provided by Ray et al. (2005), there is now clear evidence that Pitheciidae form the basal family among New World monkeys and that Atelidae and Cebidae represent sister lineages.

As mentioned above, the monophyletic origin of the Cebidae family including Callithrichinae and the genera *Cebus*, *Saimiri* and *Aotus* is confirmed by ten integrations, which is concordant with other molecular studies (Goodman et al., 1998; Singer et al., 2003; Steiper and Ruvolo, 2003; Ray et al., 2005; Opazo et al., 2006). The common origin of Callithrichinae is also supported by morphological investigations (Rosenberger, 1981, 1984; Ford, 1986; Kay, 1990), but they differ in positioning *Cebus*, *Saimiri* and *Aotus* among New World monkeys. Although all molecular studies clearly confirm a grouping of these genera together with Callithrichinae, the relationships among them remain unresolved. We detected three independent insertions concerning the phylogenetic relationships of *Saimiri*, *Aotus* and *Cebus*. The first one supports a sister grouping of *Aotus* and *Cebus* in a polytomy with *Saimiri* and Callithrichinae. However, the second one indicates a close affiliation of *Aotus* to *Saimiri* in a polytomy with *Cebus* and Callithrichinae and the third one shows the third possible combinations with *Saimiri* and *Cebus* as sister lineages. These contradicting results might be explained by incomplete lineage sorting of ancestral polymorphisms, which are common when species diverge from each other within a relative short time period. Support for such a radiation-like diversification of the four lineages is gained from the small number of observed informative and the large number of autapomorphic markers. Further support for an adaptive radiation is also provided by the relative short time period separating these four lineages as estimated from molecular sequence data (Opazo et al., 2006; Schrago, 2007).

In contrast, the monophyly and relationships within Callithrichinae are strongly supported by several integrations. We detected five integrations shared by all Callithrichinae. Together with twelve earlier described insertions (Ray et al., 2005; Singer et al., 2003), a common origin of the Callithrichinae subfamily is strongly supported. Both molecular and morphological studies agree in this respect, but the branching order within the subfamily is disputed - mainly whether *Saguinus*, *Leontopithecus* or *Callimico* form the basal genus (Rosenberger, 1981; Ford, 1986; Goodman et al., 1998; Pastorini et al., 1998; Schneider et al., 2001; Singer et al., 2003). Especially, the position of *Callimico* was challenging.

Although most studies placed the genus as basal among Callithrichinae (Napier and Napier, 1967; Szalay and Delson, 1979), others included it either in Cebidae (Martin, 1990) or in its own family Callimiconidae (Hershkovitz, 1977). Molecular studies support a classification among Callithrichinae and specifically as sister lineage to *Callithrix* (Harada et al., 1995; Horovitz and Meyer, 1995; Schneider et al., 1996; Porter et al., 1997; Pastorini et al., 1998; Canavez et al., 1999; Chaves et al., 1999; von Dornum and Ruvolo, 1999; Singer et al., 2003). In the present study, we obtained seven insertions shared by *Callimico* and *Callithrix*, confirming their sister group relationship. Three further markers unite *Callithrix*, *Callimico* and *Leontopithecus,* therefore suggesting that *Saguinus* forms the basal form among all Callithrichinae. This arrangement is in agreement with several molecular studies (Schneider et al., 1996; Chaves et al., 1999; Singer et al., 2003; Opazo et al., 2006), however, contradicts with the proposed sister grouping of *Saguinus* and *Leontopithecus* (Schneider et al., 1993; Harada et al., 1995; Porter et al., 1997; Goodman et al., 1998).

The monophyly of Atelidae is supported by five retropositional events. With the insertions detected by Ray et al. (2005), a common origin of the four Atelidae genera *Alouatta*, *Ateles*, *Lagothrix* and *Brachyteles* is now confirmed by eleven integrations. Combining morphological and molecular studies, this arrangement seems to be evident (Rosenberger, 1981; Ford, 1986; Kay, 1990; Porter et al., 1995; von Dornum and Ruvolo, 1999; Opazo et al., 2006; de Lima et al., 2007). Moreover, three markers are present in *Ateles*, *Lagothrix* and *Brachyteles*, but absent in *Alouatta*, supporting other molecular (Schneider et al., 1993; Harada et al., 1995; Porter et al., 1997; Goodman et al., 1998; Horovitz et al., 1998; Canavez et al., 1999; von Dornum and Ruvolo, 1999) and some morphological investigations (Rosenberger, 1981; Ford, 1986). Accordingly, the splitting of the family into the two subfamilies Alouatinae with the single genus *Alouatta* and Atelinae including the remaining three genera is appropriate. The relationships among the three Atelinae genera are not resolved, due to the lack of informative markers.

Four insertions support the monophyly of Pitheciidae, including *Callicebus*. The position of this unique genus was highly disputed, mainly due to morphological autapomorphies (Rosenberger, 1981; Ford, 1986; Kay, 1990), but molecular data clearly supported its close affiliations to *Pithecia*, *Chiropotes* and *Cacajo* (Goodman et al., 1998; Steiper and Ruvolo, 2003; Opazo et al., 2006). Based on our four integrations and three already described ones (Ray et al., 2005), the unity of the family is confirmed. However, no further markers, which elucidate the branching order among the four Pitheciidae genera were detected.

2.5 Conclusions

By analyzing the presence/absence pattern of 128 SINE insertions, the present study deepens our knowledge on the phylogenetic relationships among New World monkey genera. Based on our and other SINE data, a common origin of the three families Cebidae, Atelidae and Pitheciidae can be regarded as settled as well as a sister grouping of Cebidae and Atelidae to the exclusion of Pitheciidae.

Although the monophyly of Cebidae and its subfamily Callithrichinae is confirmed, the phylogenetic positions of *Saimiri*, *Aotus* and *Cebus* in the Cebidae clade remain unresolved. Due to the obtained contradicting relationships, which might be affected by incomplete lineage sorting, further efforts have to be undertaken to solve this issue. Our results further provide evidence for the branching order among Callithrichinae genera as well as for a basal position of *Alouatta* among Atelidae. As indicated by this and earlier studies, retropositional events provide a powerful molecular cladistic tool to reconstruct reliable phylogenies, which form a solid platform for various behavioral, morphological and molecular investigations.

Acknowledgements

We are grateful to the staff of the zoos in Cologne, Dresden, Gettorf, Basel, Romagne, Landau and Dortmund and the German Primate Center (DPZ) for samples, as well as to Stefan Müller for providing DNA from *Brachyteles arachnoides*. Finally, we thank Christiane Schwarz and Nico Westphal for their excellent technical assistance in the laboratory. The work was financially supported by the DPZ.

3 A PCR-based marker to simply identify *Saimiri sciureus* and *S. boliviensis boliviensis*

Martin Osterholz[1], Jan Vermeer[2], Lutz Walter[1,3], and Christian Roos*[1,3]

[1]Primate Genetics, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany

[2]La Vallée des Singes, Le Gureau, 86700 Romagne, France

[3]Gene Bank of Primates, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany

*Corresponding author

American Journal of Primatology 2008;70: 1-4.

Abstract

Squirrel monkeys, mainly *Saimiri sciureus* and *S. boliviensis*, are common in zoos and widely used in biomedical research. However, an exact species identification based on morphological characteristics is difficult. Hence, several molecular methods were proposed, but all of them are expensive and require extensive laboratory work. In contrast, we describe an Alu integration, which is present in *S. boliviensis boliviensis* and absent in *S. sciureus*. Among analysed *S. b. peruviensis* specimens various presence / absence patterns of the integration were detected indicating that this study population might have originated from a natural hybrid zone. Based on the size of the Alu element (~300 bp), the presence / absence pattern of the integration can easily be traced by PCR and followed by agarose gel electrophoresis.

3.1 Introduction

Squirrel monkeys (genus *Saimiri*) are New World monkeys and belong to the family Cebidae. The genus is distributed, apart from some isolated populations in Costa Rica and Panama, in tropical forests of northern South America to Bolivia and central Brazil (Hershkovitz, 1984; Rowe, 1996; Groves, 2001). The taxonomic classification of squirrel monkeys changed several times with one to seven species and seven to sixteen subspecies (Elliot, 1913; Lönnerg, 1940; von Pusch, 1942; Hill, 1960; Hershkovitz, 1984; Thorington, 1985; Costello et al., 1993; Rowe, 1996; Groves, 2001). Based on morphological, acoustic, chromosomal, molecular and behavioral differences, most authors agree in distinguishing two different species groups, *S. boliviensis* including *S. vanzolinii*, and *S. sciureus* with *S. ustus* and *S. oerstedii* (Hershkovitz, 1984; Boinski and Newman, 1988; Moore et al., 1990; Rowe, 1996; Boinski and Cropp, 1998; Schreiber et al., 1998; Groves, 2001).

Squirrel monkeys, mainly *S. sciureus* and *S. boliviensis*, are common in zoos and widely used in biomedical research (Mittermeier et al., 1994; Vermeer, 1996). Accordingly, many breeding colonies exist, and in order to minimize inbreeding, animals are regularly exchanged among them. However, because the exact origin of founder animals is mainly unknown, the (sub)species identity of such animals is difficult to assess with morphological characteristics (Vermeer, 1996). As a result, several hybrids were produced between species or subspecies in recent decades (Vermeer, 1996; Schreiber et al., 1998). Although both *S. sciureus* and *S. boliviensis* are currently classified as only "least concern" (IUCN, 2007), pure breeding is of interest for conservation issues. In biomedical research, pure breed lineages are also important, because species differ in critical biological parameters as e.g. susceptibility to diseases (VandeBerg et al., 1990; VandeBerg and Williams-Blangero, 1997; Lavergne et al., 2003). However, since an accurate identification of species and hybrids based on external characteristics is difficult, methods not reliant on phenotype (e.g. molecular markers) should be applied. Some molecular tests based on allozyme and microsatellite polymorphisms or sequencing of marker genes have been proposed (VandeBerg et al., 1990; Silva et al., 1993; Boinski and Cropp, 1998; Schreiber et al., 1998; Cropp and Boinski, 2000; Lavergne et al., 2003), but all of them are expensive and time demanding.

Short INterspersed Elements (SINE) represent a class of retrotransposons integrating via an RNA intermediate into the genome (Okada, 1991). The integration of a SINE at a new locus is irreversible and precise excision is highly unlikely (Shedlock and Okada, 2000; van de Lagemaat et al., 2005). Orthology can be verified and homoplasy can be excluded by tracing direct repeats which flank the integration (Schmitz et al., 2005). Alu elements are primate-specific SINEs and have been widely propagated in their genomes.

In this study, we tested the presence / absence pattern of an Alu insertion in 93 squirrel monkeys.

3.2 Material & Methods

3.2.1 Database approach

Bacterial Artificial Chromosome (BAC) clones from various New World monkey species (*S. b. boliviensis, Callithrix jacchus*, *Aotus nancymaae*, *Ateles geoffroyi* and *Callicebus moloch*) were obtained from GenBank database. Using BLAT search, orthologous BAC clones were identified and aligned with MAFFT software (Katoh et al., 2005). Subsequently, Alu integrations were identified with RepeatMasker (Jurka et al., 2005). At least 400 bp sequence information from both sides of the integration site was selected and subjected to BLAT search to find orthologous loci in the genomes of *Homo sapiens, Pan troglodytes* and *Macaca mulatta*. Based on this information, conserved oligonucleotide primers were constructed, which bind in the flanking regions of the Alu insertion. In the frame of this study, one AluTa15 integration (*SscSbo*) was detected, which was present in *S. b. boliviensis*, but absent in other platyrrhines.

3.2.2 Laboratory methods

To further test the presence / absence pattern of the *SscSbo* integration, blood samples from 93 squirrel monkeys kept in European institutions were collected. The species identity of study specimens was determined by fur coloration and other external characteristics. Samples from phenotypically pure *S. sciureus* were provided by Dresden zoo (n=2), Gettorf zoo (n=5) and Schwerin zoo (n=1), phenotypically pure *S. b. boliviensis* from Mannheim zoo (n=2), Nuremberg zoo (n=3) and Romagne primate park (n=3) and *S. b. peruviensis* from Romagne primate park (n=53). Samples from animals, which were identified phenotypically as hybrids between *S. boliviensis* and *S. sciureus* were obtained from the German Primate Center (n=14) and from Madrid zoo (n=10). We have adhered to the guidelines for the use of animals in research and the legal requirements of Germany.

DNA from blood samples was extracted using the Qiagen DNA Mini kit. PCR amplifications were performed with the locus-specific oligonucleotide primers 5'-AGTTCCTCTCTACCTTGTACC-3'

and 5'-GCCCTACTCTTGCATTAATGC-3'. The expected fragment lengths of the PCR product are ~450 and ~750 bp in the case of presence and absence of the AluTa15 integration, respectively (Fig. 3.1a). PCR conditions were 94°C initial denaturation for 2 minutes, followed by 40 cycles each with 94°C denaturation for 1 minute, 58°C annealing for 1 minute and 72°C extension for 1 minute. The final extension step at 72°C was performed for 5 minutes. Results of PCR amplifications were analysed using a 1 % agarose gel. To confirm the orthology of the integration, PCR products from each one individual of *S. sciureus* and *S. b. boliviensis* were sequenced. Therefore, PCR products were excised from the gel and purified with the Wizard gel purification kit (Promega). Sequencing reactions were run on an ABI 3100-Avant sequencer using the Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems), and the primers mentioned above. To confirm the orthology of the integration, sequences were edited and manually aligned in BioEdit (Hall, 1999).

3.3 Results

In an alignment including BAC clones from *S. b. boliviensis* (AC188239), *C. jacchus* (AC188222), *A. geoffroyi* (AC188259) and *C. moloch* (AC188270), an AluTa15 (Ray and Batzer, 2005) insertion was detected, which is present in *S. b. boliviensis* and absent in the remaining three taxa. To further check the presence and absence of the integration, conserved primers binding in the flanking region of the insertion were constructed and tested in a panel of various New World monkeys. Interestingly, all tested species including *S. sciureus* showed an absence of the Alu insertion at that locus. Consequently, further squirrel monkey individuals were examined to test whether the integration is specific for *S. boliviensis*. Among the 93 squirrel monkeys examined, in 39 individuals the integration is present (+/+), whereas in 18 the integration is absent (-/-). In another 36 individuals, a heterozygous pattern (+/-) was detected, indicating that both alleles are present (Fig. 3.1a). All studied specimens which were phenotypically identified as *S. sciureus* showed a homozygous absence of the integration, while phentotypically *S. b. boliviensis* showed a homozygous presence of the insertion. The specimens identified as *S. b. peruviensis* showed an insertion pattern with all possible combinations (n=26: +/-, n=25: +/+, n=2: -/-). Similar results were obtained for 24 individuals, which were identified phenotypically as hybrids (n=10: +/-, n=5: +/+, n=9: -/-).

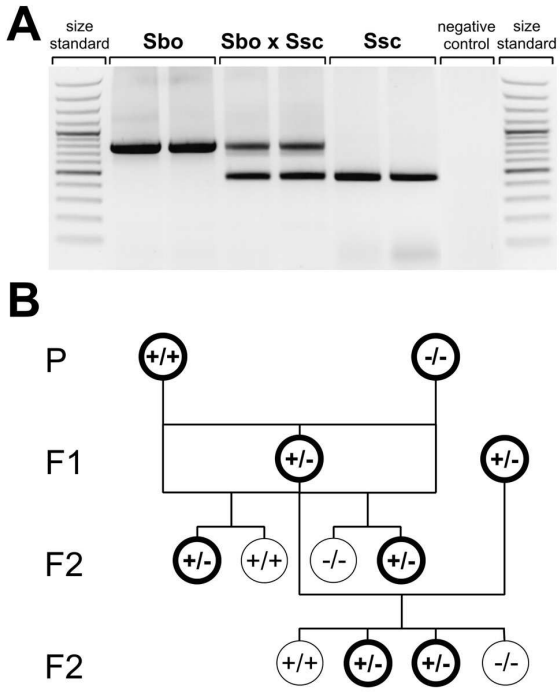

Figure 3.1: (A) Presence / absence analysis of the Alu integration as revealed by agarose gel electrophoresis. Pure *S. boliviensis* (Sbo) and *S. sciureus* (Ssc) show PCR product sizes of ~750 bp and ~450 bp, respectively. Hybrids (Sbo x Ssc) possess both alleles. (B) Inheritance scheme of alleles with (+) and without (-) Alu integration. F1 hybrids between pure *S. sciureus* (-/-) and *S. boliviensis* (+/+) show a heterozygous (+/-) pattern, whereas in the F2 generation heterozygous (+/-) as well as homozygous (+/+; -/-) individuals are possible. Accordingly, the latter are falsely classified as pure breed animals (indicated by thin lines).

3.4 Discussion

The results of this study indicate that the analysed Alu insertion is specific for *S. b. boliviensis*, and hence, can be used to distinguish *S. b. boliviensis* from *S. sciureus*. However, contradicting results were obtained for *S. b. peruviensis*, since we found homozygous positive, homozygous negative and heterozygous insertion patterns. It is possible that our study specimens of *S. b. peruviensis* originated from a natural hybrid zone between *S. sciureus* and *S. b. peruviensis*, which has been reported from the margins of the Ucayali river in the Peruvian Amazonia (Silva et al., 1992). Moreover it cannot be excluded that *S. b. peruviensis* is the result of ancestral hybridization at all, as indicated by the fact that they are phenotypically intermediate between *S. sciureus*, and *S. b. boliviensis*, showing male head coloration as in the former and female head coloration as in the latter. However, recent molecular studies indicate a close affiliation of *S. b. peruviensis* and *S. b. boliviensis* (Boinski and Cropp, 1998; Cropp and Boinski, 2000). Therefore and due to the partial presence of the integration it seems likely that *S. b. peruviensis* from outside the hybrid zone may be homozygous positive.

Among the 24 phenotypically identified hybrids, ten individuals showed a heterozygous pattern indicating indeed that these animals are hybrids between *S. b. boliviensis* and *S. sciureus*. The other 14 animals show either homozygous presence or absence of the integration. With the herein presented marker, all F1 hybrids can be clearly defined, whereas in F2 hybrids only 50 %, those with

heterozygous pattern, are traceable. F2 hybrids with either homozygous presence or absence patterns would be falsely classified as either pure *S. b. boliviensis* or *S. sciureus* (Fig. 3.1b)

Although with some drawbacks, we identified a potential molecular cladistic marker to distinguish between *S. sciureus* and *S. b. boliviensis*. The advantage of this marker is that the size of PCR products with and without integration differs by ~300 bp, so that only agarose and no acrylamide gels or sequencing analyses are necessary. Accordingly, results can easily be determined and laboratory costs are relative low. However, since no other species than S. *boliviensis* and *S. sciureus* were analysed, it remains open which presence / absence pattern the other species will show. Another drawback of the marker is that only F1 hybrids and not those in further generations are traceable with significance. This will be overcome if more such markers will become available. Nevertheless, the presented marker provides a useful tool to easily distinguish pure breed *S. sciureus* and *S. b. boliviensis*, which may help improve captive breeding management.

Acknowledgements

We thank the staff of the zoos in Dresden, Gettorf, Madrid, Mannheim, Nuremberg, Romagne, Schwerin and the German Primate Center for providing samples of squirrel monkeys. We have adhered to the guidelines for the use of animals in research and the legal requirements of Germany.

4 Phylogenetic position of the langur genera *Semnopithecus* and *Trachypithecus* among Asian colobines, and affiliations of species groups

Martin Osterholz[1], Lutz Walter[1,2], and Christian Roos*[1,2,3]

[1]Primate Genetics and [2]Gene Bank of Primates, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany,

[3]IUCN/SSC Primate Specialist Group, Section Southeast Asia

*Corresponding author

BMC Evolutionary Biology 2008 Feb 25;8:58.

Abstract

Background: The evolutionary history of the Asian colobines is less understood. Although monophyly of the odd-nosed monkeys was recently confirmed, the relationships among the langur genera *Presbytis*, *Semnopithecus* and *Trachypithecus* and their position among Asian colobines remained unclear. Moreover, in *Trachypithecus* various species groups are recognized, but their affiliations are still disputed. To address these issues, mitochondrial and Y chromosomal sequence data were phylogenetically related and combined with presence / absence analyses of retroposon integrations.

Results: The analysed 5 kb fragment of the mitochondrial genome allows no resolution of the phylogenetic relationships among langur genera, but five retroposon integrations were detected which link *Trachypithecus* and *Semnopithecus*. According to Y chromosomal data and a 573 bp fragment of the mitochondrial cytochrome b gene, a common origin of the species groups *T.* [*cristatus*], *T.* [*obscurus*] and *T.* [*francoisi*] and their reciprocal monophyly is supported, which is also underpinned by an orthologous retroposon insertion. *T.* [*vetulus*] clusters within *Semnopithecus*, which is confirmed by two retroposon integrations. Moreover, this species group is paraphyletic, with *T. vetulus* forming a clade with the Sri Lankan, and *T. johnii* with the South Indian form of *S. entellus*. Incongruence between gene trees was detected for *T.* [*pileatus*], in that Y chromosomal data link it with *T.* [*cristatus*], *T.* [*obscurus*] and *T.* [*francoisi*], whereas mitochondrial data affiliates it with the *Semnopithecus* clade.

Conclusions: Neither relationships among the three langur genera nor their position within Asian colobines can be settled with 5kb mitochondrial sequence data, but retroposon integrations confirm at least a common origin of *Semnopithecus* and *Trachypithecus*. According to Y chromosomal and 573 bp mitochondrial sequence data, *T.* [*cristatus*], *T.* [*obscurus*] and *T.* [*francoisi*] represent true members of the genus *Trachypithecus*, whereas *T.* [*vetulus*] clusters within *Semnopithecus*. Due to paraphyly of *T.* [*vetulus*] and polyphyly of *Semnopithecus*, a split of the genus into three species groups (*S. entellus* - North India, *S. entellus* - South India + *T. johnii*, *S. entellus* - Sri Lanka + *T. vetulus*) seems to be appropriate. T. [*pileatus*] posses an intermediate position between both genera, indicating that the species group might be the result of ancestral hybridization.

4.1 Background

The Old World monkeys are traditionally divided into the two subfamilies Cercopithecinae and Colobinae, which differ from each other by numerous morphological, behavioral and ecological characteristics (Napier, 1970; Szalay and Delson, 1979; Davies and Oates, 1994; Groves, 2001). While detailed information on the evolutionary history of cercopithecines (baboons, mangabeys, macaques and guenons) is at hand, knowledge on colobines is still scarce. Although some molecular studies (Rosenblum et al., 1997; Wang et al., 1997; Zhang and Ryder, 1998; Roos, 2003; Geissmann et al., 2004; Roos, 2004; Nadler et al., 2005; Whittaker et al., 2006; Roos et al., 2007) exist, they mainly focus on relationships within genera or species groups and not on the general phylogeny of the subfamily. Recently, a first mitochondrial phylogeny of colobine genera was established (Sterner et al., 2006), which confirmed some previous assumptions, but also led to confusions, calling for further research to definitively elucidate their evolutionary history.

Based on distribution and morphology, the colobines are traditionally divided into an African and an Asian clade (Davies and Oates, 1994; Groves, 2001), while Asian colobines are more diverse than their African relatives. Hence, the Asian forms are further split into the odd-nosed monkey (*Pygathrix, Rhinopithecus, Nasalis, Simias*) and langur (*Presbytis, Trachypithecus, Semnopithecus*) groups,

which are both believed to be monophyletic. Accordingly, langurs were originally combined in the single genus *Presbytis* (Napier and Napier, 1967; Groves, 1970; Delson, 1975) or *Semnopithecus* (Reichenbach, 1862), but based on neonatal coloration and cranial morphology, they were split into the three genera *Semnopithecus, Trachypithecus* and *Presbytis* (Pocock, 1935), and a fourth genus (*Kasi*) was added (Hill, 1934). Alternatively, *Semnopithecus* was separated from *Presbytis*, with *Trachypithecus* forming a subgenus of the former (Brandon-Jones, 1984; Strasser and Delson, 1987; Brandon-Jones, 1995), but recent classifications use a subdivision of langurs into the three genera *Presbytis, Trachypithecus* and *Semnopithecus* (Weitzel et al., 1988; Groves, 1989; Davies and Oates, 1994; Groves, 2001; Brandon-Jones et al., 2004).

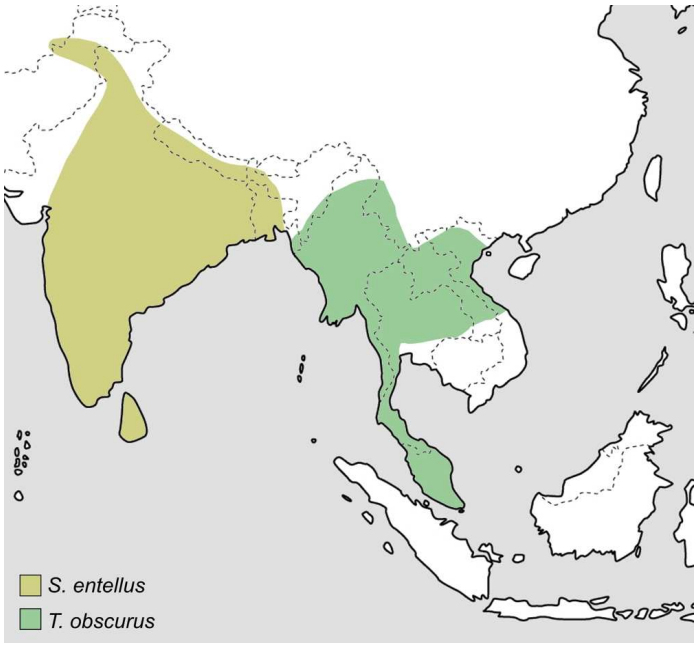

Figure 4.1: Distribution of the genus *Semnopithecus* and *Trachypithecus* [*obscurus*]. Genus affiliations and species groups after Groves (2001).

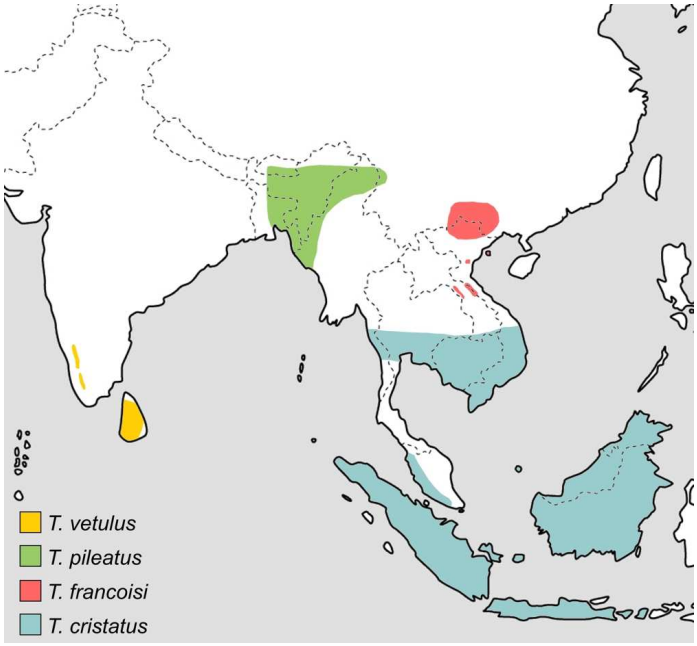

Figure 4.2: Distribution of the *Trachypithecus* species groups *T.* [*vetulus*], *T.* [*pileatus*], *T.* [*francoisi*] and *T.* [*cristatus*]. Genus affiliations and species groups after Groves (2001).

Within the different langur genera, several species are recognized, which are lumped into species groups due to similar fur coloration, behavior, ecology or distribution. With five species groups (*T.* [*obscurus*], *T.* [*francoisi*], *T.* [*cristatus*], *T.* [*pileatus*] and *T.* [*vetulus*]) (Groves, 2001), the genus *Trachypithecus* is the most diverse of all langurs and posses also the widest distribution, ranging from South India and Sri Lanka through mainland Southeast Asia to the Sundaland (Fig. 4.1, Fig. 4.2). Although all of them are morphologically similar, *T.* [*vetulus*] was sometimes separated in its own genus *Kasi* (Hill, 1934), and recent mitochondrial sequence data indicate a closer affiliation of *T.* [*vetulus*] and *T.* [*pileatus*] to *Semnopithecus* than to *Trachypithecus* (Zhang and Ryder, 1998; Geissmann et al., 2004). Accordingly, the two *T.* [*vetulus*] members were recognized as species of *Semnopithecus* (Brandon-Jones et al., 2004). In contrast to *Trachypithecus*, the genus *Semnopithecus* is restricted to the Indian subcontinent (Fig. 4.1) and traditionally regarded as monotypic with the only species *S. entellus* (Davies and Oates, 1994; Rowe, 1996), although recently several subspecies were elevated to species status (Groves, 2001). The third langur genus, *Presbytis,* includes several species, which occur solely in the Sundaland, but are not lumped into distinct species groups (Groves, 2001).

The phylogenetic relationships among the different Asian colobine genera are disputed. Although a common origin of the odd-nosed monkeys was recently confirmed (Sterner et al., 2006), evidence for monophyly of its putative sister clade, the langur group, is still lacking. Moreover, available data depict *Trachypithecus* and *Presbytis* as sister taxa to the exclusion of *Semnopithecus* (Sterner et al., 2006), which contradicts with traditional classifications, in which *Trachypithecus* and *Semnopithecus* are believed to form a clade to the exclusion of *Presbytis* (Brandon-Jones, 1984; Strasser and Delson, 1987; Brandon-Jones, 1995; Groves, 2001). These findings raise the question of what positions the langur genera occupy among Asian colobines and whether the langurs form a monophyletic clade in general. Moreover, the affiliations of different *Trachypithecus* species groups, especially *T.* [*vetulus*] and *T.* [*pileatus*], are disputed, and hence, led to different classifications. Currently, only few genetic data are available (Zhang and Ryder, 1998; Bigoni et al., 2003; Geissmann et al., 2004), so that further information from other markers is required to definitively establish their relationships.

To address all these issues, mitochondrial and Y chromosomal sequence data were phylogenetically related and combined with presence / absence analysis of retroposon integrations. This approach was used to simultaneously analyse paternal-, maternal- and biparental-inherited markers, which allow the detection of incongruence between different gene trees indicating possible hybridization or introgression events between different lineages (Evans et al., 2003; Tosi et al., 2002; Arnold and Meyer, 2006). To determine the phylogenetic position of the langur genera among Asian colobines, a 5 kb fragment of the mitochondrial genome was sequenced from eight colobine genera, and combined with presence / absence analysis of retroposon integrations. Retroposon insertion events are nearly homoplasy-free and precise excision of elements is highly unlikely (Shedlock and

Okada, 2000; Batzer and Deininger, 2002). Accordingly, retroposon insertions are powerful informative markers, which were already successfully applied to elucidate phylogenetic relationships in various primate lineages (Schmitz et al., 2001; Roos et al., 2004; Ray et al., 2005; Schmitz et al., 2005; Xing et al., 2005). To study relationships among different langur species groups and their genus affiliations, a 573 bp fragment of the mitochondrial cytochrome b gene and a 777 bp portion of the SRY (sex-determining region, Y chromosome) gene was sequenced from at least one representative per species group, and complemented with retroposon analysis.

4.2 Results and Discussion

4.2.1 Genus level phylogeny

To elucidate the phylogenetic relationships among the different langur genera and their position among Asian colobines, mitochondrial sequence studies were combined with presence / absence analysis of retroposon insertions.

The herein analysed 5 kb fragment of the mitochondrial genome was assembled from sequences derived from 1-2 kb long and partly overlapping PCR products, whereby no inconsistencies in overlapping sequence fragments were detected. As template material, mainly DNA extracted from feces was used, which minimizes the amplification of nuclear pseudogenes (Thalmann et al., 2004), and comparisons of the data with sequences already deposited at GenBank revealed only intra-species or -generic variation, indicating that no nuclear pseudogenes were amplified.

To determine the phylogenetic relationships among analysed genera, tree reconstructions were conducted with different algorithms, which all led to the same tree topology (Fig. 4.3). Moreover, most relationships are significantly supported and congruent with previous classifications (Napier, 1970; Szalay and Delson, 1979; Davies and Oates, 1994; Groves, 2001), indicating the reliability of the data set. In detail, the reconstructions confirm relationships among cercopithecine genera (*Macaca, Papio, Chlorocebus*), and reciprocal monophyly of cercopithecines and colobines, as well as of African and Asian colobines. Within the Asian clade, relationships are not resolved, although at least a common origin of the odd-nosed monkeys is depicted. Neither the assumed monophyly of the langur genera, nor the expected close affinity of *Trachypithecus* and *Semnopithecus* or the recently indicated sister grouping of *Trachypithecus* and *Presbytis* (Sterner et al., 2006) can be verified with significance. Moreover, all alternative tree topologies, in which the three langur genera were regarded as monophyletic, or variously recognized as sister clades to each other, to the odd-nosed monkey clade, or even as basal among Asian colobines, were not rejected (P=0.097 - 0.776). Accordingly, an

unresolved polytomy among the three langur genera and the odd-nosed monkey clade may best reflect the relationships among Asian colobines based on the mitochondrial data.

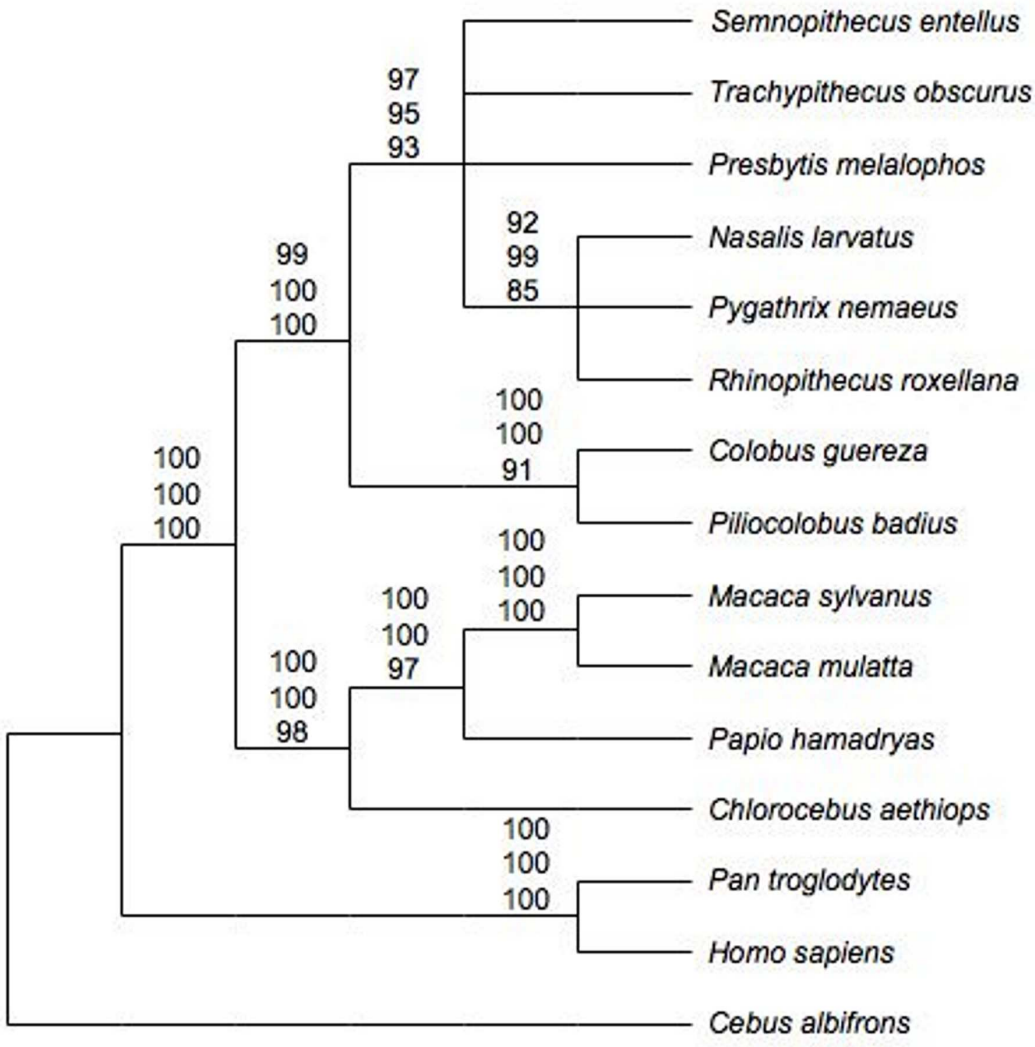

Figure 4.3: Phylogenetic relationships among colobines and related genera based on mitochondrial data. Numbers on nodes indicate support values >80% (first: ML, second: NJ, third: MP).

Several retroposon insertions were detected, which provide insights into the branching pattern of colobines (Fig. 4.4). Together with seven already published integrations (Xing et al., 2005), now 14 loci confirm monophyly of the Asian colobines to the exclusion of the African genus *Colobus*. Among Asian colobines, one insertion indicates a common origin of the odd-nosed monkeys, which is verified by three further loci (Xing et al., 2005). Most prominent are the five retroposon integrations, which confirm a sister grouping of *Trachypithecus* and *Semnopithecus*. No loci were detected, which link *Presbytis* either with the other two langur genera or the odd-nosed monkeys.

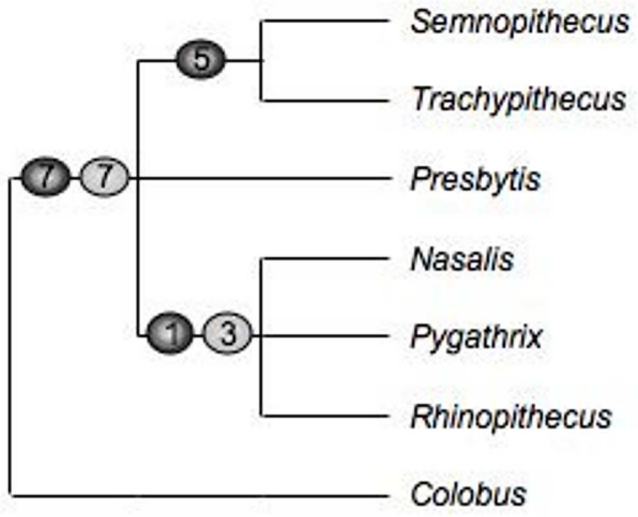

Figure 4.4: Phylogenetic relationships among Asian colobine genera based on retroposon integrations. Dark dots represent new generated data, whereas light dots refer to already published data (Xing et al., 2005). Numbers in dots indicate single integration events.

Although in general the mitochondrial data are suitable to elucidate relationships among the different genera, as indicated by the correct and significantly supported branching patterns among all other studied genera, the relationships among the langurs are unresolved, which is concordant with previous results (Sterner et al., 2006). In contrast, the presence / absence analysis of retroposon integrations provides evidence for a monophyletic odd-nosed monkey clade and a common origin of *Trachypithecus* and *Semnopithecus*, which is in agreement with morphological hypotheses (Napier, 1970; Brandon-Jones, 1984; Strasser and Delson, 1987; Davies and Oates, 1994; Brandon-Jones, 1995; Groves, 2001). Regardless which markers were used, the phylogenetic position of *Presbytis* among Asian colobines and accordingly the unity of the langurs remains unclear and needs further investigations.

4.2.2 Species group phylogeny

In order to settle affiliations among the different *Trachypithecus* species groups and their members, mitochondrial and Y chromosomal sequence data were combined with information on retroposon integrations.

The mitochondrial phylogeny was established on the basis of 573 bp long cytochrome b gene sequences, generated from most species recognized in the genus and its sister genus *Semnopithecus*. In all tree reconstructions, identical relationships were obtained, with most branches being significantly supported (Fig. 4.5a). Accordingly, the different species groups are divided into two major groups, with one including solely groups of the genus *Trachypithecus*, whereas the second one includes representatives of *Trachypithecus* and *Semnopithecus*. In the mixed clade, *T.* [*vetulus*] and *T.* [*pileatus*] members are lumped together with *S. entellus*. Whereby *T.* [*pileatus*] is monophyletic, the members of *T.* [*vetulus*] are paraphyletic, with *T. vetulus* forming a sister clade to *S. entellus* from Sri Lanka, and *T.*

johnii with *S. entellus* from South India. Furthermore, a fourth lineage in the mixed clade was detected, which is represented by *S. entellus* from North India. In contrast, the three species groups (*T.* [*obscurus*], *T.* [*francoisi*], *T.* [*cristatus*]) in the clade comprising solely *Trachypithecus* groups are all monophyletic. Alternative relationships, in which *T.* [*vetulus*] is recognized as monophyletic, either *T.* [*vetulus*] or *T.* [*pileatus*] belongs to *Trachypithecus* or even both are members of *Trachypithecus,* were tested, but all of them were rejected ($P<0.05$).

With some exceptions, the Y chromosomal data provide a similar picture (Fig. 4.5b), but due to the low number of polymorphic sites, support values are in general not as high as in the mitochondrial tree. According to the reconstructions, the species groups are divided into two major clades, with one comprising *T.* [*obscurus*], *T.* [*cristatus*], *T.* [*francoisi*] and *T.* [*pileatus*], and the other, *T.* [*vetulus*] and *Semnopithecus*. Relationships among the latter are not resolved. All alternative tree topologies, in which either *T.* [*vetulus*] belongs to *Trachypithecus* or *T. pileatus* groups with *Semnopithecus*, were rejected ($P<0.05$).

Retroposon insertions further deepened our knowledge on the species group relationships. Altogether, three informative integrations were analysed (Fig. 4.5c), with one occurring in *T.* [*obscurus*], *T.* [*cristatus*] and *T.* [*francoisi*], and the other two in *T.* [*vetulus*] and *Semnopithecus*. Interestingly, all three integrations are absent in *T. pileatus*.

With the exception of the varying position of *T.* [*pileatus*], the affiliations of the remaining species groups are congruent among different gene trees. Accordingly, all analysed markers relate *T.* [*vetulus*] with *Semnopithecus*, indicating that this species group is a real member of the genus *Semnopithecus* and not of *Trachypithecus* as assumed by morphological similarities (Groves, 2001). These similarities may be the results of adaptations to similar ecological conditions (*Semnopithecus* is semi-terrestrial and lives in deciduous forest, whereas *Trachypithecus* including *T.* [*vetulus*] is arboreal and occurs in wet evergreen forests). Although the Y chromosomal data allow no resolution within the *Semnopithecus* - *T.* [*vetulus*] clade, the mitochondrial data indicate paraphyly of the two *T.* [*vetulus*] species, with *T. vetulus* clustering with *S. entellus* from Sri Lanka and *T. johnii* with *S. entellus* from South India, which is concordant with their geographical distribution. These findings indicate paraphyly of *S. entellus,* whereby North Indian representatives form a further distinct lineage. Accordingly, the langurs of the Indian subcontinent should be split into three species groups, with one occurring solely on Sri Lanka, one in Southern India and a third one in Northern India, whereas the Gondavari river seems to be barrier between the latter two.

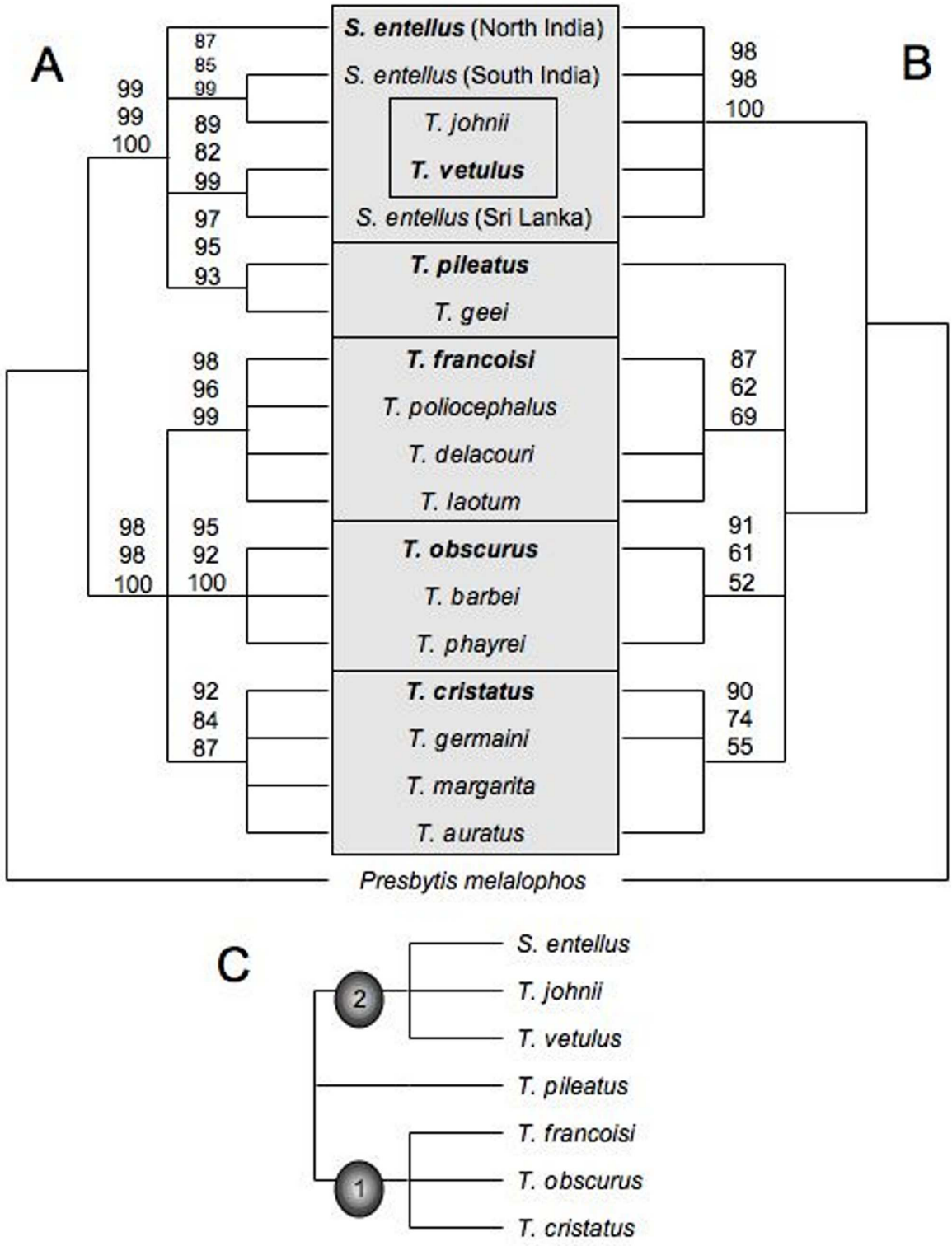

Figure 4.5: Phylogenetic relationships among *Semnopithecus* and *Trachypithecus* species groups based on a) mitochondrial data, b) Y chromosomal data, and c) retroposon integrations. Numbers on nodes indicate support values (first: ML, second: NJ, third: MP), and boxed species belong to a species group, with species in bold giving the name of the group.

Monophyly of each of *T.* [*obscurus*], *T.* [*cristatus*] and *T.* [*francoisi*] and their close affiliation is depicted in all gene trees, so that all of them can be regarded as true members of *Trachypithecus*. These findings confirm previous molecular studies (Roos, 2003; Geissmann et al., 2004; Roos, 2004; Nadler et al., 2005; Roos et al., 2007) and are in general agreement with recent classifications (Groves, 2001; Brandon-Jones et al., 2004).

The only discrepancies between different gene trees were obtained for *T.* [*pileatus*]. Whereas the mitochondrial data link the species group with *Semnopithecus* and *T.* [*vetulus*], the Y chromosomal data affiliates it with *T.* [*obscurus*], *T.* [*cristatus*] and *T.* [*francoisi*]. These findings might be explained by incomplete lineage sorting of ancestral mitochondrial or Y chromosomal haplotypes. Accordingly, the ancestor of *Trachypithecus, Semnopithecus* and *T.* [*pileatus*] carried multiple DNA lineages with one lineage being randomly fixed in two taxa, but not in the third. Alternatively, the varying position of *T.* [*pileatus*] in different gene trees might be explained by past hybridization between *Semnopithecus* and *Trachypithecus*. As depicted by the three retroposon insertions, this putative hybridization event would have occurred between ancestral forms of *Semnopithecus* and *Trachypithecus*, before both genera diverged into distinct species groups. The hybridization hypothesis is also supported by some intermediate morphological characteristics (Groves, 2001) and the distribution of *T.* [*pileatus*], which is sandwiched between those of *Semnopithecus* and other *Trachypithecus* species groups (Fig. 4.1, Fig. 4.2).

4.3 Conclusions

The present study provides detailed insights into the evolutionary history of Asian colobines and underpins the tremendous power of retroposon integrations as cladistic markers. Although mitochondrial data proved to be useful to elucidate and confirm several relationships among studied taxa, the data set failed to resolve the affiliations among the langur genera and to settle their position among Asian colobines. In contrast, retroposon insertions provided clear evidence for a sister grouping of *Semnopithecus* and *Trachypithecus*, but no integrations were detected, which link *Presbytis* either with the other two langur genera or with the odd-nosed monkeys, so that further research is required to solve this issue. Moreover, to definitively explain the evolutionary history of colobines, further molecular markers should be analysed, especially regarding possible discrepancies among gene trees due to hybridization or introgression, as such events are important speciation mechanisms in primates (Arnold and Meyer, 2006), and as it was possibly detected in the present study in the case of *T.* [*pileatus*]. Although further investigations are necessary to fully understand the evolution of colobines, the present study provides sufficient data to revise the current classification of the genera *Trachypithecus* and *Semnopithecus* (Table 4.1). Accordingly, within *Trachypithecus*, three reciprocal monophyletic species groups (*T.* [*obscurus*], *T.* [*cristatus*] and *T.* [*francoisi*]) should be recognized, whereas *T.* [*vetulus*] should be included in *Semnopithecus*.

Table 4.1: Proposed classification of *Semnopithecus* and *Trachypithecus* species based on the herein presented data.

Species group	*Semnopithecus*	*Trachypithecus*
*S. entellus** (North India)	*S. entellus**	
*S. entellus** (South India)	*S. entellus**	
	S. johnii	
*S. entellus** (Sri Lanka)	*S. entellus**	
	S. vetulus	
*T. pileatus***	*T. pileatus*	
	T. geei	
T. francoisi		*T. francoisi*
		T. poliocephalus
		T. delacouri
		T. laotum
T. obscurus		*T. obscurus*
		T. barbei
		T. phayrei
T. cristatus		*T. cristatus*
		T. germaini
		T. margarita
		T. auratus

* species / subspecies designation has to be assessed.
** provisionally classified as member of *Trachypithecus*.

Moreover, due to their paraphyletic origin, *T. vetulus* and *T. johnii* should not be lumped into a single species group, and *S. entellus* should not be regarded as monotypic. A classification into three species groups with one occurring in Northern India, one in Southern India and one on Sri Lanka may best reflect the evolutionary relationships among the langurs of the Indian subcontinent. However, this arrangement is tentative, because further research is required to confirm their distinctiveness not only on mitochondrial, but also on nuclear DNA level. *T.* [*pileatus*] might be the result of an ancient hybridization event between *Semnopithecus* and *Trachypithecus*, and hence, its classification is difficult. While a separation in its own distinct genus may be appropriate, we provisionally accept its traditional recognition as member of *Trachypithecus*.

4.4 Methods

4.4.1 Sample collection, DNA extraction and preventing contaminations

The species analysed in this study are presented in Table 4.2. All study specimens were identified by fur coloration and other external characteristics. Hanuman langur (*S. entellus*) samples were collected only from founder animals, of which the area of capture was at least roughly known. Total genomic DNA was extracted from blood, tissue or feces using the DNeasy or Stool Mini Kits from

Qiagen. When hair follicle cells were used, 1-3 hairs were directly implemented into the PCR reaction after they were washed with sterile water and 95% ethanol. To prevent contaminations, sample collection and laboratory procedures followed described standard protocols (Taberlet et al., 1999; Goossens et al., 2000; Nsubuga et al., 2004; Karanth et al., 2005). In detail, all fecal and hair samples were collected with gloves and stored in sterile tubes or plastic bags before further processing. DNA extraction, PCR, gel extraction and sequencing was performed in separate laboratories and repeated after several months, while always only one individual per species was tested. Moreover, from most specimens two different sample types were available, which both were used as template material. Sequences from independent analyses were identical. Finally, all PCR reactions were performed with negative (distilled water) controls.

Table 4.2: Species analysed, their origin, material type and GenBank accession numbers.

Species	Origin	Material type	mtDNA (5 kb)	mtDNA (573 bp)	SRY (777 bp)
Cebus albifrons	GenBank	sequence	NC_002763	-	-
Homo sapiens	GenBank	sequence	X93334	-	-
Pan troglodytes	GenBank	sequence	NC_001643	-	-
Chlorocebus aethiops	GenBank	sequence	NC_007009	-	-
Papio hamadryas	GenBank	sequence	NC_001992	-	-
Macaca mulatta	GenBank	sequence	NC_005943	-	-
M. sylvanus	GenBank	sequence	NC_002764	-	-
Colobus guereza	Cologne Zoo	tissue, feces	EU004483	-	-
Piliocolobus badius	MPI Leipzig	tissue, feces	EU004482	-	-
Pygathrix nemaeus	Cologne Zoo	tissue, feces	EU004481	-	-
Rhinopithecus avunculus	EPRC	tissue	EU004480	-	-
Nasalis larvatus	Wilhelma Stuttgart	blood	EU004476	-	-
Presbytis melalophos	Howletts Zoo	tissue, feces	EU004479	part of 5 kb	EU004456
Semnopithecus entellus (North India)	Dresden Zoo	blood, feces	EU004478	part of 5 kb	EU004457
S. entellus (South India)	Hannover Zoo	blood, feces	-	EU004471	EU004458
S. entellus (Sri Lanka)	Krefeld Zoo	hairs	-	AY519452	EU004459
Trachypithecus vetulus	Bristol Zoo	blood, feces	-	AY519454	EU004461
T. johnii	Erfurt Zoo	hairs, feces	-	AY519453	EU004460
T. pileatus	ZMB	tissue	-	EU004472	EU004462
T. geei	GenBank	sequence	-	AF294618	-
T. obscurus	Wuppertal Zoo	blood, feces	EU004477	part of 5 kb	EU004463
T. phayrei	ZMB	tissue	-	AY519460	EU004464
T. barbei	GenBank	sequence	-	AY519462	-
T. auratus	Wilhelma Stuttgart	blood	-	AY519455	EU004468
T. cristatus	Singapore Zoo	blood, feces	-	EF465128	EU004470
T. germaini	ACCB	feces	-	AY519457	EU004469
T. margarita	GenBank	sequence	-	EF465147	-
T. francoisi	Bristol Zoo	hairs, feces	-	AY519458	EU004467
T. poliocephalus	EPRC	feces	-	EU004473	-
T. delacouri	EPRC	blood, feces	-	EU004474	EU004465
T. laotum	EPRC	blood, feces	-	EU004475	EU004466

Abbreviations: ACCB - Angkor Center for Conservation of Biodiversity, Cambodia; EPRC - Endangered Primate Rescue Center, Vietnam; MPI Leipzig - Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany; ZMB - Zoologisches Museum Berlin, Germany

4.4.2 Mitochondrial sequence analysis

To determine the phylogenetic position of langur genera among Asian colobines, a ~5 kb fragment of the mitochondrial genome was sequenced from all colobine genera with the exception of *Simias* and *Procolobus*. This region spans the cytochrome b gene, the control region, the 12S rDNA and 16S rDNA, and the intermediate tRNAs. To exclude contaminations of the dataset with nuclear pseudogenes, mainly DNA extracted from feces was used as template material, and ~1-2 kb long and ~200-400 bp overlapping fragments were amplified. PCR products were generated via hot-start technique using a set of 24 primers (Table 4.3) and PCR conditions comprising a pre-denaturation step at 94°C for 2 min, followed by 40 cycles at 94°C for 1 min, 60°C for 1 min and 72°C for 1-2 min, and a final extension step at 72°C for 5 min. The results of the PCR amplifications were checked on agarose gels. PCR products were cleaned with the Qiagen PCR Purification Kit and subsequently sequenced on an ABI 3100-Avant sequencer using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems).

Table 4.3: Primers for amplifying and sequencing the 5 kb mitochondrial fragment.

Primer*	Sequence (5'-3')
L14182	AAACCATCGTTGTATTTCAACTA
L14621	GAGGACAAATATCATTTTGAGG
H14886	GTAGGGGTGGAAGGGAATTT
L15139	CACAAATCCAAACAACAAAGCA
H15246	ACCGGTTGGCTTCCAATTCA
L15322	CTCCTCAAATGAACTTGCCC
H15556	GCAGTAATGCACGAATTACATA
L15873	CCATCCTCCGTGAAATCAATA
H16117	TGCAGACCAGAGATAAAAGATA
L16248	GGTGTTATTTAATCCATGCTTG
H16402	TGTTTTTGGGGTTTGGCAAAG
L104	TTAGCAAGATTACACATGCAAG
H284	CATAGCTTAGTTAAACTTTCGTT
L501	CACTATGCTTAGCCCTAAACT
H572	AAGCTGTTGCTTGTAGTGTTC
L827	AAGAGTCCAAGGAGGATTTAG
H999	CCAGTACACTTACCATGTTAC
L1331	ACGAGCTACCCAAAAACAGC
H1535	TAAAGAGCTGTCCCTCTTTAG
L1608	TTAAGAAAGCGTTCAGCTCAA
H1883	TCCTTTTACTTTTTTTAACCTTTC
L2090	CCTGACCGTGCAAAGGTAG
H2342	TCCGAGGTCACCCCAACC
H2670	ATTACCGGGCTCTGCCATC

*Numbers refer to positions in the *Macaca sylvanus* mitochondrial genome, and H and L refer to the heavy and light strands, respectively.

To obtain a comprehensive overview on the phylogeny of colobines, the dataset was expanded with further sequences from related taxa deposited at GenBank (Table 4.2). Accordingly, the final dataset comprised 15 taxa, including eight colobines, four cercopithecines, two hominoides and a New World monkey, which was used as outgroup taxon. Sequences were aligned with ClustalW (Thompson et al., 1994) and subsequently checked by eye. Gaps and poorly aligned positions were removed with the G-blocks software (Castresana, 2000), which reduced the final dataset to 4336 bp. Based on this alignment, phylogenetic trees were constructed with the maximum-parsimony (MP), neighbor-joining (NJ) and maximum-likelihood (ML) algorithms as implemented in PAUP 4.0b10 (Swofford, 2002) and TREEPUZZLE 5.0 (Strimmer and von Haeseler, 1996). For MP analyses, all characters were treated as unordered and equally weighted throughout. A heuristic search was performed with the tree-bisection-reconnection (TBR) algorithm with random addition of sequences. The maximum number of trees was set to 100. NJ and ML trees were constructed with the GTR + I (= 0.2526) + Γ (= 0.4816) model of sequence evolution as it was selected as best-fitting model under the Akaike information criterion with MODELTEST 3.06 (Posada and Crandall, 1998). Relative support of internal nodes was performed by bootstrap analyses with 1,000 replications (MP, NJ), or by the quartet puzzling support values on the basis of 1,000 puzzling steps (ML). Finally, to evaluate the reliability of the depicted phylogenetic position of the langur genera among Asian colobines, alternative tree topologies were evaluated with the Kishino-Hasegawa (Kishino and Hasegawa, 1989) and Shimodaira-Hasegawa (Shimodaira and Hasegawa, 1999) tests in PAUP, both performed with full optimization and 1,000 bootstrap replications. Therefore, all three langur genera were regarded as monophyletic or variously recognized as sister clade to each other, to the odd-nosed monkey clade or even as sister clade to all other Asian colobines.

To determine phylogenetic affiliations of species groups, a 573 bp long fragment of the mitochondrial cytochrome b gene was analysed from all species of the different groups (Table 4.2). The generation of sequences followed laboratory methods as described (Geissmann et al., 2004; Nadler et al., 2005; Roos et al., 2007). The final alignment, which was easily generated by eye due to the absence of insertions or deletions, comprised 19 individuals including the outgroup taxon *Presbytis melalophos*. Phylogenetic trees were constructed as described above. As best-fitting model, MODELTEST selected the TIM + I (= 0.5977) + Γ (= 2.3137) model, which was applied for NJ and ML reconstructions. As for the 5 kb fragment, several alternative tree topologies, in which *T.* [*vetulus*] is recognized as monophyletic, either *T.* [*vetulus*] or *T.* [*pileatus*] belongs to *Trachypithecus*, or even both are members of *Trachypithecus*, were tested.

4.4.3 Y chromosomal sequence analysis

The SRY gene was selected as it represents a single-copy gene and as it is proved to be informative in reconstructing the Y chromosomal evolutionary history of macaques (Tosi et al., 2000). PCR conditions and primers were applied as described (Tosi et al., 2000). To amplify the SRY gene from fecal material, two overlapping fragments were generated with published primers (Tosi et al., 2000) and the newly generated internal primers 5'-TGGGCGGAGTTGAGAGGGGT-3' and 5'-TAGCGGTCCCGTTGCTGCGG-3'. The final alignment of 777 bp comprised 15 taxa representing all species groups of the genera *Semnopithecus* and *Trachypithecus* as well as *Presbytis melalophos*, which was used as outgroup. To reconstruct NJ and ML trees, the K80 model of sequence evolution, determined with MODELTEST, was used. MP trees were generated as described above. The reliability of the depicted position of *T.* [*vetulus*] and *T.* [*pileatus*] was tested in PAUP by using alternative tree topologies, in which either *T.* [*vetulus*] belongs to *Trachypithecus* or *T. pileatus* groups with *Semnopithecus*.

4.4.4 Retroposon analysis

Due to their high copy number and relative small size (~ 300 bp), the primate specific Alu elements were selected as cladistic markers. The presence or absence of Alu elements in different colobines at specific loci was tested via PCR using primers occupying the flanking region of the insertion site. Details on analysed loci, primers and studied species are listed in Table 4.4.

To detect new loci, a subtractive hybridization approach (Mamedov et al., 2005) was performed with some modifications. As tracer and driver, different colobine genera were selected. Genomic DNA of tracer and driver was digested with RSAI (Fermentas), and subsequently, the adapters AdapA1/AdapAA1 (5'- TGTAGCGTGAAGACGACAGAAAGGGCGTGGTGCGGAGGGCGGT-3' / 5'- ACCGCCCTCCG-3') and AdapA2/AdapAA2 (5'-TGTAGCGTGAAGACCTGTCTTAGGGCGTGG TGGCCAGGGCCGT-3' / 5'- ACGGCCCTGGC-3') were ligated to the tracer fragments. Each of ~15 ng tracerA1 and tracerA2 were hybridized with ~1,500 ng driver for 20 h at 60°C. 2 µl of the hybridization result was used as template to amplify solely tracer fragments using the adapter primers A1 (5'-TGTAGCGTGAAGACGACAGAA-3') and A2 (5'-TGTAGCGTGAAGACCTGTCTT-3'). The PCR program consisted of a pre-extension step at 72°C for 6 min to fill in adaptor ends, followed by 25 cycles, each with a denaturation step at 95°C for 1min, annealing at 60°C for 1 min and extension at 72°C for 2 min. To enrich fragments with Alu insertions, a semi-nested PCR was added using either primer A1 or A2 and the Alu-specific AluY (5'-GGAGAATGGCGTGAACCCGGGA-3') oligonueclotide. The PCR products were separated on agarose gels and fragments over 500 bp were excised from the gel. After purification, the fragments were cloned into the pGEMTeasy vector (Promega) and transformed into electro-competent TOP10 cells (Invitrogen).

Table 4.4: Presence / absence analysis of retroposon insertions.

Locus	Primer (5'-3')	Presence	Absence
Asia1	AAGAATCCCAGGGAAGAACACT TTGCTGGCAAAGTGACTCCT	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
Asia2	CCTGCCACTTCTGTCCATCT AGAACAAACACCAAGACAACAGC	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
Asia3	GCTTTGCCACATAAAGAGCTG GGTTAGGTGCAAATGGGAAAC	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
Asia4	TCAATCTTCCAGGGAAAATAAAG GAATATTAGTTGAAATATTTAGGC	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
Asia5	GACCATGGTAAGACAAATGTG GACTCAGGCTTAATTTTAAGTC	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
Asia6	CACCAAGCACAACTGTGAGG, TCTGCCATAGCCATCAGTCA	TOBS, PMEL, PNEM, NLAR	CGUE
Asia7	CTCTTGGTTGGGGTGAAGC GATGGTTGAACAGTGAGACTTGA	SENT, TOBS, PMEL, PNEM, NLAR	CGUE
ST1	TGATTAAAGTCAGATGAACACC GTGTAATGGGATGAAGAACAC	SENT, TVET, TDEL, TOBS, TAUR	PMEL, PNEM, NLAR
ST2	ATACATAGCATTGACTTAACTCT GATCCTGAGCCCACTATTCT	SENT, TOBS	PMEL, PNEM, NLAR
ST3	ACATCAGTGACATCAAATAAGG GAGGAAAAGATACTTTCTCATG	SENT, TOBS	PMEL, PNEM, NLAR
ST4	GGATTGAGAGCAATTTTAAAAGGA GTTCACTCCCAAATCATACTTC	SENT, TOBS	PMEL, PNEM, NLAR, CGUE
ST5	TGTAGCCAGGGAAGCCTCT TGGGATTTCTAATACTATGCCTTTG	SENT, TOBS	PMEL, PNEM, NLAR, CGUE
odd1	AGAAAGTCCCTCCCCAACAC AAGTTGGCAAAGTGGATTGC	PNEM, NLAR, RAVU	SENT, TOBS, PMEL, CGUE
T1	GAAGATTAATACTAGAAGAATCC TTGAACTTTGATCCATGGTGC	TDEL, TOBS, TAUR	TPIL, SENT, TVET, PMEL, PNEM
S1	CAAATTGTGGCTCCTTCAGTTA GGCAATGTACAGCTAACTCTGCT	SENT, TVET, TJOH	TPIL, TDEL, TOBS, TAUR, PMEL, NLAR
S2	CCCATGTGCCTTGGTTTAG GGAAGAAAGTTTGGAATGTGTG	SENT, TVET, TJOH	TPIL, TDEL, TOBS, TAUR, PMEL, NLAR

Abbreviations: CGUE - *Colobus guereza*, NLAR - *Nasalis larvatus*, PMEL - *Presbytis melalophos*, PNEM - *Pygathrix nemaeus*, RAVU - *Rhinopithecus avunculus*, SENT - *Semnopithecus entellus*, TAUR - *Trachypithecus auratus*, TDEL - *T. delacouri*, TJOH - *T. johnii*, TOBS - *T. obscurus*, TPIL - *T. pileatus*, TVET - *T. vetulus*.

Bacterial clones were collected in 96-well microtiter plates and re-screened via PCR with the primers A1 or A2 and AluY. Positive clones were sequenced and analysed with REPEATMASKER and BLAST as implemented in NCBI and EMBL. Based on the generated alignments, locus-specific primers were constructed, with the forward primer occupying a region 5'-end upstream of the insertion site, which is conserved among the colobine and human or chimp sequences. Due to the absence of the 3'-end downstream sequence of the tested colobine species, reverse primers were constructed solely on the basis of human or chimp sequences. Subsequently, the presence or absence of respective Alu insertions in different colobine species was tested via PCR. The orthology of insertions was confirmed by sequencing of at least one species per genus or species group. Sequences were deposited at GenBank and are available under the accession numbers EU004484-EU004537 and EU006662-EU006692.

Acknowledgements

We are grateful to the staff of the zoos in Cologne, Dresden, Hannover, Krefeld, Erfurt, Stuttgart, Wuppertal, Howletts, Bristol and Singapore, and to Manfred Ade (Naturhistorisches Museum Berlin, Germany), Linda Vigilant (Max-Planck-Institute of Evolutionary Anthropology, Leipzig, Germany), Kai-Olaf Krüger (Angkor Centre for Conservation of Biodiversity, Cambodia) and Tilo Nadler (Endangered Primate Rescue Center, Vietnam) for samples. Finally, we thank Christiane Schwarz for her excellent technical assistance in the laboratory. The work was financially supported by the German Primate Center (DPZ).

5 Mitochondrial versus nuclear DNA: The complex evolutionary history of colobine monkeys

Martin Osterholz*[1], Dietmar Zinner[2], Christiane Schwarz[1], Stephen Nash[3], Lutz Walter[1,3], Christian Roos*[1,4,5]

[1]Department of Primate Genetics, [2]Department of Cognitive Ethology, [4]Gene Bank of Primates, German Primate Center, Kellnerweg 4, 37077 Goettingen, Germany,
[3]Department of Anatomical Sciences, HSC, SUNY, Stony Brook, NY,
[5]IUCN/SSC Primate Specialist Group, Section Southeast Asia

*Corresponding authors

to be submitted in a modified form to Current Biology

Abstract

Background: From a genetic perspective, colobine monkeys are one of the most neglected primate subfamilies. Although various classifications and phylogenetic relationships among colobine genera, mainly based on morphological characteristics, were proposed, their evolutionary history is still controversially discussed. In recent years, an increasing number of molecular studies have been carried out, but due to contradicting gene trees, elucidating the real relationships among colobine genera became even more complicated. Insufficient data, homoplasy, differential lineage sorting or hybridization are possible explanations for gene tree discordance. However, hybridization as possible reason is traditionally neglected, since its effect in speciation processes is regarded as minimal, at least in animals. To further illuminate the evolutionary and biogeographic history of colobine genera and to determine the reasons for gene tree incongruences, we combined presence/absence analyses of Alu insertions with autosomal, Y chromosomal and mitochondrial sequence data (in total ~25,000 bp).

Results and conclusions: According to our data, a common origin of colobines, as well as of solely Asian representatives and of odd-nosed monkeys is strongly supported. However, different data sets varied in positioning *Piliocolobus* among colobines, and concerning the relationships among the odd-nosed monkey clade and the three langur genera *Presbytis*, *Semnopithecus* and *Trachypithecus*. Hence, the monophyly of African colobines and the langur group is questioned. Although homoplasy, inaccurate data and incomplete lineage sorting might have influenced gene tree topologies to some degree, they provide no sufficient explanation for the observed discordance. Hence, hybridization as main reason is favoured, which might have occurred between the two African genera *Colobus* and *Piliocolobus* and the two Asian langur genera *Semnopithecus* and *Trachypithecus*. Based on phylogenetic relationships and estimated divergence ages, colobines most likely originated in Africa. In the late Miocene, they invaded Asia and diversified within a relative short time period into respective genera. Genetic exchange between *Trachypithecus* and *Semnopithecus* occurred until the late Pliocene.

5.1 Introduction

Gene trees do not necessarily reflect species trees and even discrepancies among various gene trees are common. Homoplasy, insufficient data, differential lineage sorting, or introgressive hybridization are possible explanations for these incongruences (Avise, 2000; Barton, 2001; Hewitt, 2001; Funk and Omland, 2003; Seehausen, 2004). Accumulating evidence suggest that natural hybridization as one reason seems to be more common than previously thought. However, its implications for evolutionary processes are still unclear, at least in animals (Grant and Grant, 2002; Seehausen, 2004; Grant et al., 2005; Jorgensen and Mauricio, 2005; Arnold and Meyer, 2006; Nolte et al., 2006). Is gene flow through hybrid populations really as ineffective in speciation as it was assumed to be (Darwin, 1859; Mayr, 1942, 1963)? Several studies proposed reticulate evolution or hybridization events in various taxa, acting as driving forces for novel traits and diversification (McCracken and Sorenson, 2005; Berthier et al., 2006; Patterson et al., 2006; Pidancier et al., 2006; Koblmüller et al., 2007; Bicca-Marques et al., 2008; McDonald et al., 2008). Also for primates, several examples of natural hybridization between species (e.g. *Eulemur* sp., Wyner et al., 2002; *Lepilemur* sp., Rumpler et al., in press; *Alouatta* sp., Cortes-Ortiz et al., 2007; *Macaca* sp., Evans et al., 2003; *Papio* sp., Alberts and Altmann, 2001; *Gorilla* sp., Thalmann et al., 2007) and even the formation of new species as result of hybridization (e.g. *Macaca arctoides*, Tosi et al., 2000; *Macaca munzala*, Chakraborty et al., 2007) are

known (for review see Arnold and Meyer, 2006). However, evidence for introgressive hybridization between genera of free-ranging primates was only found between *Semnopithecus* and *Trachypithecus* (Osterholz et al., 2008), *Theropithecus* and *Papio* (Jolly et al., 1997) and the early hominine and chimpanzee lineages shortly after they initially diverged from each other (Patterson et al., 2006). These intra- and intergeneric examples suggest that natural hybridization may indeed play a role in primate speciation and diversification, but its actual impact has to be further investigated.

Colobines (subfamily Colobinae) are one of the genetically least studied primate taxa. Only a few molecular studies have been published and sequence data are limited. The classification of colobine genera and their phylogenetic relationships are still disputed and numerous controversial phylogenetic hypotheses have been proposed. Recent studies detected substantial incongruences among gene trees of some Asian colobine genera, which might be caused by ancestral hybridizations or other effects as mentioned above (Osterholz et al., 2008; Ting et al., 2008). However, a genetic analysis of colobine monkeys including sequence data from maternal-, paternal- and biparental-inherited loci or the presence/absence pattern of SINE insertions is still lacking.

Colobines form a diverse group of Old World monkeys. They share some peculiar autapomorphic features, which make them unique among primates and provide evidence for a monophyletic origin. Most prominent is the complex, at least three-chambered and ruminant-like stomach, which enables colobine monkeys to digest celluloses (Chivers and Hladik, 1980). Further typical characteristics include a high molar relief, elongated legs and tails and a reduced or even absent pollex (Chivers and Hladik, 1980; Strasser and Delson, 1987). Colobines are distributed in Africa and Asia and they most likely diverged from their closest relatives, the cheek-pouched monkeys (subfamily Cercopithecinae) in the mid-Miocene (Delson, 1994; Sterner et al., 2006).

Colobines are traditionally arranged into an African and an Asian clade, mainly based on morphological traits and geographic distribution (Napier, 1970; Szalay and Delson, 1979; Oates et al., 1994; Groves, 2001). African colobines have been further split into the three genera *Colobus, Procolobus* and *Piliocolobus* (Groves, 2001), while the latter two are sometimes combined in the genus *Procolobus* (Pocock, 1936; Kuhn, 1967; Oates et al., 1994; Grubb et al., 2003). The three African genera are believed to form a monophyletic group (Groves, 2001; Napier, 1970; Oates et al., 1994; Szalay and Delson, 1979), but paraphyly was also proposed (Jablonski, 1998). Molecular studies including all three African genera are limited. However, those in which besides *Colobus* also *Procolobus* and/or *Pilicolobus* were included, suggest a common origin of African colobines (Messier and Stewart, 1997; Sterner et al., 2006; Ting, 2008; Ting et al., 2008).

Asian colobines are more diverse than their African relatives and various alternative classifications have been proposed (Pocock, 1935, 1939; Napier and Napier, 1967; Brandon-Jones, 1984; Oates et al., 1994; Brandon-Jones et al., 2004). According to Groves (2001), the most suitable classification of Asian colobines is the division into the odd-nosed monkey group, comprising the

genera *Pygathrix*, *Rhinopithecus*, *Nasalis* and *Simias*, and the langur group, with *Semnopithecus*, *Trachypithecus* and *Presbytis*. A common origin of Asian colobines is well supported by chromosomal (Bigoni et al., 2003; Bigoni et al., 2004) and molecular studies (Collura et al., 1996; Messier and Stewart, 1997; Stewart and Disotell, 1998; Zhang and Ryder, 1998; Page et al., 1999; Xing et al., 2005; Sterner et al., 2006; Whittaker et al., 2006; Karanth et al., 2008; Osterholz et al., 2008; Ting et al., 2008). Likewise the monophyly of the odd-nosed monkey group is widely accepted and confirmed by morphological (Bennett and Davies, 1994; Groves, 2001) and genetic data (Sterner et al., 2006; Osterholz et al., 2008; Ting et al., 2008). In contrast to recent genetic data, which show that *Simias* and *Nasalis* are closely related and that both are nested within the odd-nosed monkey clade (Whittaker et al., 2006), earlier studies regarded *Simias* alone or clustered with *Nasalis* as basal among colobines (Groves, 1989; Jablonski, 1998).

The putative sister clade of the odd-nosed monkeys are the langurs. Its common origin was mainly inferred from morphological traits, reflected also by early classifications in which all three genera were combined in a single genus (*Presbytis* or *Semnopithecus*) (Napier and Napier, 1967; Groves, 1970; Delson, 1975). However, molecular evidence for the monophyly of langurs is still lacking. Nuclear and mitochondrial data revealed contrasting relationships among langur genera and the odd-nosed monkey group (Zhang and Ryder, 1998; Sterner et al., 2006; Osterholz et al., 2008; Ting et al., 2008). Morphological similarities between *Semnopithecus* and *Trachypithecus* lead to the exclusion of *Presbytis* (Brandon-Jones, 1984; Strasser and Delson, 1987; Groves, 2001), a pattern which is also supported by nuclear data (Osterholz et al., 2008; Ting et al., 2008). However, mitochondrial sequence data do either not resolve these relationships (Osterholz et al., 2008) or even suggest a clade consisting of *Presbytis* and *Trachypithecus* (Sterner et al., 2006).

SINE (Short INterspersed Element) insertions, in particular the primate-specific Alu elements, have proven to be useful at various levels of phylogenetic analyses and are widely applied as a powerful molecular-cladistic tool (Schmitz et al., 2001; Roos et al., 2004; Ray et al., 2005; Schmitz et al., 2005; Xing et al., 2005; Herke et al., 2007; Xing et al., 2007; Osterholz et al., 2008). The probability of an Alu element to insert independently into the same locus in different genomes is insignificant and a true orthology can be verified by their flanking direct repeats. Furthermore, the insertion is irreversible and the precise removal is extremely unlikely (Shedlock and Okada, 2000; Schmitz et al., 2005; van de Lagemaat et al., 2005). Consequently, the lack of an Alu element at a certain locus can be considered to be the ancestral state and therefore a common ancestor can be determined by a shared insertion (Okada, 1991; Shedlock and Okada, 2000; Schmitz et al., 2005). Based on these characteristics, SINE insertions can be regarded as noise-free Hennigian synapomorphies (Shedlock and Okada, 2000).

To further elucidate the supposed complex evolutionary history of colobine monkeys, we examined the presence/absence pattern of SINE insertions and compared the inferred phylogeny with

that derived from mitochondrial and nuclear sequence data. To detect possible hybridization events among genera, we analysed maternal-, paternal- and biparental-inherited loci. Accordingly, five autosomal loci, which map to different human chromosomes, and six Y chromosomal loci were PCR-amplified and sequenced. As maternal-inherited marker, we reanalysed available mitochondrial genome data.

5.2 Results

5.2.1 Alu insertions

In total, 189 formerly undetected Alu insertion loci were tested for presence or absence of respective insertions. Among them, only 13 loci proved to be phylogenetically informative. All other Alu elements were present only in a single genus or the amplification in essential genera was not possible. To increase the number of informative loci, another 153 insertions known from earlier studies were included (Salem et al., 2003; Xing et al., 2005; Osterholz et al., 2008), of which some were further examined in previously untested genera (Supplementary table 5.2). The orthology of insertions was confirmed by sequencing. In all cases, direct repeats flanking the insertion as well as the original target site prior transposition were detected. Moreover, no insertions showing conflicting relationships were found.

Based on these insertions, an unequivocal phylogeny of colobine genera and outgroup taxa could be established (Fig. 5.1A). The sister grouping of *Homo* and *Pan* (A-I) is supported by 16 insertions (Salem et al., 2003). Among Cercopithecinae 15 insertions support a common origin of *Macaca* and *Papio* (A-II) and the clade consisting of the latter two genera and *Chlorocebus* (A-III) is supported by 33 insertions (Xing et al., 2005). The monophyly of all Cercopithecidae representatives (A-IV) is supported by 35 insertions (Xing et al., 2005).

20 Alu insertions (A-V) unite all colobines and strongly support their monophyly. Among colobines, three unlinked Alu insertions (A-VI) were found in *Piliocolobus* and all Asian colobines, but not in *Colobus* and hence, a paraphyly of African colobines is indicated. A common origin of solely Asian genera (A-VII) is supported by 26 Alu insertions. Among Asian colobines, ten Alu elements unite *Semnopithecus* and *Trachypithecus* (A-VIII) to the exclusion of all other colobines, while two insertions link *Presbytis* with *Nasalis/Pygathrix* (A-IX). The monophyly (A-X) of the latter two genera is supported by six insertions.

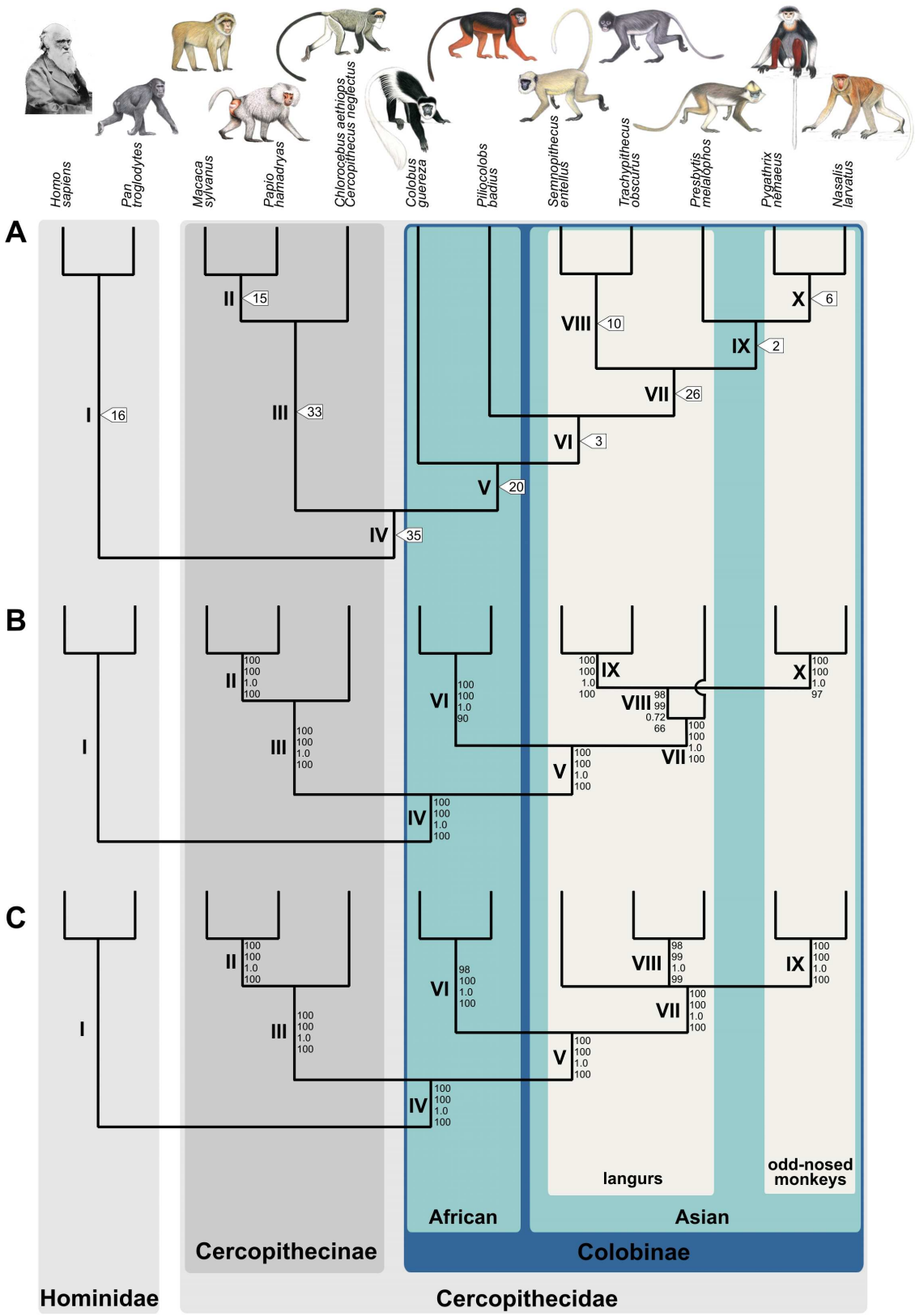

Figure 5.1: Phylogenetic relationships among colobine genera and outgroup taxa as inferred from (A) Alu insertion loci, (B) nuclear sequence data, and (C) mitochondrial genome data. Numbers in flags and beside branches indicate the number of Alu insertions and bootstrap support or posterior probability values (top to bottom: MP, NJ, Bayesian, ML), respectively.

5.2.2 Nuclear phylogeny

Our complete nuclear data set, including five autosomal and six Y chromosomal loci, comprises 8,917 bp, of which 1,291 sites are variable and 722 parsimony-informative. Phylogenetic analyses yielded identical and for most cases significantly supported branching patterns among studied taxa irrespectively of the applied algorithm (maximum-parsimony: MP, neighbor-joining: NJ, maximum-likelihood: ML, Bayesian). Tree topologies for individual loci or concatenated autosomal and Y chromosomal data sets were in general congruent with the tree derived from the combined data, but in some cases no resolution of relationships were obtained and support values were in general lower (data not shown).

Based on nuclear data (Fig. 5.1B), significant support was obtained for a common origin of Cercopithecidae (B-IV) to the exclusion of hominids (B-I) and its division into Cercopithecinae (B-III) and Colobinae (B-V). Among cercopithecines, *Macaca* and *Papio* (B-II) form a strongly supported sister clade to *Cercopithecus*. Among colobines, African (B-VI) and Asian (B-VII) colobine genera form reciprocal monophyletic groups, a division, which is significantly supported by MP, NJ and Bayesian algorithms, but ML bootstrap values for a sister grouping of the two African genera do not exceed 90%. The Asian genera further diverged into a lineage leading to *Presbytis* and another to all the remaining genera (B-VIII), which later split into two clades, with one consisting of *Trachypithecus* and *Semnopithecus* (B-IX) and the second of *Nasalis* and *Pygathrix* (B-X). Both clades are strongly supported.

Although most relationships are significantly supported, some of them are not in agreement with the tree topology as obtained from presence/absence data of Alu elements. Accordingly, we tested the reliability of the depicted relationships by evaluating alternative tree topologies with the Kishino-Hasegawa and Shimodaira-Hasegawa tests. Based on these data, a paraphyly of African colobines with *Piliocolobus* being closer related to Asian colobines than to *Colobus* as proposed by three Alu insertions is not rejected ($P>0.05$). Traditionally, *Presbytis* is recognized as member of the langur group, which also includes *Semnopithecus* and *Trachypithecus*, but nuclear data indicate a basal position of *Presbytis* among Asian colobines and two Alu insertions support a close relationship to the odd-nosed monkeys. Alternative positions of *Presbytis* among Asian colobines, in which all these proposed relationships were tested, are not rejected. However, a close affiliation of *Presbytis* to either *Semnopithecus* or *Trachypithecus* is rejected on a significance level of $P<0.005$.

Based on these findings, it is highly likely that nuclear sequence data do not reflect the real relationships among colobine monkeys, which might be caused by the relative low number of polymorphic sites. Hence, divergence ages from nuclear sequence data were estimated on the basis of an *a priori* fixed tree topology as proposed by Alu insertions. Accordingly, the major split among cercopithecids leading to the two subfamilies Colobinae and Cercopithecinae occurred ~17.4 ± 1.4 million years ago (mya), while among the latter *Cercopithecus* diverged from the *Macaca/Papio* clade

~9.6 ± 0.9 mya (Table 5.1). Among colobines, *Colobus* and *Piliocolobus* split off from the Asian genera ~10.7 ± 1 mya and ~10.4 ± 1 mya, respectively. The Asian lineage further diverged ~7.6 ± 0.8 mya into a group consisting of *Trachypithecus*/*Semnopithecus* and a clade with *Presbytis*, *Nasalis* and *Pygathrix*. Among the latter, *Presbytis* split off ~7.6 ± 0.8 mya, before finally also *Nasalis* and *Pygathrix* diverged ~5.7 ± 0.8 mya. The most recent split among Asian genera occurred between *Trachypithecus* and *Semnopithecus* (~2.3 ± 0.5 mya).

Table 5.1: Estimated divergence times

Divergence	**nucDNA** HKY+G, C3	**mtDNA** GTR+I+G, C3
Homo – Pan	6.1 ± 0.2	6.0 ± 0
Cercopithecoidea – Hominoidea	34.3 ± 0.5	28.2 ± 1.7
Cercopithecinae – Colobinae	17.4 ± 1.4	17.1 ± 0.9
Cercopithecini – Papionini	9.6 ± 0.9	10.1 ± 0.5
Papio – *Macaca*	6.3 ± 0.3	7.9 ± 0.1
Colobus – (Piliocolobus, Asian colobines)	10.7 ± 1	-
Piliocolobus – Asian colobines	10.4 ± 1	-
African colobines – Asian colobines	-	11.1 ± 0.6
Colobus – *Piliocolobus*	-	8.3 ± 0.6
Asian split	7.6 ± 0.8	8.8 ± 0.4
Trachypithecus – *Semnopithecus*	2.3 ± 0.5	-
Presbytis – odd-nosed monkeys	7.6 ± 0.8	-
Presbytis – *Trachypithecus*	-	7.7 ± 0.5
Nasalis – *Pygathrix*	5.7 ± 0.8	6.4 ± 0.2

5.2.3 Mitochondrial phylogeny

Among the 15,905 bp in the mitochondrial alignment, 6,995 sites are variable and 5,131 of them parsimony-informative. Phylogenetic relationships among genera as suggested by various algorithms are mainly congruent and significantly supported (Fig. 5.1C). In detail, the data strongly support a Cercopithecidae clade (C-IV) to the exclusion of *Homo* + *Pan* (C-I), the monophyly of the two cercopithecid subfamilies Colobinae (C-V) and Cercopithecinae (C-III), and among the latter, a sister grouping of *Chlorocebus* to a *Papio/Macaca* clade (C-II). Colobines initially diverged into an African (C-VI) and an Asian (C-VII) clade. The Asian colobine clade further splits into a lineage leading to *Pygathrix/Nasalis* (C-IX), a lineage comprising *Trachypithecus* and *Presbytis* (C-VIII), and finally a lineage with *Semnopithecus*. The relationships among these three lineages are not well resolved. While maximum-parsimony, neighbor-joining and Bayesian approaches weakly support a sister grouping of *Semnopithecus* to the *Presbytis/Trachypithecus* clade, the maximum-likelihood inference indicates a clade

consisting of *Semnopithecus* and the *Nasalis/Pygathrix* clade (data not shown). Finally, the monophyly of both the *Nasalis/Pygathrix* and *Presbytis/Trachypithecus* clades are again strongly supported.

Due to incongruences between nuclear and mitochondrial trees, alternative tree topologies were also tested for the mitochondrial data set. A paraphyly of African colobines with *Piliocolobus* being closer related to Asian colobines than to *Colobus*, as indicated by Alu insertions, is significantly rejected ($P<0.001$). Nuclear and mitochondrial tree topologies differ also concerning the relationships among the three langur genera *Semnopithecus*, *Trachypithecus* and *Presbytis*, and the odd-nosed monkey clade. The tests revealed, that relationships in which *Trachypithecus* and *Presbytis* do not form a monophyletic group are significantly rejected ($P<0.005$) and also the close relationship of *Trachypithecus* and *Semnopithecus* as suggested by nuclear data is rejected ($P<0.005$). In contrast, different positions of *Semnopithecus* among Asian colobines as the grouping with either the *Trachypithecus/Presbytis* or the *Nasalis/Pygathrix* clade, as well as a basal position or an unresolved trichotomy among these three lineages are not rejected ($P>0.05$). Hence, an unresolved trichotomy might best reflect the relationships among these three lineages in the mitochondrial tree.

As for the nuclear data set, divergence ages for the mitochondrial genome data were estimated on the basis of an *a priori* fixed tree topology. Relationships were in general in agreement with those depicted in Fig. 5.1C, with exception of the unresolved trichotomy among the three major Asian lineages. Based on the estimates (Table 5.1), the initial split among cercopithecids leading to the two subfamilies Colobinae and Cercopithecinae occurred ~17.1 ± 0.9 mya. Among Cercopithecinae, *Chlorocebus* diverged from the *Macaca/Papio* clade ~10.1 ± 0.5 mya, while the latter split into its respective genera ~7.9 ± 0.1 mya. Colobines initially split into an African and an Asian lineage ~11.1 ± 0.6 mya. The two African genera finally diverged from each other ~8.3 ± 0.6 mya. The major Asian split leading to the three lineages *Semnopithecus*, *Trachypithecus/Presbytis* and *Nasalis/Pygathrix* occurred ~8.8 ± 0.4 mya. The latter two clades further diverged into respective genera ~7.7 ± 0.5 and ~6.4 ± 0.2 mya.

5.3 Discussion

By combining presence/absence analyses of Alu integration with extensive autosomal, Y chromosomal and mitochondrial sequence data (in total ~25,000 bp), the present study provides new and comprehensive insights into the evolutionary history of colobine monkeys. Most relationships are resolved and strongly supported by various algorithms and Alu insertions, and the branching patterns for which statistical support is relative low, are evidently confirmed by Alu insertions. Moreover, divergence ages and relationships as revealed from the different data sets are mainly congruent and in agreement with estimates from other molecular studies (Collura et al., 1996; Stewart

and Disotell, 1998; Zhang and Ryder, 1998; Raaum et al., 2005; Xing et al., 2005; Sterner et al., 2006; Karanth et al., 2008; Osterholz et al., 2008; Ting, 2008; Ting et al., 2008).

Identical branching patterns among different gene trees were obtained for the division of Cercopithecidae into two subfamilies (Colobinae, Cercopithecinae), a sister grouping of the *Macaca/Papio* clade to Cercopithecini, represented by either *Cercopithecus* or *Cholorocebus*, and a common origin for both an Asian colobine, and an odd-nosed monkey clade within the Asian colobines. Also divergence times for respective splits are comparative among gene trees and to previous data (Goodman et al., 1998; Stewart and Disotell, 1998; Page et al., 1999; Raaum et al., 2005; Steiper and Young, 2006; Sterner et al., 2006; Ting, 2008). Accordingly, the major split among cercopithecids occurred ~17.5 ± 1.4 mya. The *Macaca/Papio* clade diverged from Cercopithecini ~9.7 ± 1 mya, while the last common ancestor of Asian colobines and odd-nosed monkeys lived ~8.1 ± 1.2 mya and ~5.8 ± 0.8 mya, respectively.

However, our study revealed also massive discrepancies among gene trees. First, mitochondrial data significantly support a sister grouping of the two African colobine genera *Colobus* and *Piliocolobus*, whereas Alu insertions provide evidence for a closer affiliation of *Piliocolobus* to Asian colobines than to *Colobus*. Although nuclear sequence data support the mitochondrial pattern, an alternative relationship with *Piliocolobus* forming a sister genus to Asian colobines to the exclusion of *Colobus* is not rejected. Accordingly, mitochondrial DNA supports the monophyly of African colobines, while nuclear data indicate their paraphyly. The observed discrepancy becomes also obvious by considering the estimated divergence ages. Based on mitochondrial data, African and Asian colobines diverged ~11.1 ± 0.6 mya and the split between *Colobus* and *Piliocolobus* occurred ~8.3 ± 0.6 mya. In contrast, the nuclear-based estimates indicate a separation of *Colobus* from *Piliocolobus* and Asian colobines ~10.7 ± 1 mya, and a division of the latter two shortly after, ~10.4 ± 1 mya.

The second example for discrepancies among gene trees concerns the relationships among the three langur genera *Presbytis*, *Semnopithecus* and *Trachypithecus*, and the odd-nosed monkey group. Nuclear sequence data and Alu insertions support close relationships between *Semnopithecus* and *Trachypithecus*, and between *Presbytis* and the odd-nosed monkeys. The latter clade is mainly supported by two Alu insertions and nuclear sequence data do not contradict this hypothesis. In contrast to nuclear data, mitochondrial DNA strongly links *Trachypithecus* with *Presbytis*. The position of *Semnopithecus* relative to the clades *Presbytis/Trachypithecus* and odd-nosed monkeys is not resolved. Hence, the mitochondrial relationships among the three Asian lineages *Semnopithecus*, *Presbytis/Trachypithecus* and odd-nosed monkeys might be best displayed as trichotomic split. Based on nuclear sequence data, the split between *Semnopithecus/Trachypithecus* and *Presbytis*/odd-nosed monkeys occurred ~7.6 ± 0.8 mya. *Presbytis* diverged from the odd-nosed monkeys shortly afterwards, while *Trachypithecus* and *Semnopithecus* had a common ancestor until ~2.3 ± 0.5 mya. According to mitochondrial DNA, the initial split among Asian colobines into the three lineages *Semnopithecus,*

Presbytis/Trachypithecus and odd-nosed monkeys occurred ~8.8 ± 0.4 mya, while *Presbytis* and *Trachypithecus* diverged from each other ~7.7 ± 0.5 mya.

Inadequate data, homoplasy, incomplete lineage sorting or hybridization are possible explanations for the observed incongruence (Avise, 2000; Funk and Omland, 2003). For the mitochondrial data set, at least for the *Colobus/Piliocolobus* and *Presbytis/Trachypithecus* clades, incorrect branching patterns due to inadequate data or homoplasy are unlikely, since a sufficient phylogenetic resolution with long branches was obtained. However, for the unresolved position of *Semnopithecus* among Asian colobines, inadequate data are the most likely explanation. Differential sorting of ancestral mitochondrial lineages can also be excluded. If the *Colobus/Piliocolobus* and *Presbytis/Trachypithecus* clades would be indeed the result of incomplete sorting of mitochondrial lineages, the mitochondrial divergence between respective genera should predate the nuclear splitting times, which is not the case. For the nuclear data set, homoplasy, incomplete lineage sorting and insufficient data can be excluded as well, at least for the branching of *Trachypithecus* and *Semnopithecus*, because ten independent Alu insertions and sequence data from six Y chromosomal and five unlinked autosomal loci clearly confirm their close relationship. However, the ambiguous relationships of *Piliocolobus* and *Presbytis* among colobines in the nuclear sequence data set might be affected by homoplasy, but insufficient data might be an alternative explanation therefore. Hence, the positioning of *Presbytis* and *Piliocolobus* among colobines does not rely on nuclear sequence data, but mainly on Alu insertions. Homoplasy and incomplete lineage sorting for Alu insertions is unlikely, because the probability that several unlinked Alu elements integrate independently at orthologous sites in different genomes is implausible and neglectable (Shedlock and Okada, 2000; Schmitz et al., 2005).

Although inadequate data, homoplasy and differential lineage sorting may have affected the tree topologies to some degree, these factors provide no sufficient explanation for the herein detected gene tree discordances and hence, ancestral hybridization events as possible reason are favoured therefore. A hybridization hypothesis gains further support by some natural characteristics as behavior or morphology, or distribution.

Hybridization between the two African genera *Colobus* and *Piliocolobus* is in principal possible, because they occur sympatrically over wide ranges of their distribution (Oates et al., 1994). Furthermore, in contrast to *Colobus* and most other primates, females in *Piliocolobus* tend to leave their natal groups (Newton and Dunbar, 1994; Oates et al., 1994) and finally, *Colobus* males are in average larger than males in *Piliocolobus* (Oates et al., 1994). Accordingly, we propose a scenario in which proto-*Colobus* was separated from the main colobine stem ~10.7 ± 1 mya, shortly afterwards followed by proto-*Piliocolobus*. After these initial splits, a limited number of *Piliocolobus* females hybridized with *Colobus* males leading to female (mitochondrial) introgression of *Piliocolobus* into *Colobus*. By backcrossing of hybrid females with *Colobus* males, the *Piliocolobus* mitochondrial lineage became

fixed in *Colobus*, while the original nuclear lineages of *Colobus* increased in every generation. Based on our calculations, this hybridization event occurred ~8.3 ± 0.6 mya, although the possibility that hybridization took place over a longer time period, beginning directly or sometime after the initial split ~10.7 ± 1 mya until the divergence of the mitochondrial lineages, can not be excluded.

Ancestral hybridization seems to be also the most plausible explanation for the gene tree incongruences among Asian colobines. The close relationship between *Semnopithecus* and *Trachypithecus* is also supported by morphological similarities, and hybridization events due to partly overlapping distribution zones are in general possible (Oates et al., 1994; Groves, 2001; Brandon-Jones et al., 2004). Accordingly, in an initial split, Asian colobines diverged ~8.1 ± 1.2 mya into *Semnopithecus* and a progenitor of *Presbytis*, *Trachypithecus* and the odd-nosed monkeys. Among the latter, the progenitor of odd-nosed monkeys split off shortly after the initial division, while *Presbytis* and *Trachypithecus* were separated from each other ~7.7 ± 0.5 mya. *Semnopithecus* males, which are much larger than males of *Trachypithecus* (Oates et al., 1994), hybridized with *Trachypithecus* females, leading to the fixation of mitochondrial DNA in hybrids. By backcrossing with pure *Semnopithecus* males over a long period (until ~2.3 ± 0.5 mya), the progenitor of *Trachypithecus* accumulated solely nuclear material of *Semnopithecus* (nuclear swamping), while the mitochondrial genome remained *Trachypithecus*-like.

Based on phylogenetic relationships and estimated divergence ages, colobines most likely originated in Africa, which is in agreement with earlier suggestions (Stewart and Disotell, 1998). Support for this hypothesis is provided by the fact that with exception of macaques, which mainly occur in Asia, all other recent members of the sister lineage to colobines, the cercopithecines, are solely distributed in Africa. Another strong hint for an African origin is provided by the paraphyly of African colobines, because a putative Asian origin would require two migration steps into Africa. According to our data, colobines and cercopithecines were separated from each other ~17.5 ± 1.4 mya. Among colobines, *Colobus* split off first from the main stem (~10.7 ± 1 mya). Shortly afterwards, in a time frame of only ~200,000 years, also *Piliocolobus* diverged from the progenitor of Asian colobines. Although initially separated, hybridization events between *Piliocolobus* and *Colobus* occurred until ~8.3 ± 0.6 mya. The progenitor of Asian colobines most likely invaded Eurasia via an emerging landbridge connection between Africa and the Arabian Peninsula in the late Miocene (Whybrow, 1984, 1992; Stewart and Disotell, 1998), sometimes after the separation from *Piliocolobus* 10.4 ± 1 mya and before its diversification ~8.1 ± 1.2 mya. Whether a route north or south of the Himalaya was selected is putative, but at least north of the Himalaya, on the Tibetan plateau, colobine fossils from the late Miocene were found, but not south of the Himalaya (Delson, 1994). The diversification hotspot of Asian colobines is unknown, but the Hengduan Mountains in the border region of today's Burma, India and China might have been a possible place (Peng et al., 1993; Jablonski, 1998). This is supported by the fact that all of the larger Southeast-Asian rivers (Mekong, Salween, Yangtze) emerge there,

which are well-known to act as barriers for arboreal primates (Meijaard and Groves, 2006). As first lineage, *Semnopithecus* diverged from the main stem and invaded the Indian subcontinent. Shortly afterwards, also the progenitor of *Presbytis* and *Trachypithecus* separated from the odd-nosed monkey ancestor and migrated into southern mainland Asia. Finally, *Presbytis* entered Sundaland and diverged from the southern mainland form *Trachypithecus* ~7.7 ± 0.5 mya. The odd-nosed monkeys also split into different lineages, which further expanded their range into Indochina and Sundaland ~5.8 ± 0.8 mya. *Trachypithecus* and *Semnopithecus* came into secondary contact and exchanged genetic material until ~2.3 ± 0.5 mya, most likely in the region of today's Bangladesh, Burma and the Northeast of India.

5.4 Conclusions

Our study gives new and detailed insights into the evolutionary and biogeographic history of colobine monkeys, and provides evidence for hybridization events among ancestral genus lineages. The analysis of presence/absence patterns of Alu insertions proved to be the most reliable tool to reconstruct phylogenetic relationships, although also pure sequence data contributed to some degree. However, as this study has shown, Alu insertion data alone are not sufficient to confirm putative hybridization events, because data from maternal-, paternal- and biparental inherited loci are required to fully elucidate such events. Hybridization among taxa is traditionally recognized as factor, which leads to limited diversification, reproductive isolation and lowered fitness (Darwin, 1859, Mayr, 1963). However, our and other data clearly support that hybridization plays a role in the diversification of primates, although its extend has to be further investigated.

5.5 Experimental Procedures

5.5.1 Samples

Total genomic DNA was extracted from blood or tissue samples using the DNeasy Blood & Tissue Kit from Qiagen. Samples were obtained from *Colobus guereza* (Guereza, Cologne Zoo, Germany), *Piliocolobus badius* (Western red colobus, Taï National Park, Ivory Coast; provided by Linda Vigilant, Max-Planck-Institut Leipzig, Germany), *Semnopithecus entellus* (Hanuman langur, Dresden Zoo, Germany), *Trachypithecus obscurus* (Dusky leaf monkey, Wuppertal Zoo, Germany), *Presbytis melalophos* (Sumatran surili, Howletts Wild Animal Park, UK), *Pygathrix nemaeus* (Red-shanked douc langur, Cologne Zoo, Germany), *Nasalis larvatus* (Proboscis monkey, Wilhelma Stuttgart, Germany; provided by Werner Schempp, University of Freiburg, Germany), *Macaca sylvanus* (Barbary macaque,

Nuremberg Zoo, Germany), *Papio hamadryas* (Hamadryas baboon, Munich Zoo, Germany), *Cercopithecus neglectus* (De Brazza's Monkey, Leipzig Zoo, Germany), *Chlorocebus aethiops* (Grivet, provided by Roland Plesker, Paul-Ehrlich-Institut Langen, Germany) and *Pan troglodytes* (Chimpanzee, Munich zoo, Germany). The only colobines not included in this study are *Simias, Rhinopithecus* and *Procolobus*, since all of them are closely related to other genera (Groves, 1970; Delson, 1975; Whittaker et al., 2006). Outgroup taxa include representatives from all relevant catarrhine groups (*Homo*, *Pan*, *Macaca*, *Papio*, *Cercopithecus* or *Chlorocebus*). For a more comprehensive analysis, further sequences deposited at GenBank were added to the data sets.

5.5.2 Alu insertions

Due to their high copy number and relative small size (~300 bp), the primate specific Alu elements were selected as molecular-cladistic markers. The presence or absence of Alu elements in different colobines at specific loci was tested via PCR using primers occupying the flanking region of the insertion site. The orthology of insertions was confirmed by sequencing. Details on analysed loci, primers and presence/absence pattern of Alus in studied species are listed in supplementary table 5.2.

To detect new Alu insertion loci, a subtractive hybridization approach was performed as described in Mamedov et al. (2005) and Diatchenko et al. (1996), but with some modifications (Osterholz et al., 2008). To avoid biased hybridization results, various species combinations were used as tracer and driver (hybridization (hyb.) 1: tracer *Nasalis/Pygathrix*, driver *Presbytis*; hyb. 2: tracer *Nasalis/Pygathrix*, driver *Semnopithecus*; hyb. 3: tracer *Trachypithecus/Presbytis*, driver *Pygathrix*; hyb. 4: tracer *Presbytis*, driver *Semnopithecus*). A detailed description of respective laboratory methods is outlined in Osterholz et al. (2008). To increase the number of informative loci, another 153 insertions (Salem et al., 2003; Xing et al., 2005; Osterholz et al., 2008) were included, of which colobine markers were further examined in previously untested genera (Supplementary table 5.2).

5.5.3 PCR and sequencing of nuclear loci

Exonic, but mainly intronic sequences were generated for six single-copy genes mapping to the non-recombining portion of the Y chromosome, and for five autosomal loci, which map to different human chromosomes. With exception of the SRY gene (sex-reversal, Y chromosome), all other Y chromosomal loci (DBY5: Dead Box, intron 5; SMCY7: SMC mouse homologue, intron 7; SMC11: SMC mouse homologue, intron 11; UTY18: ubiquitous TPR motif, intron 18; ZFYLI: Zinc finger, last intron) have homologues on the X chromosome (X degenerate). PCR conditions and primers for amplifying Y chromosomal loci were applied as listed in supplementary table 5.1. Intron 11 of the von Willebrand Factor (vWF11), located on the short arm of human chromosome 12, was amplified following methods as described in Chaves et al. (1999). For intron 3 of the serum albumin gene (ALB3, human chromosome 4), intron 3 of the interstitial retinol-binding protein, (IRBP3, human chromosome 10),

intron 1 of the transition protein 2 (TP2, human chromosome 16) and intron 1 of the transthyretin gene (TTR1, human chromosome 18), new primers were designed on the basis of available primate sequences in GenBank. Oligonucleotide primers, PCR conditions and the size of PCR products are listed in supplementary table 5.1. The results of the PCR amplifications were checked on agarose gels. PCR products were cleaned with the Qiagen PCR Purification Kit and subsequently sequenced on an ABI 3730xl sequencer using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). Sequences were deposited at GenBank and are available under the accession numbers xxxxxx-xxxxxx.

5.5.4 Phylogenetic reconstructions

Phylogenetic analyses were conducted on the basis of nuclear sequence data, which were newly generated in this study (see above), and with complete mitochondrial genome sequences, which were obtained from GenBank (Supplementary table 5.3). Both data sets included seven colobine (*Colobus, Piliocolobus, Trachypithecus, Semnopithecus, Presbytis, Nasalis, Pygathrix*), three cercopithecine (*Macaca, Papio, Chlorocebus* or *Cercopithecus*) and two hominid genera (*Homo, Pan*), while the latter two were used as outgroup taxa. Sequences were aligned with ClustalX (Thompson et al., 1994) and manually adjusted. After the manual removal of indel positions (758 bp), the concatenated nuclear alignment comprised 8,917 bp. In the mitochondrial data set, 895 ambiguous positions were removed with G-blocks (Castresana, 2000), resulting in an alignment with 15,905 bp in length. Phylogenetic trees were constructed with MP and NJ algorithms as implemented in PAUP 4.0b10 (Swofford, 2002) as well as with ML and Bayesian algorithms, using the programs GARLI 0.951 (Zwickl, 2006) and MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003), respectively. Calculations were performed for each locus separately, for concatenated autosomal and Y chromosomal data sets and for a data set, in which all nuclear data were combined. However, mitochondrial and nuclear data were never combined. For MP analyses, all characters were treated as unordered and equally weighted throughout. A heuristic search was performed with the maximum number of trees set to 100. NJ, ML and Bayesian trees were constructed with the respective best-fitting models as selected under the Akaike information criterion with MODELTEST 3.7 (Posada and Crandall, 1998) (see also supplementary table 5.1). Relative support of internal nodes was performed by bootstrap analyses with 10,000 (MP, NJ) or 500 replications (ML). In GARLI, only the model specifications settings were adjusted according to the respective data set, while all other settings were left at their default value. ML majority-rule consensus trees were calculated in PAUP. For Bayesian analyses, four Monte Carlo Markov Chains (MCMC) with the default temperature of 0.1 were used. Four repetitions were run for 10,000,000 generations with tree and parameter sampling occurring every 100 generations. The first 25% of samples were discarded as burnin, leaving 75,001 trees per run. Posterior probabilities for each split were calculated from the posterior density of trees.

To evaluate the reliability of the depicted relationships among colobines, alternative tree topologies were evaluated with the Kishino-Hasegawa (Kishino and Hasegawa, 1989) and Shimodaira-Hasegawa (Shimodaira and Hasegawa, 1999) tests with full optimization and 1,000 bootstrap replications in PAUP. Therefore, various hypothetical sister group relationships among Asian colobine lineages as well as a putative paraphyly of African colobines was tested.

5.5.5 Divergence age estimation

A Bayesian MCMC method, which employs a relaxed molecular clock approach, as implemented in the BEAST 1.4.8 package (Drummond and Rambaut, 2007), was applied to estimate divergence times. Calculations for the concatenated nuclear and mitochondrial data sets were performed separately and on the basis of *a priori* fixed tree topologies. For all estimations, a relaxed lognormal model of lineage variation, a Yule prior for branching rates and the optimal nucleotide substitution model as listed in supplementary table 5.1 were used. As calibration points we applied the split between *Homo* and *Pan*, which has been dated at 6-7 mya (Steiper and Young, 2006), the divergence between *Papio* and *Macaca* which is estimated at 6-8 mya (Steiper and Young, 2006), and the spilt between hominids and cercopithecids, which is estimated at 26.9-36.4 mya (Steiper and Young, 2006). Four replicates were run for 10 million generations with tree and parameter sampling occurring every 100 generations. The adequacy of a 10% burnin and convergence of all parameters were assessed by visual inspection of the trace of the parameters across generations using the software TRACER 1.4 (Rambaut and Drummond, 2007). Subsequently, the sampling distributions were combined (25% burnin) using LogCombiner 1.4.8 and a consensus chronogram with node height distribution was generated and visualized with the software TreeAnnotator 1.4.8 and FigTree v1.1.2 (Rambaut, 2008).

Acknowledgements

We thank the zoos in Cologne, Dresden, Howletts, Leipzig, Munich, Nuremberg, Stuttgart and Wuppertal, and Linda Vigilant, Roland Plesker and Werner Schempp for providing samples. This study was financially supported by the German Primate Center.

6 Summary and Conclusion

6.1 Summary

6.1.1 New World monkeys

Although already two SINE based studies, carried out by Singer et al. (2003) and Ray et al. (2005), are available, several important questions in New World monkey phylogeny remained. The reason therefore is that only few insertion events were analysed (Singer et al., 2003) or that a relative large number of genera were not tested (Ray et al., 2005). The first and most important issue in New World monkey phylogeny is the branching order among the three families Cebidae, Atelidae and Pitheciidae, which was not solved with significance in earlier studies. Associated with this issue is the question, whether the family Cebidae is monophyletic, more precisely whether *Saimiri*, *Aotus* and *Cebus* have to be included therein. Moreover, various contradicting phylogenetic affiliations among *Saimiri*, *Aotus* and *Cebus* were proposed by molecular and morphological studies, but never tested by SINE analyses. Furthermore, Ray et al. (2005) did not include *Callimico*, *Leontopithecus*, *Cebus*, *Brachyteles*, *Cacajao* and *Chiropotes* in their study. Hence, a complete phylogeny of Callithrichinae as well as Atelidae and Pitheciidae is still lacking.

This thesis includes all New World monkey genera combined with a relatively large dataset (128 SINE insertions were analysed), in contrast to earlier SINE based studies. In general, all relationships as depicted in earlier studies were confirmed and now a common origin of New World monkeys can be regarded as settled. Moreover, this thesis provides strong conclusive evidence for the unity of Pitheciidae and Atelidae and most importantly for a monophyletic origin of Cebidae, including the genera *Saimiri*, *Aotus* and *Cebus*. The monophyly of Atelidae is also confirmed, although Ray et al. (2005) provided already significant evidence for that. Hence, the diversification of New World monkeys into three different families Cebidae, Atelidae and Pitheciidae (Goodman et al., 1998; Opazo et al., 2006) is strongly supported. The branching order of these three families is now elucidated, with Pitheciidae as the most basal family, resulting in a grouping of Atelidae and Cebidae.

Within the Cebidae the phylogenetic positions of *Saimiri*, *Aotus* and *Cebus* could not be solved. Three independent SINE insertions were found, each of them showing different relationships among these three genera: either *Saimiri* and *Cebus*, *Saimiri* and *Aotus* or *Aotus* and *Cebus* cluster together. These contradicting results could be caused by incomplete lineage sorting of ancestral polymorphic

insertions. This would be also supported by studies (e.g. Schrago, 2007) which estimated a relative short time period in which these three genera and Callithrichinae diverged from each other. Moreover, this assumption is further supported by a relatively large number of detected autapomorphic SINE insertions, at least in *Saimiri* and *Aotus*. Alltogether these ascertainments could explain why numerous molecular studies proposed different phylogenetic affiliations of those three genera within Cebidae and morphological studies even in the whole infraorder.

In contrast to the ambiguous relationships of *Saimiri*, *Cebus* and *Aotus* among cebids, this thesis provides clear evidence for the branching order among Callithrichinae genera. The monophyletic origin of all callithrichines was confirmed and is in agreement with other molecular studies. The positioning of *Callimico* within callitrichines was a challenging issue for a long time but at least molecular studies indicated a sister grouping to *Callithrix*. This sister-relationship is conclusively supported herein by seven SINE insertions. Moreover, this thesis shows evidence for a close affiliation of *Leontopithecus*, *Callimico* and *Callithrix* to the exclusion of *Saguinus*. Therefore, earlier classifications with *Saguinus* and *Leontopithecus* as sister genera can be rejected (e.g. Goodman et al., 1998).

Within the Atelidae, *Alouatta* could clearly be separated from *Ateles*, *Lagothrix* and *Brachyteles*, which supports molecular as well as morphological analyses (e.g. Ford, 1986; Goodman et al., 1998). The splitting of the family into the two subfamilies Atelinae and Alouatinae seems thus to be appropriate. Unfortunately no further SINE insertions could be found resolving the remaining relationships within Atelinae and Pitheciidae.

In the frame of this New World monkey study, we identified an Alu insertion, distinguishing between *Saimiri sciureus* and *S. b. boliviensis*. Squirrel monkeys are common in zoos and widely used in biomedical research. This molecular cladistic marker may help to improve captive breeding management, since an exact species identification based on morphological characteristics and other molecular markers is delicate. Moreover, the presence/absence pattern among the analyzed *S. b. peruviensis* specimen indicates that this study population, and therefore most of European captive animals of this subspecies, might have originated from a natural hybrid zone in the margins of the Ucayali River in the Peruvian Amazonia. This molecular marker can easily improve the breeding management of captive populations, which is of great interest for conservation, but also for biomedical research, since both species differ in critical biological parameters, as e.g. the susceptibility to diseases.

6.1.2 Colobine monkeys

A comprehensive mobile element based phylogeny of colobine monkeys is still missing. Xing et al. (2005) performed a first attempt, which addressed Cercopithecidae phylogeny in general and therefore included just four (five) from actual ten different colobine genera. Crucial questions concerning a monophyletic origin of Asian and African colobines as well as the affiliations of the langurs *Trachypithecus* and *Semnopithecus* and the position of *Presbytis* within Asian colobines could not be answered. Moreover, the unity of the odd-nosed monkeys was merely indicated. Within this thesis these decisive long-standing questions could be definitely answered with surprising and unexpected results.

Monophyly of African and Asian colobines was not unquestioned. Although most morphological and molecular studies (e.g. Zhang and Ryder, 1998; Page et al., 1999; Ting, 2008) clearly proposed a monophyletic origin of Asian and African colobines, others (e.g. Groves, 1989; Peng et al., 1993; Jablonski, 1998) considered a paraphyly of Asian or a polyphyly of African colobines. This thesis strongly provides reliable evidence for paraphyly of African colobines and a monophyletic origin of Asian colobines. More precisely, this result includes three different Alu insertions, shared by all Asian colobines and the African genus *Piliocolobus*. However, to underpin or disclaim this finding mitochondrial and nuclear markers were incorporated as well. Mitochondrial data clearly supported a monophyletic origin of Asian and African colobines respectively, whereas nuclear data indicate paraphyly of African genera. Inadequate data, homoplasy or incomplete lineage sorting may have affected the tree topologies to some degree (Funk and Omland, 2003). However, these factors provide no sufficient explanation for the observed incongruences. Homoplasy and incomplete lineage sorting for Alu insertions is unlikely. Hence, it seems most likely that this pattern was caused by ancestral hybridizations, which is also supported by socioecological characteristics, morphology and distribution. In contrast to most other primate females, those of *Piliocolobus* tend to leave their natal groups (Newton and Dunbar, 1994). Furthermore, *Colobus* males are in average larger than *Pilicolobus* males and both genera occur over wide ranges in sympatry (Oates et al., 1994).

This thesis adduces further evidence for the monophyly of Asian colobines, whereas the evolution within Asian colobines seems to be indeed complex. Phylogenetic affiliations within Asian colobines and whether langurs and odd-nosed monkeys form monophyletic groups were major parts of debate. Concerning odd-nosed monkeys the question remains whether they represent more than an informal group as it was proposed 2001 by Groves. Not only sequence data clearly supported their monophyletic origin (Sterner et al., 2006), but also SINE data (Xing et al., 2005). This thesis supplies conclusive evidence for their monophyly and the relationships among them. Two Alu insertions

identified *Rhinopithecus* as the most basal genus, followed by the separation of *Pygathrix*, whereas *Nasalis* and *Simias* form a sister relationship and diverged relative recently. These data are not shown and will be shortly implemented in another publication combined with comprehensive mitochondrial genome and nuclear sequence data (Roos et al., 2008, in prep).

Nuclear data and Alu insertions support a close relationship of *Trachypithecus* and *Semnopithecus*, and between *Presbytis* and the odd-nosed monkeys. In contrast, mitochondrial data result in an unresolved position of *Semnopithecus* relative to a *Presbytis*/*Trachypithecus* and an odd-nosed monkey clade. Positioning *Presbytis* among Asian colobines rely on Alu insertions. As in African colobines, ancestral hybridization seems to be the most plausible explanation for the incongruent gene trees. The close relationship of *Semnopithecus* and *Trachypithecus* is also supported by morphological characteristics. Moreover, hybridization events are in general possible due to partial sympatry (Oates et al., 1994). Accordingly, these incongruences among nuclear and mitochondrial gene trees indicate a complex evolutionary scenario, in which *Semnopithecus* males hybridized with *Trachypithecus* females leading to a fixation of the mitochondrial lineage in hybrids. By backcrossing with *Semnopithecus* males over a relative long time period (until ~2.3 ± 0.5 mya), the progenitor of *Trachypithecus* enriched solely nuclear material of *Semnopithecus*, while the mitochondrial genome remained *Trachypithecus*-like.

By combining the obtained phylogenies and estimated divergence ages with paleoenvironmental data, some biogeographic suggestions can be provided. Colobines most likely originated in Africa, which is in agreement to Stewart and Disotell (1998). This is supported by the herein confirmed paraphyly, because a putative Asian origin would require two independent migrations into Africa. Furthermore, this is also supported by the fact that most cercopithecines (except Asian macaques) are solely distributed in Africa. The progenitor of Asian colobines most likely invaded Eurasia via an emerging landbridge connection between Africa and the Arabian Peninsula in the late Miocene, sometimes after the separation from *Piliocolobus* (~10.4 ± 1 mya) and before its diversification (~8.1 ± 1.2 mya). The Hengduan Mountains in the border region of today's Burma, India and China were proposed as a possible diversification hotspot for Asian colobines (Peng et al., 1993) which is supported by the fact that most of the larger Southeast-Asian rivers (Mekong, Salween, Yangtze) emerge there. Whether the ancestor of Asian colobines took a route north or south of the Himalaya from Africa to this hotspot is putative, although colobines from the later Miocene were found on the Tibetan plateau (Delson, 1994). After successive migration events, *Trachypithecus* and *Semnopithecus* came into secondary contact and exchanged genetic material until ~2.3 ± 0.5 mya, most likely in the region of today's Bangladesh, Burma and the Northeast of India.

A further part of this thesis provides detailed insight into the genus affiliations of the species groups of *Semnopithecus* and *Trachypithecus*. According to our data, the genus *Trachypithecus* comprises three species groups (*T.* [*obscurus*], *T.* [*cristatus*] and *T.* [*francoisi*]), whereas two other species,

previously also recognized as species of *Trachypithecus* (*Trachypithecus johnii* and *Trachypithecus vetulus*), should be included *Semnopithecus*. Moreover, the mitochondrial results support a diversification of the langurs from the Indian subcontinent into three species groups: one in North India comprising solely *Semopithecus entellus* populations, another in South India including *T. johnii* and *S. entellus* from South India, and finally a third group on Sri Lanka with *T. vetulus* and Sri Lankan *S. entellus*. Furthermore, the results show substantial discrepancies among the analysed gene trees for *Trachypithecus pileatus* and give therefore a hint for an ancient hybridization event between *Semnopithecus* and *Trachypithecus*. This hybridization hypothesis is also supported by some intermediate morphological characters (Groves, 2001) and the distribution of *Trachypithecus pileatus*, which is between *Semnopithecus* and other *Trachypithecus* species groups. In a recent study (Wangchuk et al., 2008) however, all tested *T. pileatus* cluster together with other *Trachypithecus* representatives, which contradicts with our (Osterholz et al. 2008) and other findings (Karanth et al., 2008). To explain these results, Wangchuk et al. (2008) proposed a hybrid origin for the individual analysed by Karanth et al. (2008). However, the individual analysed in our study is a museum specimen and was taken from the wild. Accordingly, an artificial hybrid origin can be excluded. A possible explanation for these findings might be that mitochondrial haplotypes of both *Semnopithecus* and *Trachypithecus* ancestors are still present in the *T. pileatus* population. To definetly solve this issue, further individuals and molecular markers have to be tested. Nevertheless, if the results in this thesis can be further supported in ongoing studies, this will underpin that hybridization plays a more important role in speciation mechanisms as previously thought, at least in colobines.

6.2 Contributions to Primatology and Evolution

This thesis resolves several highly disputed phylogenetic issues in primatology, especially in the evolutionary history of New World and colobine monkeys. Together with other studies, a SINE based phylogeny for the entire primate order is nearly complete. However, there are still some questions concerning the branching among Lemuriformes families, Galagidae genera or gibbon genera, which are not settled yet. A supervised diploma thesis (Kuska, 2008) was addressed to the remaining questions in strepsirrhine phylogeny. Unfortunately no further Alu insertions were detected, which elucidate remaining questions in Galagidae and Lemuriformes phylogeny, although former studies were generally confirmed. In the course of this thesis an attempt was made to resolve gibbon phylogeny, but without noteworthy results. Several solely genus specific insertion events were found by computational and experimental approaches, which is not surprising regarding the relative short time period in which the four genera diverged from each other (Roos and Geissmann, 2001). Some questions also remained in New World monkey phylogeny, like relationships within Atelidae and

Pitheciidae, and the positions of *Saimiri*, *Aotus* and *Cebus*. Atelidae and Pitheciidae evolution could be resolved by further mobile element based analyses. To resolve relationships within both families an experimental approach would be more promising than a computational approach, because only few sequences are available so far. Furthermore, the *Callithrix* genome is lately available and could act as reference sequence for an improved oligonucleotide design. The positions of *Saimiri*, *Aotus* and *Cebus* seem to be more complex issues with important and decisive evolutionary implications. Many more efforts are required to solve respective relationships. Few insertions can be found in radiation-like diversification events because splitting times are too rapid for a fixation in a population. In this thesis contradicting results were solely found for *Saimiri*, *Aotus* and *Cebus*. It is worth mentioning that retroposons could also represent a strong indication for these types of diversification causing incomplete lineage sorting phenomena and contradicting results of different sequence-based topologies. Taking all disadvantages (e.g. homoplasy) of other systems into account it would be necessary to reconsider phylogenetic concepts, which are solely based on sequence data.

With this nearly completely resolved phylogeny of primates, a basis could be provided for other fields, like socioecology, forensic science or biomedical research. The *Saimiri* marker presented in Chapter 3 is a further example for the importance of retroposons in a valid identification of organisms. Retroposons provide an ideal alternative approach requiring a minimum of equipment and time. The identification is often the first step after receiving samples from unkown origin, the wild or zoos and hence informative retroposons are also important for conservation interests (e.g. illegal wildlife trade). Furthermore, they have proven to be useful for forensic purposes in quantifying limited DNA as well as identifying human geographic origins or gender (e.g. Hedges et al., 2003; Walker et al., 2003).

The results of this thesis underpin that retroposons are still one of the most promising tools in molecular cladistic, because they represent a reliable, nearly homoplasy free and unbiased system. The only pitfall in analysing retroposons remains the possibility of differential lineage sorting on weak supported branches.

An important achievement of this thesis is the finding that – at least in primates - hybridization occurs among genera and hence probably plays a decisive role in diversification. The data provide incentives for modifications of main evolutionary principles. Several studies proposed similar hypotheses for several years. This thesis impressively demonstrates that retroposons can contribute to unequivocally reveal phylogenetic relationships, which are otherwise not resolvable by pure nuclear sequence data. Comparing them with relationships obtained from mitochondrial sequence data can give a hint to such hybridization events. Colobine monkeys and especially their diversification in Southeast Asia form the most prominent example. Consequently, we have to reconsider concepts, where hybridization leads principally to a limited diversification, reproductive isolation or lower fitness, like it was proposed by Darwin (1859) or Mayr (1963).

Bibliography

Alberts, S. C. and Altmann, J. 2001. Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. Am. J. Primatol. 53:139-154.

Andrews, T. D., Jermiin, L. S. and Easteal, S. 1998. Accelerated evolution of cytochrome b in simian primates: adaptive evolution in concert with other mitochondrial proteins? J. Mol. Evol. 47:249-257.

Andriaholinirina, N., Fausser, J. L., Roos, C. et al. 2006. Molecular phylogeny and taxonomic revision of the sportive lemurs (*Lepilemur*, Primates). BMC Evol. Biol. 6:17.

Arnason, U., Adegoke, J. A., Bodin, K., Born, E. W., Esa, Y. B., Gullberg, A., Nilsson, M., Short, R. V., Xu, X. and Janke, A. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proc. Natl. Acad. Sci. U S A 99:8151-8156.

Arnold, M. L. and Meyer, A. 2006. Natural hybridization in primates: one evolutionary mechanism. Zoology 109:261-276.

Avise, J. C. 2000. Phylogeography - The History and Formation of Species. Harvard University Press, Cambridge, MA.

Barton, N. H. 2001. The role of hybridization in evolution. Mol. Ecol. 10:551-568.

Batzer, M. A. and Deininger, P. L. 2002. Alu repeats and human genomic diversity. Nat. Rev. Genet. 3:370-379.

Batzer, M. A., Deininger, P. L., Hellmann-Blumberg, U., Jurka, J., Labuda, D., Rubin, C. M., Schmid, C. W., Zietkiewicz, E. and Zuckerkandl, E. 1996. Standardized nomenclature for Alu repeats. J. Mol. Evol. 42:3-6.

Bennett, E. L. and Davies, A. G. 1994. The ecology of Asian colobines. Pp. 129-171 *in* A. G. Davies, and J. F. Oates, eds. Colobine Monkeys: Their Ecology, Behaviour, and Evolution. Cambridge University Press, Cambridge.

Berthier, P., Excoffier, L. and Ruedi, M. 2006. Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. Proc. Biol. Sci. 273:3101-3109.

Bicca-Marques, J. C., Prates, H. M., de Aguiar, F. R. and Jones, C. B. 2008. Survey of *Alouatta caraya*, the black-and-gold howler monkey, and *Alouatta guariba clamitans*, the brown howler monkey, in a contact zone, State of Rio Grande do Sul, Brazil: evidence for hybridization. Primates.

Bigoni, F., Houck, M. L., Ryder, O. A., Wienberg, J. and Stanyon, R. 2004. Chromosomal painting shows that *Pygathrix nemaeus* has the most absal karyotype among Asian colobines. Int. J. Primatol. 25:679-688.

Bigoni, F., Stanyon, R., Wimmer, R. and Schempp, W. 2003. Chromosome painting shows that the proboscis monkey (*Nasalis larvatus*) has a derived karyotype and is phylogenetically nested within Asian Colobines. Am. J. Primatol. 60:85-93.

Boinski, S. and Cropp, S. 1998. Disparate data sets resolve squirrel monkey (*Saimiri*) taxonomy: implications for behavioral ecology and biomedical usage. Int. J. Primatol. 20:237-256.

Boinski, S. and Newman, J. D. 1988. Preliminary observations of squirrel monkey (*Saimiri oerstedii*) vocalizations in Costa Rica. Am. J. Primatol. 14:329-343.

Boubli, J. P., da Silva, M. N. F., Amado, M. V., Hrbek, T., Pontual, F. B. and Farias, I. P. 2008. A taxonomic reassessment of *Cacajao melanocephalus* Humboldt (1811), with the description of two new species. Int. J. Primatol. 29:723-741.

Brandon-Jones, D. 1995. A revision of the Asian pied leaf monkeys (Mammalia: Cercopithecidae: Superspecies *Semnopithecus auratus*), with a description of a new subspecies. The Raffles Bulletin of Zoology 43:3-43.

Brandon-Jones, D. 1984. Colobus and leaf monkeys. Pp. 398-403 *in* D. MacDonald, ed. The Encylopedia of Mammals. Allen & Unwin, London.

Brandon-Jones, D., Eudey, A. A., Geissmann, T., Groves, C. P., Melnick, D. J., Morales, J. C., Shekelle, M. and Stewart, C. B. 2004. Asian Primate Classification. Int. J. Primatol. 25:97-164.

Britten, R. J. and Kohne, D. E. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science 161:529-540.

Brown, W. M., Prager, E. M., Wang, A. and Wilson, A. C. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. 18:225-239.

Canavez, F. C., Moreira, M. A., Ladasky, J. J., Pissinatti, A., Parham, P. and Seuanez, H. N. 1999. Molecular phylogeny of new world primates (Platyrrhini) based on beta2-microglobulin DNA sequences. Mol. Phylogenet. Evol. 12:74-82.

Cann, R. L., Stoneking, M. and Wilson, A. C. 1987. Mitochondrial DNA and human evolution. Nature 325:31-36.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540-552.

Chakraborty, D., Ramakrishnan, U., Panor, J., Mishra, C. and Sinha, A. 2007. Phylogenetic relationships and morphometric affinities of the Arunachal macaque *Macaca munzala*, a newly described primate from Arunachal Pradesh, northeastern India. Mol. Phylogenet. Evol. 44:838-849.

Chaves, R., Sampaio, I., Schneider, M. P., Schneider, H., Page, S. L. and Goodman, M. 1999. The place of *Callimico goeldii* in the Callitrichine phylogenetic tree: evidence from von Willebrand factor gene intron II sequences. Mol. Phylogenet. Evol. 13:392-404.

Chivers, D. J. and Hladik, C. M. 1980. Morphology of the gastrointestinal tract in primates: comparisons with other mammals in relation to diet. J. Morphol. 166:337-386.

Collins, A. C. 2004. Atelinae phylogenetic relationships: the trichotomy revived? Am. J. Phys. Anthropol. 124:285-296.

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

Collura, R. V. and Stewart, C. B. 1995. Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. Nature 378:485-489.

Cortes-Ortiz, L., Duda, T. F., Jr., Canales-Espinosa, D., Garcia-Orduna, F., Rodriguez-Luna, E. and Bermingham, E. 2007. Hybridization in large-bodied New World primates. Genetics 176:2421-2425.

Costello, R. K., Dickinson, C., Rosenberger, A. L., Boinski, S. and Szalay, F. S. 1993. A multidisciplinary approach to squirrel monkey (genus *Saimiri*) species taxonomy *in* W. Kimbel, and L. Martin, eds. Species, Species Concepts, and Primate Evolution. Plenum Press, New York.

Cropp, S. and Boinski, S. 2000. The Central American squirrel monkey (*Saimiri oerstedii*): introduced hybrid or endemic species? Mol. Phylogenet. Evol. 16:350-365.

Darwin, C. 1859. On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle of Life. John Murray, London.

Davenport, T. R., Stanley, W. T., Sargis, E. J., De Luca, D. W., Mpunga, N. E., Machaga, S. J. and Olson, L. E. 2006. A new genus of African monkey, *Rungwecebus*: morphology, ecology, and molecular phylogenetics. Science 312:1378-1381.

Davies, A. G. and Oates, J. F. 1994. Colobine monkeys: their ecology, behaviour, and evolution. Cambridge University Press, Cambridge.

de Jong, W. W., van Dijk, M. A., Poux, C., Kappe, G., van Rheede, T. and Madsen, O. 2003. Indels in protein-coding sequences of Euarchontoglires constrain the rooting of the eutherian tree. Mol. Phylogenet. Evol. 28:328-340.

de Lima, M. M., Sampaio, I., Vieira Rdos, S. and Schneider, H. 2007. Spider monkey, Muriqui and Woolly monkey relationships revisited. Primates 48:55-63.

Deininger, P. L. and Batzer, M. A. 1993. Evolution of retroposons. Evol. Biol. 27:157-196.

Deininger, P. L., Batzer, M. A., Hutchison, C. A., 3rd and Edgell, M. H. 1992. Master genes in mammalian repetitive DNA amplification. Trends Genet. 8:307-311.

Delson, E. 1975. Evolutionary history of the Cercopithecidae. Contrib. Primatol. 5:167-217.

Delson, E. 1994. Evolutionary history of the colobine monkeys in paleoenvironmental perspective. Pp. 45-74 *in* A. G. Davies, and J. F. Oates, eds. Colobine Monkeys: Their Ecology, Behaviour, and Evolution. Cambridge University Press, Cambridge.

Delson, E. 1992. Evolution of old world monkeys. Pp. 217-222 *in* J. S. Johns, R. D. Martin, D. Pilbeam, and S. Bunney, eds. The Cambridge Encyclopedia of Human Evolution. Cambridge University Press, Cambridge.

Diatchenko, L., Lau, Y. F., Campbell, A. P. et al. 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. U S A 93:6025-6030.

Disotell, T. R. 1994. Generic level relationships of the Papionini (Cercopithecoidea). Am. J. Phys. Anthropol. 94:47-57.

Drummond, A. J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Elliot, D. G. 1913. A Review of the Primates. American Museum of Natural History, New York.

Evans, B. J., Supriatna, J., Andayani, N. and Melnick, D. J. 2003. Diversification of Sulawesi macaque monkeys: decoupled evolution of mitochondrial and autosomal DNA. Evolution 57:1931-1946.

Farris, S., Albert, V., Källersjö, M., Lipscomb, D. and Kluge, A. 1996. Parsimony Jackknifing Outperforms Neighbor-Joining. Cladistics 12:99-124.

Felsenstein, J. 1978. Cases in Which Parsimony and Compatibility Methods Will Be Positively Misleading. Syst. Zool. 27:401-410.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17:368-376.

Felsenstein, J. 1985. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution 39:783-791.

Feng, Q., Moran, J. V., Kazazian, H. H., Jr. and Boeke, J. D. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905-916.

Fleagle, J. F. 1988. Primate Adaptation and Evolution. Academic Press, New York.

Flynn, J. J. and Nedbal, M. A. 1998. Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. Mol. Phylogenet. Evol. 9:414-426.

Ford, S. M. 1986. Systematics of the New World monkeys. Pp. 73-135 *in* D. R. Swindler, and J. Erwin, eds. Comparative Primate Biology, Vol. 1: Systematics, Evolution, and Anatomy. A. R. Liss, New York.

Funk, D. J. and Omland, K. E. 2003. Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397-423.

Geissmann, T. 2003. Vergleichende Primatologie. Springer, Berlin/Heidelberg.

Geissmann, T., Groves, C. P. and Roos, C. 2004. The Tenasserim Lutung, *Trachypithecus barbei* (Blyth, 1847) (Primates: Cercopithecidae): Description, of a live specimen, and a reassessment of phylogenetic affinities, taxonomic history, and distribution. Contrib. Zool. 73:271-282.

Gibbs, R. A., Rogers, J., Katze, M. G. et al. 2007. Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222-234.

Goodman, M., Porter, C. A., Czelusniak, J., Page, S. L., Schneider, H., Shoshani, J., Gunnell, G. and Groves, C. P. 1998. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. Mol. Phylogenet. Evol. 9:585-598.

Goodman, M., Tagle, D. A., Fitch, D. H., Bailey, W., Czelusniak, J., Koop, B. F., Benson, P. and Slightom, J. L. 1990. Primate evolution at the DNA level and a classification of hominoids. J. Mol. Evol. 30:260-266.

Goossens, B., Chikhi, L., Utami, S. S., De Ruiter, J. R. and Bruford, M. W. 2000. A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an aboreal ape. Conserv. Genet. 1:157-162.

Grant, P. R. and Grant, B. R. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296:707-711.

Grant, P. R., Grant, B. R. and Petren, K. 2005. Hybridization in the recent past. Am. Nat. 166:56-67.

Groves, C. P. 2001. Primate Taxonomy. Pages Smithsonian Press, Washington, DC.

Groves, C. P. 1989. A Theory of Human and Primate Evolution. Clarendon Press, Oxford.

Groves, C. P. 1970. The forgotten leaf-eaters and the phylogeny of Colobinae. Pp. 555-586 *in* J. R. Napier, and P. H. Napier, eds. Old World Monkeys: Ecology, Systematics, and Behaviour. Academic Press, New York.

Grubb, P., Butynski, T. M., Oates, J. F., Bearder, S. K., Disotell, T. R., Groves, C. P. and Struhsaker, T. T. 2003. Assessment of the diversity of African primates. Int. J. Primatol. 24:1301-1357.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Res. 41:95-98.

Harada, M. L., Schneider, H., Schneider, M. P., Sampaio, I., Czelusniak, J. and Goodman, M. 1995. DNA evidence on the phylogenetic systematics of New World monkeys: support for the sister-grouping of *Cebus* and *Saimiri* from two unlinked nuclear genes. Mol. Phylogenet. Evol. 4:331-349.

Hedges, D. J., Walker, J. A., Callinan, P. A., Shewale, J. G., Sinha, S. K. and Batzer, M. A. 2003. Mobile element-based assay for human gender determination. Anal. Biochem. 312:77-79.

Hellborg, L. and Ellegren, H. 2003. Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. Mol. Ecol. 12:283-291.

Hennig, W. 1950. Grundzüge einer Theorie der phylogenetischen Systematik. Deutscher Zentralverlag, Berlin.

Herke, S. W., Xing, J., Ray, D. A., Zimmerman, J. W., Cordaux, R. and Batzer, M. A. 2007. A SINE-based dichotomous key for primate identification. Gene 390:39-51.

Hershkovitz, P. 1984. Taxonomy of the squirrel monkeys, genus *Saimiri* (Cebidae, Platyrrhini): A preliminary report with description of a hitherto unnamed form. Am. J. Primatol. 6:257-281.

Hershkovitz, P. 1977. Living New World monkeys, part 1 (Platyrrhini) with an introduction to primates. Chicago University Press, Chicago.

Hewitt, G. M. 2001. Speciation, hybrid zones and phylogeography - or seeing genes in space and time. Mol. Ecol. 10:537-549.

Hill, W. C. O. 1934. A monography on the purple-faced leaf-monkeys (*Pithecus vetulus*). Ceylon J. Sci. 9:23-88.

Hill, W. C. O. 1960. Primates, Comparative Anatomy and Taxonomy. IV. Cebidae, Part A. Interscience, New York.

Hofer, H. O. 1976. Preliminary study of the comparative anatomy of the external nose of South America monkeys. Folia Primatol. 25:193-214.

Horovitz, I. and Meyer, A. 1995. Systematics of New World monkeys (Platyrrhini, Primates) based on 16S mitochondrial DNA sequences: a comparative analysis of different weighting methods in cladistic analysis. Mol. Phylogenet. Evol. 4:448-456.

Horovitz, I., Zardoya, R. and Meyer, A. 1998. Platyrrhine systematics: a simultaneous analysis of molecular and morphological data. Am. J. Phys. Anthropol. 106:261-281.

Horvath, J. E., Weisrock, D. W., Embry, S. L., Fiorentino, I., Balhoff, J. P., Kappeler, P., Wray, G. A., Willard, H. F. and Yoder, A. D. 2008. Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. Genome Res. 18:489-499.

Huelsenbeck, J. P. and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.

IUCN. 2007. 2007 Red List of Endangered Animals. IUCN, Gland, Switzerland.

Jablonski, N. G. 1998. The Evolution of the Doucs and Snub-nosed monkeys and the question of the phyletic unity of the odd-nosed colobines. Pp. 13-52 *in* N. G. Jablonski, ed. Natural History of the Doucs and Snub-nosed monkeys. World Scientific Publishing Company, New Jersey.

Jolly, C. J., Woolley-Barker, T., Beyene, S., Disotell, T. R. and Phillips-Conroy, J. E. 1997. Intergeneric hybrid baboons. Int. J. Primatol. 18:597-627.

Jones, T., Ehardt, C. L., Butynski, T. M., Davenport, T. R., Mpunga, N. E., Machaga, S. J. and De Luca, D. W. 2005. The highland mangabey *Lophocebus kipunji*: a new species of African monkey. Science 308:1161-1164.

Jorgensen, S. and Mauricio, R. 2005. Hybridization as a source of evolutionary novelty: leaf shape in a Hawaiian composite. Genetica 123:171-179.

Jurka, J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc. Natl. Acad. Sci. U S A 94:1872-1877.

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O. and Walichiewicz, J. 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110:462-467.

Kajikawa, M. and Okada, N. 2002. LINEs mobilize SINEs in the eel through a shared 3' sequence. Cell 111:433-444.

Kappeler, P. M., Rasoloarison, R. M., Razafimanantsoa, L., Walter, L. and Roos, C. 2005. Morphology, behaviour and molecular evolution of giant mouse lemurs (*Mirza* spp.) gray, 1970, with description of a new species. Primate Report 71:3-26.

Karanth, K. P., Delefosse, T., Rakotosamimanana, B., Parsons, T. J. and Yoder, A. D. 2005. Ancient DNA from giant extinct lemurs confirms single origin of Malagasy primates. Proc. Natl. Acad. Sci. U S A 102:5090-5095.

Karanth, K. P., Singh, L., Collura, R. V. and Stewart, C. B. 2008. Molecular phylogeny and biogeography of langurs and leaf monkeys of South Asia (Primates: Colobinae). Mol. Phylogenet. Evol. 46:683-694.

Katoh, K., Kuma, K., Toh, H. and Miyata, T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33:511-518.

Kay, R. 1990. The phyletic relationships of extant and fossil Pitheciinae (Platyrrhini, Anthropoidea). Pp. 175-208 *in* J. F. Fleagle, and A. L. Rosenberger, eds. The Platyrrhine Fossil Record. Academic Press, New York.

Kazazian, H. H., Jr. 2004. Mobile elements: drivers of genome evolution. Science 303:1626-1632.

Kazazian, H. H., Jr. and Moran, J. V. 1998. The impact of L1 retrotransposons on the human genome. Nat. Genet. 19:19-24.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. and Haussler, D. 2002. The human genome browser at UCSC. Genome Res. 12:996-1006.

Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J. Mol. Evol. 29:170-179.

Koblmüller, S., Duftner, N., Sefc, K. M., Aibara, M., Stipacek, M., Blanc, M., Egger, B. and Sturmbauer, C. 2007. Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika--the result of repeated introgressive hybridization. BMC Evol. Biol. 7:7.

Kriegs, J. O., Churakov, G., Kiefmann, M., Jordan, U., Brosius, J. and Schmitz, J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PLoS Biol. 4:e91.

Kriegs, J. O., Matzke, A., Churakov, G., Kuritzin, A., Mayr, G., Brosius, J. and Schmitz, J. 2007. Waves of genomic hitchhikers shed light on the evolution of gamebirds (Aves: Galliformes). BMC Evol. Biol. 7:190.

Kuhn, H.-J. 1967. Zur Systematik der Cercopithecidae. Pp. 25-46 *in* D. Starck, R. Schneider, and H.-J. Kuhn, eds. Neue Ergebnisse der Primatologie. Gustav Fischer, Stuttgart.

Kuryshev, V. Y., Skryabin, B. V., Kremerskothen, J., Jurka, J. and Brosius, J. 2001. Birth of a gene: locus of neuronal BC200 snmRNA in three prosimians and human BC200 pseudogenes as archives of change in the Anthropoidea lineage. J. Mol. Biol. 309:1049-1066.

Kuska, B. 2008. Molekulare Phylogenie der Strepsirrhini. Diploma thesis, University of Göttingen.

Lander, E. S., Linton, L. M. and IHGS, C. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Lavergne, A., Catzeflis, F., Lacote, S., Barnaud, A., Bordier, M., Mercereau-Puijalon, O. and Contamin, H. 2003. Genetic analysis of the *Saimiri* breeding colony of the Pasteur Institute (French Guiana): development of

a molecular typing method using a combination of nuclear and mitochondrial DNA markers. J. Med. Primatol. 32:330-340.

Li, W. H., Gu, Z., Wang, H. and Nekrutenko, A. 2001. Evolutionary analyses of the human genome. Nature 409:847-849.

Linnaeus, C. N. 1758. Systema Naturae Per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, Cum Characteribus Differentiis, Synonymis, Locis.

Lönnerg, E. 1940. Notes on some members of the genus *Saimiri*. Ark. Zool. 32:1-18.

Luan, D. D., Korman, M. H., Jakubczak, J. L. and Eickbush, T. H. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72:595-605.

Luckett, W. P. 1980. Monophyletic or diphyletic origins of Anthropoidea and Hytricognathi: evidence of the fetal membranes. Pp. 347-368 *in* R. L. Ciochon, and A. B. Chiarelli, eds. Evolutionary Biology of the New World Monkeys and Continental Drift. Plenum Press, New York.

Mamedov, I. Z., Arzumanyan, E. S., Amosova, A. L., Lebedev, Y. B. and Sverdlov, E. D. 2005. Whole-genome experimental identification of insertion/deletion polymorphisms of interspersed repeats by a new general approach. Nucleic Acids Res. 33:e16.

Martin, R. D. 1990. Primate Origins and Evolution. A Phylogenetic Reconstruction. Chapman & Hall, London.

Matveev, V., Nishihara, H. and Okada, N. 2007. Novel SINE families from salmons validate Parahucho (Salmonidae) as a distinct genus and give evidence that SINEs can incorporate LINE-related 3'-tails of other SINEs. Mol. Biol. Evol. 24:1656-1666.

Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

Mayr, E. 1963. Animal Species and Evolution. Belknap Press, Cambridge, MA.

McCracken, K. and Sorenson, M. 2005. Is homoplasy or lineage sorting the source of incongruent mtdna and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? Syst. Biol. 54:35-55.

McDonald, D. B., Parchman, T. L., Bower, M. R., Hubert, W. A. and Rahel, F. J. 2008. An introduced and a native vertebrate hybridize to form a genetic bridge to a second native species. Proc. Natl. Acad. Sci. U S A 105:10837-10842.

Meijaard, E. and Groves, C. P. 2006. The geography of mammals and rivers in mainland Southeast Asia. Pp. 305-329 *in* S. M. Lehman, and J. G. Fleagle, eds. Primate Biogeography: Progress and Prospects. Springer, New York.

Messier, W. and Stewart, C. B. 1997. Episodic adaptive evolution of primate lysozymes. Nature 385:151-154.

Mittermeier, R. A., Konstant, W. R., Hawkins, F. et al. 2006. Lemurs of Madagascar. Conservation International, Washington, DC.

Mittermeier, R. A., Konstant, W. R. and Mast, R. B. 1994. Use of neotropical and Malagasy primate species in biomedical research. Am. J. Primatol. 34:73-80.

Mivart, S. G. J. 1873. On *Lepilemur* and *Cheirogaleus* and on the zoological rank of the Lemuroidea. Proc. Zool. Soc. Lond. 1873:484-510.

Moore, C. M., Harris, C. P. and Abee, C. R. 1990. Distribution of chromosomal polymorphisms in three subspecies of squirrel monkeys (genus *Saimiri*). Cytogenet. Cell. Genet. 53:118-122.

Mourier, T., Hansen, A. J., Willerslev, E. and Arctander, P. 2001. The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus. Mol. Biol. Evol. 18:1833-1837.

Murata, S., Takasaki, N., Saitoh, M. and Okada, N. 1993. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc. Natl. Acad. Sci. U S A 90:6995-6999.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A. and O'Brien, S. J. 2001a. Molecular phylogenetics and the origins of placental mammals. Nature 409:614-618.

Murphy, W. J., Eizirik, E., O'Brien, S. J. et al. 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348-2351.

Nadler, T., Walter, L. and Roos, C. 2005. Molecular evolution, systematics and distribution of the taxa within the silvered langur species group (*Trachypithecus* [*cristatus*]) in Southeast Asia. Zoologischer Garten (NF) 75:238-247.

Napier, J. R. 1970. Paleoecology and catarrhine evolution. Pp. 53-95 *in* J. R. Napier, and P. H. Napier, eds. Old World Monkeys: Ecology, Systematics, and Behaviour. Academic Press, New York.

Napier, J. R. and Napier, P. H. 1967. A Handbook of Living Primates. Academic Press, London.

Newton, P. N. and Dunbar, R. I. M. 1994. Colobine monkey society. Pp. 311-346 *in* A. G. Davies, and J. F. Oates, eds. Colobine Monkeys: Their Ecology, Behaviour and Evolution. Cambridge University Press, Cambridge.

Nikaido, M., Rooney, A. P. and Okada, N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. U S A 96:10261-10266.

Nishihara, H., Satta, Y., Nikaido, M., Thewissen, J. G., Stanhope, M. J. and Okada, N. 2005. A retroposon analysis of Afrotherian phylogeny. Mol. Biol. Evol. 22:1823-1833.

Nolte, A. W., Freyhof, J. and Tautz, D. 2006. When invaders meet locally adapted types: rapid moulding of hybrid zones between sculpins (*Cottus*, Pisces) in the Rhine system. Mol. Ecol. 15:1983-1993.

Nsubuga, A. M., Robbins, M. M., Roeder, A. D., Morin, P. A., Boesch, C. and Vigilant, L. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Mol. Ecol. 13:2089-2094.

Oates, J. F., Davies, A. G. and Delson, E. 1994. The diversity of living colobines. Pp. 45-73 *in* A. G. Davies, and J. F. Oates, eds. Colobine Monkeys: Their Ecology, Behaviour and Evolution. Cambridge University Press, Cambridge.

Okada, N. 1991. SINEs. Curr. Opin. Genet. Dev. 1:498-504.

Opazo, J. C., Wildman, D. E., Prychitko, T., Johnson, R. M. and Goodman, M. 2006. Phylogenetic relationships and divergence times among New World monkeys (Platyrrhini, Primates). Mol. Phylogenet. Evol. 40:274-280.

Osterholz, M., Walter, L. and Roos, C. 2008. Phylogenetic position of the langur genera *Semnopithecus* and *Trachypithecus* among Asian colobines, and genus affiliations of their species groups. BMC Evol. Biol. 8:58.

Pace, J. K., 2nd and Feschotte, C. 2007. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. Genome Res. 17:422-432.

Page, S. L., Chiu, C. and Goodman, M. 1999. Molecular phylogeny of Old World monkeys (Cercopithecidae) as inferred from gamma-globin DNA sequences. Mol. Phylogenet. Evol. 13:348-359.

Page, S. L. and Goodman, M. 2001. Catarrhine phylogeny: noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human-chimpanzee clade. Mol. Phylogenet. Evol. 18:14-25.

Pastorini, J., Forstner, M. R., Martin, R. D. and Melnick, D. J. 1998. A reexamination of the phylogenetic position of *Callimico* (primates) incorporating new mitochondrial DNA sequence data. J. Mol. Evol. 47:32-41.

Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S. and Reich, D. 2006. Genetic evidence for complex speciation of humans and chimpanzees. Nature 441:1103-1108.

Peng, Y. Z., Pan, R. L. and Jablonski, N. G. 1993. Classification and evolution of Asian colobines. Folia Primatol. 60:106-117.

Pidancier, N., Jordan, S., Luikart, G. and Taberlet, P. 2006. Evolutionary history of the genus *Capra* (Mammalia, Artiodactyla): discordance between mitochondrial DNA and Y-chromosome phylogenies. Mol. Phylogenet. Evol. 40:739-749.

Piskurek, O., Austin, C. C. and Okada, N. 2006. Sauria SINEs: Novel short interspersed retroposable elements that are widespread in reptile genomes. J. Mol. Evol. 62:630-644.

Pocock, R. I. 1936. The external characters of a female red colobus monkey (*Procolobus badius waldroni*). Proceedings of the Zoological Society of London (1935):939-944.

Pocock, R. I. 1935. The monkeys of the genera *Pithecus* (or *Presbytis*) and *Pygathrix* found to the east of the Bay of Bengal. Proceedings of the Zoological Society of London (1934):895-961.

Pocock, R. I. 1939. The Fauna of British India, Including Ceylon and Burma: Mammals. 1. Primates and Carnivores (in part), Families Felidae and Viverridae. Taylor & Francis, London.

Porter, C. A., Czelusniak, J., Schneider, H., Schneider, M. P., Sampaio, I. and Goodman, M. 1997. Sequences of the primate epsilon-globin gene: implications for systematics of the marmosets and other New World primates. Gene 205:59-71.

Porter, C. A., Sampaio, I., Schneider, H., Schneider, M. P., Czelusniak, J. and Goodman, M. 1995. Evidence on primate phylogeny from epsilon-globin gene sequences and flanking regions. J. Mol. Evol. 40:30-55.

Posada, D. and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Prychitko, T., Johnson, R. M., Wildman, D. E., Gumucio, D. and Goodman, M. 2005. The phylogenetic history of New World monkey beta globin reveals a platyrrhine beta to delta gene conversion in the atelid ancestry. Mol. Phylogenet. Evol. 35:225-234.

Raaum, R. L., Sterner, K. N., Noviello, C. M., Stewart, C. B. and Disotell, T. R. 2005. Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. J. Hum. Evol. 48:237-257.

Rambaut, A. 2008. FigTree v.1.1.2. Available from http://tree.bio.ed.ac.uk/software/figtree.

Rambaut, A. and Drummond, A. J. 2007. Tracer v1.4. Available from http://beast.bio.ed.ac.uk/Tracer.

Ray, D. A. and Batzer, M. A. 2005. Tracking Alu evolution in New World primates. BMC Evol. Biol. 5:51.

Ray, D. A., Xing, J., Hedges, D. J. et al. 2005. Alu insertion loci and platyrrhine primate phylogeny. Mol. Phylogenet. Evol. 35:117-126.

Reichenbach, H. G. L. 1862. Die vollständige Naturgeschichte der Affen. Die vollständige Naturgeschichte des In- und Auslandes. Central-Atlas für Zoologische Gärten, Dresden.

Ronquist, F. and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Roos, C. 2003. Molecular phylogeny and systematics of Vietnamese leaf monkeys. Pp. 19-23 *in* T. Nadler, F. Momberg, N. X. Dang, and N. Lormee, eds. Vietnam Primate Conservation Status Review 2002. Fauna & Flora International and Frankfurt Zoological Society, Hanoi.

Roos, C. 2004. Molecular Evolution and Systematics of Vietnamese leaf monkeys. Pp. 23-28 *in* T. Nadler, U. Streicher, and H. T. Long, eds. Conservation of Primates in Vietnam. Haki Publishing, Hanoi.

Roos, C. and Geissmann, T. 2001. Molecular phylogeny of the major hylobatid divisions. Mol. Phylogenet. Evol. 19:486-494.

Roos, C., Nadler, T. and Walter, L. 2008. Mitochondrial phylogeny, taxonomy and biogeography of the silvered langur species group (*Trachypithecus cristatus*). Mol. Phylogenet. Evol. 47:629-636.

Roos, C., Schmitz, J. and Zischler, H. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. Proc. Natl. Acad. Sci. U S A 101:10650-10654.

Roos, C., Thanh, V. N., Walter, L. and Nadler, T. 2007. Molecular systematics of Indochinese Primates. Vietn. J. Primatol. 1:41-53.

Rosenberger, A. L. 1977. *Xenothrix* and ceboid phylogeny. J. Hum. Evol. 6:461-481.

Rosenberger, A. L. 1984. Fossil New World monkeys dispute the molecular clock. J. Hum. Evol. 13:737-742.

Rosenberger, A. L. 1981. Systematics: The higher taxa. Pp. 9-27 *in* A. F. Coimbra-Filho, and R. A. Mittermeier, eds. Ecology and Behaviour of Neotropical Primates. Academia Brasileira de Ciencias, Rio de Janeiro.

Rosenblum, L. L., Supriatna, J., Hasan, M. N. and Melnick, D. J. 1997. High mitochondrial DNA diversity with little structure within and among leaf monkey populations (*Trachypithecus cristatus* and *Trachypithecus auratus*). Int. J. Primatol. 18:1005-1028.

Rowe, N. 1996. The Pictorial Guide to the Living Primates. Pogonias Press, East Hampton, New York.

Rumpler, Y., Warter, S., Hauwy, M., Fausser, J. L., Roos, C. and Zinner, D. in press.Comparing chromosomal and mitochondrial phylogenies of sportive lemurs (Genus *Lepilemur*, Primates). Chromosome Res.

Ryan, S. C. and Dugaiczyk, A. 1989. Newly arisen DNA repeats in primate phylogeny. Proc. Natl. Acad. Sci. U S A 86:9360-9364.

Ryder, O. A. 1986. Species conservation and systematics: the dilemma of subspecies. Trends Ecol. Evol. 1:9-10.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Salem, A. H., Ray, D. A., Xing, J., Callinan, P. A., Myers, J. S., Hedges, D. J., Garber, R. K., Witherspoon, D. J., Jorde, L. B. and Batzer, M. A. 2003. Alu elements and hominid phylogenetics. Proc. Natl. Acad. Sci. U S A 100:12787-12791.

Schmid, C. and Maraia, R. 1992. Transcriptional regulation and transpositional selection of active SINE sequences. Curr. Opin. Genet. Dev. 2:874-882.

Schmitz, J., Ohme, M., Suryobroto, B. and Zischler, H. 2002a. The colugo (*Cynocephalus variegatus*, Dermoptera): the primates' gliding sister? Mol. Biol. Evol. 19:2308-2312.

Schmitz, J., Ohme, M. and Zischler, H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. Genetics 157:777-784.

Schmitz, J., Ohme, M. and Zischler, H. 2002b. The complete mitochondrial sequence of *Tarsius bancanus*: evidence for an extensive nucleotide compositional plasticity of primate mitochondrial DNA. Mol. Biol. Evol. 19:544-553.

Schmitz, J., Roos, C. and Zischler, H. 2005. Primate phylogeny: molecular evidence from retroposons. Cytogenet. Genome Res. 108:26-37.

Schmitz, J. and Zischler, H. 2003. A novel family of tRNA-derived SINEs in the colugo and two new retrotransposable markers separating dermopterans from primates. Mol. Phylogenet. Evol. 28:341-349.

Schneider, H., Canavez, F. C., Sampaio, I., Moreira, M. A., Tagliaro, C. H. and Seuanez, H. N. 2001. Can molecular data place each neotropical monkey in its own branch? Chromosoma 109:515-523.

Schneider, H., Sampaio, I., Harada, M. L., Barroso, C. M., Schneider, M. P., Czelusniak, J. and Goodman, M. 1996. Molecular phylogeny of the New World monkeys (Platyrrhini, primates) based on two unlinked nuclear genes: IRBP intron 1 and epsilon-globin sequences. Am. J. Phys. Anthropol. 100:153-179.

Schneider, H., Schneider, M. P., Sampaio, I., Harada, M. L., Stanhope, M., Czelusniak, J. and Goodman, M. 1993. Molecular phylogeny of the New World monkeys (Platyrrhini, primates). Mol. Phylogenet. Evol. 2:225-242.

Schrago, C. G. 2007. On the time scale of New World primate diversification. Am. J. Phys. Anthropol. 132:344-354.

Schreiber, A., Wang, M. and Kaumanns, W. 1998. Captive Breeding of Squirrel Monkeys, *Saimiri sciureus* and *Saimiri boliviensis*: The Problem of Hybrid Groups. Zoo Biology 17:95-109.

Seehausen, O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19:198-207.

Sen, S. K., Han, K., Wang, J., Lee, J., Wang, H., Callinan, P. A., Dyer, M., Cordaux, R., Liang, P. and Batzer, M. A. 2006. Human genomic deletions mediated by recombination between Alu elements. Am. J. Hum. Genet. 79:41-53.

Shedlock, A. M. and Okada, N. 2000. SINE insertions: powerful tools for molecular systematics. Bioessays 22:148-160.

Shedlock, A. M., Takahashi, K. and Okada, N. 2004. SINEs of speciation: tracking lineages with retroposons. Trends Ecol. Evol. 19:545-553.

Shimamura, M., Yasue, H., Ohshima, K., Abe, H., Kato, H., Kishiro, T., Goto, M., Munechika, I. and Okada, N. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature 388:666-670.

Shimodaira, H. and Hasegawa, M. 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol. Biol. Evol. 16:1114-1116.

Shoshani, J., Groves, C. P., Simons, E. L. and Gunnell, G. F. 1996. Primate phylogeny: morphological vs. molecular results. Mol. Phylogenet. Evol. 5:102-154.

Silva, B. T. F., Sampaio, M. I. C., Schneider, H., Schneider, M. P. C., Montoya, E., Encarnacion, F., Callegari-Jacques, S. M. and Salzano, F. M. 1993. Protein electrophoretic variability in *Saimiri* and the question of its species status. Am. J. Primatol. 29:183-193.

Silva, B. T. F., Sampaio, M. I. C., Schneider, H., Schneider, M. P. C., Montoya, E., Encarnacion, F. and Salzano, F. M. 1992. Natural hybridization between *Saimiri* taxa in the Peruvian Amazonia. Primates 33:107-113.

Singer, S. S., Schmitz, J., Schwiegk, C. and Zischler, H. 2003. Molecular cladistic markers in New World monkey phylogeny (Platyrrhini, Primates). Mol. Phylogenet. Evol. 26:490-501.

Sinha, A., Datta, A., Madhusudan, M. D. and Mishra, C. 2005. *Macaca munzala*: a new species from western Arunachal Pradesh, northeastern India. Int. J. Primatol. 26:977-989.

Smit, A. F. and Riggs, A. D. 1995. MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. Nucleic Acids Res. 23:98-102.

Steiper, M. E. and Ruvolo, M. 2003. New World monkey phylogeny based on X-linked G6PD DNA sequences. Mol. Phylogenet. Evol. 27:121-130.

Steiper, M. E. and Young, N. M. 2006. Primate molecular divergence dates. Mol. Phylogenet. Evol. 41:384-394.

Sterner, K. N., Raaum, R. L., Zhang, Y. P., Stewart, C. B. and Disotell, T. R. 2006. Mitochondrial data support an odd-nosed colobine clade. Mol. Phylogenet. Evol. 40:1-7.

Stewart, C. B. and Disotell, T. R. 1998. Primate evolution - in and out of Africa. Curr. Biol. 8:R582-588.

Strasser, E. and Delson, E. 1987. Cladistic analysis of cercopithecid relationships. J. Hum. Evol. 16:81-99.

Strimmer, K. and von Haeseler, A. 1996. Quartet Puzzling: A Quartet Maximum-Likelihood Method for Reconstructing Tree Topologies. Mol. Biol. Evol. 13:964-969.

Swofford, D. L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods) 4.0b10. Sinauer Associates, Sunderland.

Szalay, F. S. and Delson, E. 1979. Evolutionary History of the Primates. Academic Press, New York.

Taberlet, P., Waits, L. P. and Luikart, G. 1999. Noninvasive genetic sampling: look before you leap. Trends Ecol. Evol. 14:323-327.

Takahashi, K., Terai, Y., Nishida, M. and Okada, N. 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Mol. Biol. Evol. 18:2057-2066.

Takahashi, K., Terai, Y., Nishida, M. and Okada, N. 1998. A novel family of short interspersed repetitive elements (SINEs) from cichlids: the patterns of insertion of SINEs at orthologous loci support the proposed monophyly of four major groups of cichlid fishes in Lake Tanganyika. Mol. Biol. Evol. 15:391-407.

Tavaré, S., Marshall, C. R., Will, O., Soligo, C. and Martin, R. D. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416:726-729.

Thalmann, O., Fischer, A., Lankester, F., Paabo, S. and Vigilant, L. 2007. The complex evolutionary history of gorillas: insights from genomic data. Mol. Biol. Evol. 24:146-158.

Thalmann, O., Hebler, J., Poinar, H. N., Paabo, S. and Vigilant, L. 2004. Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. Mol. Ecol. 13:321-335.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.

Thorington, R. W. 1985. The taxonomy and distribution of squirrel monkeys (*Saimiri*) *in* L. A. Rosenblum, and C. L. Coe, eds. Handbook of squirrel monkey research. Plenum Publishing Corp, New York.

Ting, N. 2008. Mitochondrial relationships and divergence dates of the African colobines: evidence of Miocene origins for the living colobus monkeys. J. Hum. Evol. 55:312-325.

Ting, N., Tosi, A. J., Li, Y., Zhang, Y. P. and Disotell, T. R. 2008. Phylogenetic incongruence between nuclear and mitochondrial markers in the Asian colobines and the evolution of the langurs and leaf monkeys. Mol. Phylogenet. Evol. 46:466-474.

Tosi, A. J., Morales, J. C. and Melnick, D. J. 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. Mol. Phylogenet. Evol. 17:133-144.

Tosi, A. J., Morales, J. C. and Melnick, D. J. 2002. Y-chromosome and mitochondrial markers in *Macaca fascicularis* indicate introgression with Indochinese *M. mulatta* and a biogeographic barrier in the Isthmus of Kra. Int. J. Primatol. 23:161-178.

van de Lagemaat, L. N., Gagnier, L., Medstrand, P. and Mager, D. L. 2005. Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. Genome Res. 15:1243-1249.

VandeBerg, J. L. and Williams-Blangero, S. 1997. Advantages and limitations of nonhuman primates as models in genetic research on complex diseases. J. Med. Primatol. 26:113-119.

VandeBerg, J. L., Williams-Blangero, S., Moore, C. M., Cheng, M. L. and Abee, C. R. 1990. Genetic relationships among three squirrel monkey types: implications for taxonomy, biomedical research, and captive breeding. Am. J. Primatol. 22:101-111.

Vermeer, J. 1996. EEP SURVEY OF Saimiri., Apeldoorn, Apenheul Primate Park.

von Dornum, M. and Ruvolo, M. 1999. Phylogenetic relationships of the New World monkeys (Primates, platyrrhini) based on nuclear G6PD DNA sequences. Mol. Phylogenet. Evol. 11:459-476.

von Pusch, A. 1942. Die Arten der Gattung *Cebus*. Z. Säugertierk. 16:183-237.

Waddell, P. J., Kishino, H. and Ota, R. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Inform. 12:141-154.

Walker, J. A., Kilroy, G. E., Xing, J., Shewale, J., Sinha, S. K. and Batzer, M. A. 2003. Human DNA quantitation using Alu element-based polymerase chain reaction. Anal. Biochem. 315:122-128.

Wallace, A. 1858. On the Tendency of Varieties to Depart Indefinitely from the Original Type. Proceedings of the Linnean Society 3:53-62.

Wallace, R. B., Gomez, H., Felton, A. and Felton, A. M. 2006. On a new species of titi monkey, genus *Callicebus Thomas* (Primates, Pitheciidae), from western Bolivia with preliminary notes on distribution and abundance. Primate Conserv. 20:29-39.

Wang, W., Forstner, M. R. J., Zhang, Y. P., Liu, Z. M., Wei, Y., Huang, H. Q., Hu, H. G., Xie, Y. X., Wu, D. H. and Melnick, D. J. 1997. Phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences. Int. J. Primatol. 18.

Warter, S., Hauwy, M., Dutrillaux, B. and Rumpler, Y. 2005. Application of molecular cytogenetics for chromosomal evolution of the Lemuriformes (Prosimians). Cytogenet. Genome Res. 108:197-203.

Weitzel, V., Yang, C. M. and Groves, C. P. 1988. A catalogue of primates in the Singapore Zoological Reference Collection. The Raffles Bulletin of Zoology 36:1-166.

Whitfield, L. S., Lovell-Badge, R. and Goodfellow, P. N. 1993. Rapid sequence evolution of the mammalian sex-determining gene SRY. Nature 364:713-715.

Whittaker, D. J., Ting, N. and Melnick, D. J. 2006. Molecular phylogenetic affinities of the simakobu monkey (*Simias concolor*). Mol. Phylogenet. Evol. 39:887-892.

Whybrow, P. J. 1984. Geological and faunal evidence for mammal 'migrations' between Asia and Africa during the Miocene. Cour. Forsch. Inst. Senckenberg 69:189-198.

Whybrow, P. J. 1992. Land movements and species dispersal *in* S. Jones, R. E. Martin, and D. Pilbeam, eds. The Cambridge Encyclopedia of Human Evolution. Cambridge University Press, Cambridge.

Wyner, Y. M., Johnson, S. E., Stumpf, R. M. and Desalle, R. 2002. Genetic assessment of a white-collared x red-fronted lemur hybrid zone at Andringitra, Madagascar. Am. J. Primatol. 57:51-66.

Xing, J., Wang, H., Han, K., Ray, D. A., Huang, C. H., Chemnick, L. G., Stewart, C. B., Disotell, T. R., Ryder, O. A. and Batzer, M. A. 2005. A mobile element based phylogeny of Old World monkeys. Mol. Phylogenet. Evol. 37:872-880.

Xing, J., Wang, H., Zhang, Y., Ray, D. A., Tosi, A. J., Disotell, T. R. and Batzer, M. A. 2007. A mobile element-based evolutionary history of guenons (tribe Cercopithecini). BMC Biol. 5:5.

Yoder, A. D., Burns, M. M., Zehr, S., Delefosse, T., Veron, G., Goodman, S. M. and Flynn, J. J. 2003. Single origin of Malagasy Carnivora from an African ancestor. Nature 421:734-737.

Yoder, A. D., Cartmill, M., Ruvolo, M., Smith, K. and Vilgalys, R. 1996. Ancient single origin for Malagasy primates. Proc. Natl. Acad. Sci. U S A 93:5122-5126.

Zhang, Y. P. and Ryder, O. A. 1998. Mitochondrial Cytochrome b Gene Sequences of Old World Monkeys: With Special Reference on Evolution of Asian Colobines. Primates 39:39-49.

Zischler, H., Geisert, H., von Haeseler, A. and Pääbo, S. 1995. A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. Nature 378:489-492.

Zwickl, D. J. 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion. PhD thesis, The University of Texas, Austin.

Supplementary

Chapter 2

Supplementary Table 2.1: Accession number of screened BAC clones

Callithrix jacchus	*Saimiri boliviensis*	*Aotus nancymaae*	*Ateles geoffroyi*	*Callicebus moloch*
AC145530	AC146467	AC151892	AC186273	AC144655
AC150217	AC151884	AC153313	AC188244	AC146285
AC150604	AC151887	AC166195	AC188245	AC151890
AC150613	AC153429	AC168952	AC188246	AC186942
AC150814	AC161093	AC169135	AC188247	AC185388
AC151019	AC183844	AC171393	AC188248	AC186116
AC151042	AC187665	AC188295	AC188249	AC187183
AC151378	AC188233	AC188338	AC188250	AC187439
AC160512	AC188238	AC188344	AC188251	AC187643
AC166001	AC188239	AC188493	AC188252	AC187691
AC166519	AC188240	AC189157	AC188253	AC187954
AC167294	AC188241	AC190001	AC188254	AC188269
AC168185	AC188242	AC190441	AC188256	AC188270
AC182520	AC188243	AC193070	AC188257	AC188271
AC188221	AC198552	AC193272	AC188258	AC188273
AC188222	AC199259	AC195014	AC188259	AC188276
AC188223	AC199260	AC196902	AC188260	AC188359
AC188224	AC202236	AC202244	AC188261	AC188361
AC188227	AC207836	AC203482	AC199280	AC188363
AC212546	AC208270		AC199282	AC190019
	AC209604		AC200010	AC191895
			AC200012	AC196347
			AC200013	AC196348
			AC200397	AC198832
			AC201730	AC199677
			AC202262	AC199678
			AC202263	AC199679
			AC202264	AC200000
			AC203135	AC200393
			AC203518	AC200395
				AC203508
				AC203731
				AC205235

Supplementary Table 2.2: Presence/absence pattern and further details on studied SINE loci.

No.	Locus ID	Name	Source (Hyb. or BAC)	Forward Primer	Reverse Primer	AT	CP	CG	LC	SI	CA	SS	AA	BA	LL	AF	AC	CC	CAS	PP	CM	Human Ortholog Chromosome	Human Ortholog Start
1	5263	CebAte1	Hybridization Lagothrix Saimiri	CTTTAGATGGTGGATTGTGTTC	GTGTCATTCTCCCATCTTTAC	58	1			1					1	1			0			4	99813129
2	4797	Pitheciidae1	AC151884 151892 188246 151890	GATCAATTGTTATTTTCTTATTGGGC	CGCTTTGTACCTTAGATGCTA	58	0					0	0			0		1	1	1	1	8	11646425
3	5003	Pitheciidae2	AC166519 166195 203135 200393	TTGGTGAGATGGAATCTGATC	GTGGTCAGGGTCTTGAGG	58	0						0			0		1	1	1	1	20	3786998
4	5305	Pitheciidae3	AC193272 207836 186116	GACCTGCTCCGTTTCTGAC	CTCTCTGTTTGGGCTAATAAC	58	0					0	0		0			1	1	1	1	21	39644719
5	5367	Pitheciidae4	AC188223 188252 188269	CCTTCAGGATTATCATTCATTTG	CCTTGTTCAAGTGACAGCATC	58	0			0				0	0	0		1	1	1	1	17	58986618
6	4999	Atelidae1	AC199259 203482 203518 203731	GCTGATGGACATTTCATGGACA	GCCTTTTATGAGGTTAAATTTTC	58	0					0	0		1	1	1		0		0	?	?
7	5015	Atelidae2	AC166519 166195 203135 200393	ATGCTGGGCAGGTCAGTTG	CTAAAATTCTACCACTTGGGAAC	58	0					0	0		1	1	1		0		0	20	3796900
8	5361	Atelidae3	AC188223 188252 188269	CACTTCTTACTAGATGAAATGAC	AGAGTAAAACAATAGCAGTT	58	0			0					1	1	1				0	17	58943802
9	5365	Atelidae4	AC188223 188252 188269	GCCCAGATGAACAGGTGAG	CTGTCAATCTTGTAGCCTGTG	58	0			0					1	1	1				0	17	58975491
10	5369	Atelidae5	AC188223 188252 188269	GCAGATTCCAGACTCAGAAG	CACCAAGCTGGTCACAATTC	58	0			0					1	1	1				0	? (17)	?
11	5017	Atelinae1	AC203708 200011	TGCCAACCTGTGTCAAACTC	CTGAGCAATGGAGCTTTCT	58	0					0	0	1	1	1	0		0			8	42616260
12	5375	Atelinae2	AC188227 188295 188251 188276	TAAATCTGGTTCTTTGCTCTCC	GGTCATGTGCTCAGAAATGC	58	0			0			0	1	1	1	0				0	1	27090419
13	5865	Atelinae3	Ray Atel_Alu1D_159	CCAAGACAGACCCAATATCCA	CCAGCCAGCACTTTGNATCA	54	0			0		0	0	1	1	1	0			0	0	1	27354917
14	5867	Atelinae4	Ray Atel_Alu1D_32	TGAAACCTCATTGCAGTCCA	TGCTCCANTTGCAAGTATGC	57	0			0		0	0	?	1	1	0			0	0	14	37808207
15	5869	Atelinae5	Ray Lago_Alu1D_i2C_156	TTGAACAAAACCATCAGAGCA	CAGTNGAATRCCTTTTGCATCA	52	0			0		0	0	?	1	1	0			?	0	12	43212740
16	3814	Cebidae1	AC188226 188242 188294 188274	CTTCTGCTGAGGCAGTGTC	ATGATCAGGTCTATGCTGTG	58	1		1		1	1	1			0			0		0	5	149208819
17	4791	Cebidae2	AC151884 151892 188246 151890	CAACAGTGTGCTAGGCAGC	CACCATAACTTACAGAGAGGC	58	1		1		1	1	1			0					0	8	11665537
18	4809	Cebidae3	AC151884 151892 188246 151890	TAAGAACTATGGCAAGATATGG	TCCCACTCATCCTCAAACAG	58	1		1		1	1	1			0					0	8	11607395
19	5009	Cebidae4	AC166519 166195 203135 200393	CACCTTGATTTGCAGAAAAGC	CCCAGCCACACTGATGATC	58	1		1		1	1	1			0					0	20	3742303
20	5295	Cebidae5	AC193272 207836 186116	GGTACTGACCAATTACTTGAG	ACAGTGTGTACTTCTGACTTG	58	1				1	1	1		0				0		0	21	39639959
21	5297	Cebidae6	AC193272 207836 186116	ACAAAGTGAGACCCTGTCTC	AAGGTTAGGACTCTGGTGAC	58	1				1	1	1		0				0		0	21	39718257
22	5299	Cebidae7	AC193272 207836 186116	TCGGTCTCCTGTGAAGTAAC	CCTACAAATACACAATTCCTTATC	58	1	1			1	1	1		0				0		0	21	39720105
23	5289	CeAo1	AC193272 207836 186116	GAAGCAAAGTAGACTGAAATAT	CTTAATTTCAAGTGGTTCAATG	58	0	0			1	0	1		0				0		0	? (21)	?
24	5829	AoSa1	Ray Aotus_G03_2	CGTANTTACAGGGTTAGTCYGAGG	GTCTCAGCCCATCTCTTGCTGT	55	0			0	0	1	1		0	0	0			0	0	22	25050860
25	5851	CeSa1	Ray AC148188_1_5	CAGCTCYRGAAACTAGAGCACA	GCTCAGAGGAGGGCTAYAAA	55	0			0	1	1	0		0	0	0			0	0	7	115297146
26	4369	Callithrichinae1	AC188222 188239 188259 188270	TGTTGTAGGGCCTGGACTG	AATGCAGAAATATCCATGAAC	58	1		1	1	0	0	0			0			0		0	6	35516037
27	4389	Callithrichinae2	AC188222 188239 188270	CTCTGTTACTGCCTGCTAAG	CACAAGATTGAAAGCCTGGAG	58	1		1	1	0	0	0						0		0	6	35482941
28	4479	Callithrichinae3	AC188224 188254 188271	TTGACTCTGGAGTTCCAACC	ATTGAAGGCAGCATGTAGAAG	58	1		1	1	0	0	0			0	0		0		0	4	23426859
29	4513	Callithrichinae4	AC188221 188253 188233	CTTCTGAGGCCAGATTATTTC	TGCTCTTTCATGTTCCCTTTC	58	1		1	1	0	0	0			0	0		0			1	55337403
30	4847	Callithrichinae5	AC182520 188493 198832	ACCTGCATTTATGACTCTTAAC	CATAGCAGGTACTTCCTGTC	58	1	1	1	1	0	0	0								0	6	108686795
31	4447	CaCaLe1	AC150613 190441 196347	GGAATTTTGTTTCAGCTTTCTCTC	TATCACCTCATGTGACTGCAC	58	1		1	0			0				0		0		0	6	108836825
32	4481	CaCaLe2	AC188224 188254 188271	AACTTCAGTCATAAATCTCAAAC	TGAATGATTGAAGATTAGCTTTC	58	1	1	1	0		0	0			0	0		0		0	4	23443044
33	5827	CaCaLe3	Ray AC146674_1_4	GCCAGGTAGCCAACATCAGT	CCAACARGGAAATGTGGTCT	55	1	1	1	0		0	?		0	0	0				0	7	117009768
34	4387	CaCa1	AC188222 188239 188270	TCTCAGTCACAAAGCTAGAGG	GAGGGAGCGATCAACATTCAA	58	1	1	0	0		0									0	6	35484780
35	4393	CaCa2	AC188222 188239 188270	CCTTGTCTTGTCTGAGGGTC	AGCCAGAAGCCCTCTGATC	58	1	1	0	0		0									0	6	35481188
36	4503	CaCa3	AC188221 188253 188233	CTAGATCTACAGTTCTCTGTG	TGGTGACTGGTTATGAGTTAC	58	1	1	0	0		0				0						1	55314374
37	4505	CaCa4	AC188221 188253 188233	ATAGTGCCGAAAGCTGCCAC	CCTAATTTCATGAGTCCAGAC	58	1	1	0	0		0				0						1	55318310
38	4509	CaCa5	AC188221 188253 188233	TTCCAGCAAGTAACTACTCAC	GCTATTTGCTCATGTTTAGTAC	58	1	1	0			0				0						1	55333738
39	5819	CaCa6	Ray Marm_Alu1D_2D_26	GAGGGAGGTGCTGAGGAGTA	AGCAGCTGTGTAAACCCAAAA	59	1	1	0	0		0	0		0	0	0			0	0	7	34586975
40	5823	CaCa7	Ray AC146661_3_1	AATGAGGGATTYAGGGAAGG	GGGACAAGGGGRAGGAGAGT	58	1	1	0	0		0	0		0	0	?			?	0	7	116776646
New World Monkey markers																							
41	5153	NWM1	Hybridization Pithecia/Chiropotes Ateles	TGG CAT AGG CAT CAC ATA TTG	CTT TGG TGC TCT TTA GCC AG	58	1			1						1			1	1		7	139524773
42	5157	NWM2	Hybridization Pithecia/Chiropotes Ateles	GTC TTC TTA TGC TAC CAT ACA C	TAC CAT GTT ATT GGC TTT ATG C	58	1			1						1			1	1		22	33386532
43	5159	NWM3	Hybridization Pithecia/Chiropotes Ateles	ACA GCA TGT CAG CCA GAG TG	AAC AGC TGC TAC CTC AAC TC	58	1			1						1			1	1		12	6837525
44	5161	NWM4	Hybridization Pithecia/Chiropotes Ateles	CTC TTA TCC TGC TGG TCT TG	CTT AGA AGG CTC TTA TCC CAG	58	1			1						1			1	1		1	152982198
45	5171	NWM5	Hybridization Pithecia/Chiropotes Ateles	ATT GCC TGC ATA CAA CAT CTG	GAG CGA ATT TGA ATG TGT TGC	58				1						1			1	1		6	151023300
46	5173	NWM6	Hybridization Pithecia/Chiropotes Ateles	TAC AAA ACA CCT CCC AGG AG	ATC AAT GCC AGA GAG ATT AGC	58	1									1			1	1		3	40537415
47	5235	NWM7	Hybridization Lagothrix Saimiri	AGC CCT TCA GTT ATG AAG CT	CAG CAG TAC ATG TGG ATT TAC	58	1			1				1	1				1			12	102907389
48	5237	NWM8	Hybridization Lagothrix Saimiri	CAT GGC GTA GGA CAA GCT C	CAA ATC CTG CAC AGG GAA CC	58	1			1					1	1			1			16	68304701

No.	Locus ID	Name	Source (Hyb. or BAC)	Forward Primer	Reverse Primer	AT	CP	CG	LC	SI	CA	SS	AA	BA	LL	AF	AC	CC	CAS	PP	CM	Human Orthology Chromosome	Start
49	5241	NWM9	Hybridization Lagothrix Saimiri	GAG CAT GTG TTT TGA CTG GC	GAC TTT AGG CAG CCC AAT TC	58	1			1				1	1				1			12	87736926
50	5243	NWM10	Hybridization Lagothrix Saimiri	AGG TCA AGA AGC CCT GAT TC	CCT CCT CAA ACC TGG ACT TC	58	1			1				1	1				1			1	199453452
51	5249	NWM11	Hybridization Lagothrix Saimiri	GGA AGG AAT ACA GAA GAG GC	CTC CAT CTT ACC TGA GTG AC	58	1							1	1				1			12	25298561
52	5253	NWM12	Hybridization Lagothrix Saimiri	AAC TGG GTA GAG GAG TAA GTG	ATA GAG GGA GCA GTT AAA GAC	58	1			1				1	1				1			3	182071324
53	5257	NWM13	Hybridization Lagothrix Saimiri	AGT TGT GAC TAC CTA AGA ACC	CTC TAA TCA TGA GGT AGC ATG	58	1			1				1	1				1			3	186496904
54	5259	NWM14	Hybridization Lagothrix Saimiri	GAT GCA CTG AAT TAG AAT GCT C	GGT TGT ATC TGT AGA GTA ATG G	58	1			1				1	1				1			14	80738994
55	5261	NWM15	Hybridization Lagothrix Saimiri	CAG TTA CAG ACA TAT CCA GAA G	ACA CGA AAA CAT GAC ACT GAC	58	1			1				1	1				1			11	122602533
56	5265	NWM16	Hybridization Lagothrix Saimiri	AAC TGA TCA TGG CTT TGA GAC	ATA TCA CCA GGT AAC AGA AGC	58	1			1				1	1				1			8	28265475
57	5267	NWM17	Hybridization Lagothrix Saimiri	CTC ACC TCT TCT GCT TTC TTC	GTA TTC TGT ATG CTG AAG GAG	58	1			1				1	1				1			13	23597170
58	5271	NWM18	Hybridization Lagothrix Saimiri	CTG TGG CCA CAG TCT TAA TTG	TGT GTT GGT GGA TTG AAT TGT G	58	1							1	1				1			X	79954707
Genus specific markers																							
59	5147	Pithecia1	Hybridization Pithecia/Chiropotes Ateles	ACA GAG GCA ATC ATG CTT TAT C	GAT AAA CCA TGA AAG CAG CTC	58	0			0				0			0	0	1	0		18	70285374
60	5157	Pithecia2	Hybridization Pithecia/Chiropotes Ateles	GTC TTC TTA TGC TAC CAT ACA C	TAC CAT GTT ATT GGC TTT ATG C	58	0			0							0	0	1	0		22	33386532
61	5873	Pithecia3	Ray Lago_H05b	CATCTGCTGCCTTTCTATCCAG	CATCTGGGAACTCGCTATGG	50	0			0							0	0	1	0		X	14843214
62	4367	Callicebus1	AC188222 188239 188270	CTGGGATTACAGGCATGTGC	CATGAGCAAAGGCGATGCCTC	58	0			0		0		0	0		0	0	1	0		6	35518202
63	4451	Callicebus2	AC150613 190441 196347	TGG ACT ATG AGT TGA GCA GAC	TAT GCT AAT TAC TCT TCT TGG GC	58	0		0	0		0	0			0		0		1		6	108856437
64	4811	Callicebus3	AC151884 151892 188246 151890	TCC ATT ACC AGG CCA GCT AG	GAA CTT GCT GGC TTT CAT ATC	58	0	0			0	0	0		0			0	0	1		8	11606573
65	4831	Callicebus4	AC150604 193070 191895	GGA ACA TAA TAG CCT GAA AGG	GAT GGA CAA AAT GTG TAA GAA AC	58	0	0					0					0	0	1		16	61189868
66	3940	Ateles1	AC188222 188239 188260 188270	GGA GCT GTT CTC TTT CTG AG	GAG ATC AGC AGA TGA CAC TG	58	0					0		0	1					0		6	35455452
67	4841	Ateles2	AC160512 161093 188250	GTC AGC ACC TAT AAA TGA AGC	CTA GAA CCA AGT CTG GCT AAC	58	0					0	0	0	1	0						2	21133592
68	5005	Ateles3	AC166519 166195 203135 200393	CCT GGT CAC ATT TAT AAC TGT G	CCT TCC ACA GCC ATT GTA TTC	58	0			0		0	0	0	1	0						20	3787392
69	5007	Ateles4	AC166519 166195 203135 200393	GCA GTC TAA TCC TCT GTC AC	TCA TAT GGA ACA TTG CAA AGA C	58	0			0			0	?	0	1	0		0			20	3790352
70	5387	Ateles5	AC188223 188252 188269	CCATTGTCCCAATCCTCTGC	TTTCCACTGTTGTGCTGC	58	0			0			0	0	0	1	0				0	1	27165701
71	5871	Ateles6	Ray Lago_Alu1D_17a	ATGCAGCAAGAGACAGCTTG	TCCCACCAAANTAGGACTGC	56	?			?		0	?	0	0	1	0		0	0		22	30588338
72	5367	Lagothrix1	AC188223 188252 188269	CCTTCAGGATTATCATTCATTTG	CCTTGTTCAAGTGACAGCATC	58	0			?			0	0	1	0		0	0	0		17	58986618
73	5871	Lagothrix2	Ray Lago_Alu1D_17b	ATGCAGCAAGAGACAGCTTG	TCCCACCAAANTAGGACTGC	56	?			?		0	?	0	1	0		0	0	0		22	30588338
74	4459	Aotus1	AC151378 198552 188344 187641	CCA TAT CTA TTT TCT GGT TGC	GCC TAG AAG TTT CTC AGT TC	58	0			0	0	0	0	1		0		0		0		18	24103431
75	4461	Aotus2	AC151378 198552 188344 187641	CCT GAA TTC CCA GCT TAA AAT C	AAC AAA ATT ACC TCC TGG ATG G	58	0			0	0	0	0	1		0		0		0		18	24100808
76	4465	Aotus3	AC151378 198552 188344 187641	GTC TCA GGG TCT GTG ATG AC	CTT AAA TTC TGA CTG CAA TGA TC	58	0			0	0	0	0	1		0		0		0		18	24093677
77	4467	Aotus4	AC151378 198552 188344 187641	ACT GAG TTG CAC TGA TGT TAA ATC	GTA AGG AAA CAC TTA GGC ATT TAG	58	0			0	0	0	0	1		0		0		0		18	24085435
78	4621	Aotus5	AC151047 188241 189157 188258 188273	ATG CTG ATG GCT GCA AAG TC	ACT TTA AGT ACA AAA CAG CTC C	58	0			0	0	0	0	1		0		0		0		5	132066014
79	4635	Aotus6	AC150604 193070 191895	CCT GCA TGA AAT ACT GTC ATT C	ACT AAG TGC AAA CAT GTC CTC	58	0			0	0	0	0	1		0		0		0		16	61205669
80	4637	Aotus7	AC150604 193070 191895	TCT CTA TTC TAG ATC ACA GCC	CTT AGT AAT TCT ACT GGC TCA C	58	0			0	0	0	0	1		0		0		0		16	61212015
81	4849	Aotus8	AC182520 188493 198832	TTC TCC AGG TGT AGT GAA GC	TTG AGT ATA AGA CGA TTT TGC C	58	0			0	0	0	0	1		0		0		0		6	108687066
82	5853	Aotus9	Ray Aotus_1D_2D_108	AATTGGACCAAGGAAAAAGGA	GGATGNCAGGACATTCTGTTCTC	50	0			0	0	0	0	1		0		0		0	?	6	135732560
83	5855	Aotus10	Ray Aotus_Alu1D_13	AGTGGACTCCTCCATGAACAA	TYAARCCCAAATCATGTAAGG	53	0			?	0	0	0	1		0		0		0	0	1	235108348
84	5857	Aotus11	Ray Aotus_Alu1D_20	GCAAAAGCTTATGGTTTCTGAG	TAGAGCCTCCCCAAGTGCCTA	56	0			0	0	0	0	1		0		0		0	0	X	94661666
85	5859	Aotus12	Ray Aotus_Alu1D_F9	CCCAAAGCAGCAGCATAGTCTG	CACCCTAACCTGGTAGAAGAG	53	0			0	0	0	0	1		0		0		0	?	2	6969671
86	5861	Aotus13	Ray Aotus_B02	AGCAGGATTCAAGGAAGCAG	GAACAGTAAGAGAGCTGTGGAGG	56	0			0	0	0	0	1		0		0		0	0	8	60060720
87	5863	Aotus14	Ray SAlu2A_D03	CGATACTGCGATGTAGAGAGCC	GAGTTCCTCTTCTTCACAGC	57	0			0	0	0	0	1		0		0		0	0	9	96108596
88	5831	Aotus15	Ray Sag_Alu1D_132b	TGGATGTCTCACAGCTCCAA	TTGATGYCCTTCCCCAAGAC	58	0			0	0	0	0	1		0		0		0	0	7	23047208
89	4813	Cebus1	AC151884 151892 188246 151890	GCT TTC GAA ATT ACT ACT CTG C	GTG GGC TCT TAG ATT ATG AG	58	0			1	0	0		0		0				0		8	11594849
90	4849	Cebus2	AC182520 188493 198832	TTC TCC AGG TGT AGT GAA GC	TTG AGT ATA AGA CGA TTT TGC C	58	0			1	0	0		0		0				0		6	108687066
91	3896	Saimiri1	AC188241	CTT TGC CAA GTA AGA CAT TCC	GAA ATA ATC TCT TGC TGT ACC AG	58	0			0	1	0		0		0			0			5	131919489
92	3898	Saimiri2	AC188241	TTT CCA AGG CAG GCA GAG G	ATG TAG AGG TAA TCC ATT CCG	58	0			0	1	0		0		0			0			5	131919714
93	3906	Saimiri3	AC188241	AAT ACA GTG AGC AGG CAC TC	AGG GTT GCT GAG GCA CTG	58	0			0	1	0		0		0			0			5	131970950
94	3908	Saimiri4	AC188241	CAT AGT TCA TTT TAA TCA CCA GG	GTG GAT CAT TTT GCA TAG GAA C	58	0			0	1	0		0		0			0			5	131978793
95	3910	Saimiri5	AC188241	TCC CAG AGT ATG TCA TGC C	ATT ACT TCC CTG CTT CAA ACC	58	0			0	1	0		0		0			0			5	131991243
96	4373	Saimiri6	AC188222 188239 188270	GCATTGCAGAAGGGAAAAGTTG	TGGAACAACACCTGTAAGGAAC	58	0			0	1	0		0		0			0			6	35507955

																						Human Ortholog	
No.	Locus ID	Name	Source (Hyb. or BAC)	Forward Primer	Reverse Primer	AT	CP	CG	LC	SI	CA	SS	AA	BA	LL	AF	AC	CC	CAS	PP	CM	Chromosome	Start
97	4375	Saimiri7	AC188222 188239 188270	CTATTTCCAAGAGAAGCCATCC	CTACAGGCATGCACCAACTG	58	0				0	1	0						0		0	6	35508463
98	4391	Saimiri8	AC188222 188239 188270	CTCCAGGCTTTCAATCTTGTG	GGAACAGGAATACTATGCCTAG	58	0				0	1	0			0			0		0	6	35483239
99	4463	Saimiri9	AC151378 198552 188344 187641	GCA TAA TTG AGA GAT GAT ACC TC	TAA CAA GTG ACT AGA TGA AAG GC	58	0		0	0	0	1	0				0		0		0	18	24095261
100	4775	Saimiri10	AC151884 151892 188246 151890	GAA CAG AGA TTT CAG TAA ATG C	TAA TGA CAC TCC AGA GCT TAC	58	0				0	1	0			0					0	8	11719096
101	4779	Saimiri11	AC151884 151892 188246 151890	TTC CAC ACT GAT GGT CAT GTC	ATT ACA CCC GTG AAC TTA GAC	58	0				0	1	0			0					0	8	11704540
102	4781	Saimiri12	AC151884 151892 188246 151890	AAT ATA TTA TCA GAG CTA GAA GG	GTC AGA GCT CAT TTG GAT CC	58	0				0	1	0			0					0	8	11699229
103	4787	Saimiri13	AC151884 151892 188246 151890	GAT ACA CAT TCA CTC TGT GGC	CTT AGG ATT CTC CTG GCT TTG	58	0				0	1	0			0					0	8	11676398
104	4789	Saimiri14	AC151884 151892 188246 151890	AGC CCT AAG AAC GTG CAT TC	ACC AGT GAT AGG AAG TGG AC	58	0				0	1	0			0					0	8	11675605
105	4793	Saimiri15	AC151884 151892 188246 151890	AGA ATG CAG GGA ACG GAG TG	CAG TGT CTT TCT AGA AGT TGA G	58	0				0	1	0			0					0	8	11661104
106	4795	Saimiri16	AC151884 151892 188246 151890	CAC CTT CAA TAC TCA TAG TAT G	GCC TAA TAA GAA AAT AAC AAT TGA T	58	0				0	1	0			0					0	8	11647035
107	4799	Saimiri17	AC151884 151892 188246 151890	TTG CAG ATA CAG TAG CTG GC	TCA ACT CAA GTA AAG TAG ATC C	58	0				0	1	0			0					0	8	11629807
108	4805	Saimiri18	AC151884 151892 188246 151890	ATA CTC ATT TCC AAG ACA GCC	AAG CCA GTT GCA AAT GTC GC	58	0				0	1	0			0					0	8	11611797
109	4813	Saimiri19	AC151884 151892 188246 151890	GCT TTC GAA ATT ACT ACT CTG C	GTG GGC TCT TAG ATT ATG AG	58	0				0	1	0			0					0	8	11594747
110	4815	Saimiri20	AC151884 151892 188246 151890	TTT CTA CTT ATT CCC AGA GTT C	GAA GAT CCA GGA AGC CAA AC	58	0				0	1	0			0					0	8	11589485
111	4817	Saimiri21	AC151884 151892 188246 151890	GTG AGG CCC ATT TGA ATA TGC	TGT CTT GCC CAG AGG AGC AG	58	0				0	1	0			0					0	8	11585342
112	4819	Saimiri22	AC151884 151892 188246 151890	GCA GCT AAG AGC AAA GTC AC	CCA CAT CTA GGA GCA CAT TG	58	0				0	1	0			0					0	8	11583596
113	4827	Saimiri23	AC151047 188241 189157 188258 188273	CAT GGA ATG TTT GAC CAA TCA C	GAA TAT CCT AGC CAC TTT TCT C	58	0				0	1	0			0					0	5	131999206
114	5831	Saimiri24	Ray Sag_Alu1D_132d	TGGATGTCTCACAGCTCCAA	TTGATGYCCTTCCCCAAGAC	58	0			0	0	1	0		0	0	0			0	0	7	23047208
115	5833	Saimiri25	Ray Saim_1D_F3_6	AGAGCTGTAATGTGCTGACAGG	GCCGCATATGCTCAGTATCA	55	0			0	0	1	0		?	?	0			0	0	17	49770801
116	5835	Saimiri26	Ray Saim_Alu1D_17	GCTTCCAGTAGTCGTTGTTCC	CAGCTTTTGCTRTGTTAANGC	50	?			0	0	1	0		0	0	0			0	0	20	10982762
117	5837	Saimiri27	Ray Saim_Alu1D_8	CATCCCCCTGGTCAAAGTAA	TTTCCTCTTTGTTGCTTAGGG	59	0			?	0	1	0		?	0	0			0	?	9	109842485
118	5838	Saimiri28	Ray AC147845_2_13	ACAGTGAATGTGCGATRCTC	CATTTGGGAACCCAGANAAA	55	?			0	0	1	0		0	0	0			0	0	7	116805066
119	5843	Saimiri29	Ray AC148186_1_1	TCCAAGCCYGGACTTGTCTC	GRGTCCAGTNAACCCACAGA	55	0			0	0	1	0		?	?	0			0	0	7	116222906
120	5845	Saimiri30	Ray AC148186_1_13	TCCAAAGAGNAAACAAAAGGAA	GCCTAAGATGTTTGGGGAAA	55	?			0	0	1	0		0	0	0			0	0	7	116378470
121	5847	Saimiri31	Ray AC148186_1_14	TACAACACCNACCACCYTGA	RGAAATGNTCACTCCTGTCTTG	55	0			0	0	1	0		0	0	0			?	0	7	116347408
122	5849	Saimiri32	Ray AC148186_1_5	TTTGGTCAAAGATGGGTCTTC	ACATCAGGCATCTCACTTGC	55	0			0	0	1	0		0	0	0			0	0	7	116211969
123	4469	Callithrix1	AC151378 198552 188344 187641	CAA TGG ATA GTG TGA GAA ACC	TGT TGG TGA GAA AAG GCA ATC	58	1	0	0	0	0	0	0				0		0		0	18	24070102
124	4475	Callithrix2	AC151378 198552 188344	CTG TTA AAT GGA CCC TAA TAT C	TAG AAA GAC TTC AGT AAC TAA AG	58	1	0	0	0	0	0	0				0		0			18	24020510
125	4477	Callithrix3	AC188224 188254 188271	CAT GCA GTT TGG CTT CAG TC	CAT TTG AAA GTC CAG TGA ATT C	58	1	0	0	0	0	0				0	0		0		0	4	23397828
126	4833	Callithrix4	AC150604 193070 191895	CTC AGC AAG TAA AAA TTC TCT C	AGA ACT TTA GTG TAG TTA CCA C	58	1	0	0			0	0								0	16	61197483
127	5821	Callithrix5	Ray Saki_1D_46b	CATGGCCACCCTTATGCTA	GAGGGGAAATNGGAATYCAG	55	1	0	0	0		0	0		0	0	?			0	?	3	2538661
128	5825	Callithrix6	Ray AC146661_8_2	AACGTTTGGCCTTCATCCTA	GGAGCAGAATGNNATATGGAAG	55	1	0	0	0		0	0		?	0	0			0	0	7	116686081

0 and 1: absence and presence of integrations

Abbreviations: AT (Annealing temperature), CP (*Callithrix pygmaea*), CG (*Callimico goeldii*), LC (*Leontopithecus chrysomelas*), SI (*Saguinus imperator*), CA (*Cebus apella*), SS (*Saimiri sciureus*), AA (*Aotus azarae*), BA (*Brachyteles arachnoides*), LL (*Lagothrix lagothricha*), AF (*Ateles fusciceps*), AC (*Alouatta caraya*), CC (*Cacajao calvus*), CAS (*Chiropotes albinasus* x *satanas*), PP (*Pithecia pithecia*), CM (*Callicebus moloch*).

Supplementary Table 5.1: Details on studied Y-chromosomal and autosomal markers.

Locus	PCR conditions*	Size (bp)	Substitution model (AIC)	Reference	Forward	Reverse
Y chromosomal						
SRY	1	~800	TVM	Whitfield et al., 1993	CTTGAGAATGAATACATTGTCAGGG	AGGTCTTTGTAGCCAATGTTACCCG
DBY5	2	~720	TVM	Hellborg and Ellegren, 2003	CTTTGAGAAATATGATGATATA	CGAGTAAGTTCAATGTTC
SMCY7	3	~500	TVM+I	Hellborg and Ellegren, 2003	TGGAGGTGCCCRAARTGTA	AACTCTGCAAASTRTACTCCT
SMCY11	4	~600	TVM	Hellborg and Ellegren, 2003	CTGCCCTGYRCCATGCAT	TCCACCTGTTSMAGRACAT
UTY18	4	~900	GTR+G	Hellborg and Ellegren, 2003	CATCAATTTTGTAYMAATCCAAAA	TGGTAGAGAAAAGTCCAAGA
ZFYLI	5	~900	TVM	/	CCTGATTCCAGGCAGTACC	ATCAGGGCCAATAATTATTGCT
Autosomal						
vWF11	6	~950	HKY+G	Chaves et al., 1999	GAGCTGGATGTCCTGGCCATCCATGGCAAC	GAGTGCCTTGTCACTGGTCATCCCACTTCAA
IRBP3	7	~1600	HKY+G	/	CTCTGGACACACGCCCAG	CACACTGCTGGTCAGAATGA
ALB3	8	~1200	GTR	/	GCATTCAAAGTCAACCATG	ACGAAGAGTTGCAACTGTGC
TP2	9	~950	GTR	/	GCAGGTGTACAAAACCAAG	GTCTCATTAGTTGGATTTCC
TTR1	8	~950	TVM+G	/	GGCCCTACGGTGAGTGTT	ACTTTGACCATCAGAGGACA
nuc combi			TVM+G			
mtDNA			GTR+I+G			
*						
1			**5**			
94°C	2'		94°C	2'		
94°C	1'		94°C	1'		
60°C	1'	40x	58°C	1'	40x	
72°C	1'		72°C	1'		
72°C	5'		72°C	5'		
2			**6**			
95°C	10'		94°C	4'		
95°C	0.5'		94°C	1'		
55°C	1'	20x (-0.5°C/cycle)	50°C	1'	40x	
72°C	1'		72°C	3'		
95°C	0.5'		72°C	10'		
45°C	1'	20x				
72°C	1'		**7**			
72°C	10'		94°C	2'		
			94°C	1'		
3			58°C	1'	40x	
95°C	10'		72°C	2'		
95°C	0.5'		72°C	5'		
65°C	1'	20x (-0.5°C/cycle)				
72°C	1'		**8**			
95°C	0.5'		94°C	2'		
55°C	1'	20x	94°C	1'		
72°C	1'		56°C	1'	40x	
72°C	10'		72°C	1.5'		
			72°C	5'		
4						
95°C	10'		**9**			
95°C	0.5'		94°C	2'		
60°C	1'	20x (-0.5°C/cyle)	94°C	1'		
72°C	1'		54°C	1'	40x	
95°C	0.5'		72°C	1'		
50°C	1'	20x	72°C	5'		
72°C	1'					
72°C	10'					

Supplementary Table 5.2: Presence/absence pattern and further details on studied SINE loci.

Name	MM	PC	CGK	CG	PB	SE	SV	TO	TC	PM	PN	NL	Forward	Reverse	Filled	Empty	AT	Locus position (Homo)
Colobinae																		
DoucL_Yd_14Re	0	0	1	1			1		1		1	1	TTTTAAAATGTATGCTCCTTGTCA	TCACTGAACTCCTAAaTCAGCA	468	158	50	chr11:68568822
DoucL_Yd_21	0	0	1	1			1		1		1	1	TGGGAAGTTTGAAGCCTGATA	TGAAGTTCAAAGGCTTAGTTTTATTTT	792	482	57	chr14:52024695
DoucL_Yd_16	0	0	1	1			1		1		1	1	CCCTCTGAGCTCCTTCTGAAT	TGGGATCAGTCTTTGCTGAcT	566	256	50	chr2:60020556
DoucL_Yd_28	0	0	1	?			1		1		1	1	TTGAAAGAaCAGGGGAAATCA	TATCTGACAGCCCTTGACCTG	685	375	50	chr20:30349475
Kirk_Yd_2Re	0	0	1	1			1		1		1	1	GCCTCACaTTCATCTTCCAAAc	TCTATGGTTTCTGGAGGTGCTT	530	220	60	chr4:7764401
Kirk_Yd_33Re	0	0	1	1			1		1		1	1	CTCACAATTTCAGTCTTGTGTGA	GGcCTCAGAGAAGAGACTTTCC	464	154	60	chrX:127523203
SLL_Yd_1	?	0	1	1			1		1		1	1	TTTGGGGAACTTACGGTCTTT	GAGAAGCCACTCACCATTTGa	554	244	57	chr15:24453327
SLL_Yd_8	0	0	1	1			1		1		1	1	CCACCCTTTCCTAAATTTCCA	GTCTTGGCTTTCCCTTCTACG	510	200	50	chr16:46418388
SLL_PY2_25	0	0	1	1			1		1		1	1	CCTGAGGCTGCTCAGAGAAA	TGCTTATACGAGGCAACCTTTA	684	374	55	chr17:56469756
Nasalis_PY2_36	0	0	1	1			1		1		1	1	cAGAAATGTGTTTGTGAAGCAG	ACCCTAATGGCAACATTCAGTT	706	396	60	chr3:194023750
N21	0	0	1	1			1		1		1	1	ACCACTTTTGCTTGGTTCTCAT	TGTCCAGACCaAGAGACATCTG	567	257	60	chr2:242183727
N11	0	0	1	1			1		1		1	1	TGGCAAATTCTCCCTTTATCAC	TGAGTAAAGTAATACAAATTGCACCA	570	260	60	chr10:118380329
N16	?	0	1	1			1		1		1	1	CAGATTCAAAGAGGTGATGTCg	TAGGTGTCTCAGGATGGTGCTA	540	230	60	chr11:56721429
3133	0			1*	1	1*		1*		1*	1*	1	TGTCTTCTTCCCTCTGGAAATCTA	GTGGTCCCCAATCTCCAAG	450	150	58	chr20:57502278
3141	0			1*	1	1*		1*		1*	1*	1*	CCAGTGTTTTGGTGTCAAAATG	GGGGTTTAAGGTTAATGATGACAG	550	250	58	chr12:4069255
3249	0			1	1	1		1		1	1	1	ATGTCTTGAGAGTTCAGGGC	CTTAAGATACACAATATAACAAGC	600	300	58	chr2:203941433
3261	0			1	1	1		1		1	1	1	GTAATATAGCTTGGAAAGATGC	TTATTCTGTGACTTGGAAATAGT	600	300	58	chr2:212208212
3269	0			1	1	1		1		1	1	1	TGTAGCCAGGGAAGCCTCT	TGGGATTTCTAATACTATGCCTTTG	800	500/350	58	chr11:12743945
3365	0			1	1	1		1		1	1	1	GTCTTCTTCCCTCTGGAAATCT	AGCTGCCTGGATGAGACCT	500	200	58	chr20:57502300
3369	0			1	1	1		1		1	1	1	TGGTTTCATGTGTCCACTTAGG	CCAAGAATTTATTGAGCATCCA	900	600	58	chr17:35395755
PB/Asian colobines																		
2474	0			0	1	1		1		1	1	1	TGGACAAGCTGAAGACATGG	CTGGATCTAGAGCTAGCTAG	500	200	58	chr6:4376475
3371	0			0	1	1		1		1	1	1	CAGAGTGCTAAATTCATGCTTC	TGGCTGTTCCAAAGTCAGTTAG	600	300	58	chr9:111700546
3373	0			0	1	1		1		1	1	1	TCGTTTGAAGATTTCAGTTGG	CTCTGTCTTGACAGCAGTAAC	600	300	58	chr1:8704528
Asian colobines																		
DoucL_PY2_3	0	0	0	?	0		1		1	1	1	1	TGCGTATTTCcACATTTCTGAC	GGCACAGCAAATGACTACGTTA	560	250	55	chr3:44778934
DoucL_PY2_12RE	0	0	0	0	0	1	1		1	1	1	1	CCTTGATTTCATCTATGGGCTTA	TGACAAGGGAAATAGAAAGATTGA	743	433	60	chr3:32047688
DoucL_PY2_16	0	?	0	0	0	0	?	0	1	?	1	1	AGGGGCTAAAGACAtATCATGG	GGAACAAATCTTACCTCACTACGAA	759	449	54	chr17:42193854
DoucL_PY2_21a	0	0	0	?	0	1	1	1	1	1	1	1	CACATTAGAGgGCgCTAGaGTT	TCCTGTTGGTCCTTTGGTAGTT	574	264	55	chr20:41704425
DoucL_Yd_2RE	0	0	?	0	0	1	1	1	1	1	1	1	TCCCACTTCCTCTTTTTCTCAG	CTTTTAtGGTTTTGcGGCTAAT	578	268	60	chr2:158824533
DoucL_Yd_5RE	0	0	0	0	0	1	1	1	1	1	1	1	ATGTGAAGACCTCTGCCCAGTA	TCCTTTTGTTCAAACTGCTTCTT	440	130	60	chr8:123230360
DoucL_Yd_10	0	0	0	0	0		1		1	1	1	1	TCAAAGAAGCAGCCTTCAAAA	TGCAAAACtCATCTGTgCTGT	610	300	50	chr4:154270313
DoucL_Yd_22	0	0	0	?	0		1		1	1	1	?	TTTCATTTAACCCTCCAGCTATTT	TAACCAAACCTGGCCAGAAC	789	479	50	chr3:38767375
DoucL_Yd_27	0	0	0	0	0	1	?	1	?	1	1	1	CAAGGGAGGATTCTAAGTCAGG	CAAATACCTCTTGAGAGCCAGTG	624	314	60	chr8:57979210
PFL_PY2_7	0	0	0	0	0	1	1	1	1	1	1	1	AtCCTGCgGTTCACATTTCTAT	GCCAACTTCTTAGAAACAACAGG	560	250	55	chr2:1049803

Supplementary

Name	MM	PC	CGK	CG	PB	SE	SV	TO	TC	PM	PN	NL	Forward	Reverse	Filled	Empty	AT	Locus position (Homo)
PFL_PY2_9	0	0	0	0	0	1	1	1	1	1	1	1	TTCTAAGCAGGACCTAAAAAGCA	TCATTTCAGAGATTTTGCTGATG	561	251	55	chr6:123355176
PFL_Yd_1	0	0	0	0	0		1		1	1	1	1	GTCAGACAAGGTGTGGAACAA	AATgGTTATGTTTGATCTCTTTAACA	466	156	50	chr14:63230836
PFL_Yd_11	0	0	0	0	0		1		1	1	1	1	CAAATGCTAACCACCTTGCAT	GCCCCATCAAAGAATGTATTTC	555	245	48	chr11:75528664
PFL_Yd_15	0	0	0	0	0		1		1	1	1	1	TTCAGTTTCATCTGGGAGTGG	CAtGTGGAGTGACCTGTGTTG	407	97	50	chr15:63007825
SLL_Yd_2	0	0	0	0	0	1	1	1	?	1	1	1	AaCACATCAACACATGCCTCA	CCTTTGGGTTACTCTCCAGGT	645	335	57	chr2:149643161
Sll_Yd_20	0	0	0	0	0		1		1	1	1	1	TGGTTAAGTAAAGGGGGTGCT	TCATGCTACAACCACAAGCTG	574	264	60	chr1:214914371
Sll_Yd_4a	0	0	0	?	?		1		1		1	1	GCAGgGTAAGGCATGACTAAAA	TGAGGGAGAGTGAGTGTGGAT	720	410	50	chr15:40759804
N17	0	0	0	0	0		1		1	1	1	1	ACAGGGATTTCAGAGCACAAGT	CAGCTCATTATTCCCAAACACA	530	220	55	chr1:172065132
Nasalis_PY2_27	0	0	?	?	0	1	1	1	1	1	1	1	CAAATGTTCCGTTGAGTCCA	GAGTCTTGGAAGATGCAGTGA	688	378	60	chr6:43076845
Nasalis_PY2_28	0	0	0	0	?		1		1	?	1	1	GTGCTAATCCTTGGCCACAT	GAATCTGTCTGCCAAAGtCACA	487	177	60	chr2:116383219
Nasalis_PY2_32	0	0	0	0	0		1		1	1	1	1	TTTTGTAAACAGCCAAAGCTCA	TTGTTGAAAATATGGCACAAGC	624	314	55	chr10:50352041
3125	0			0	0	1		1		1	1	1	AAGAATCCCAGGGAAGAACACT	TTGCTGGCAAAGTGACTCCT	700	400	58	chr7:107109056
3131	0			0	0	1		1		1	1	1	CCTGCCACTTCTGTCCATCT	AGAACAAACACCAAGACAACAGC	450	150	58	chr3:193991928
3149	0			0	0	1		1		1	1	1	GCTTTGCCACATAAAGAGCTG	GGTTAGGTGCAAATGGGAAAC	420	120	58	chr2:109178381
3247	0			0	0	1		1		1	1	1	TCAATCTTCCAGGGAAAATAAAG	GAATATTAGTTGAAATATTTAGGC	600	300	58	chr15:48692337
3253	0			0	0	1		1		1	1	1	GACCATGGTAAGACAAATGTG	GACTCAGGCTTAATTTTAAGTC	500	200	58	chr4:39427060
3267	0			0	0	1*		1		1	1	1	CACCAAGCACAACTGTGAGG	TCTGCCATAGCCATCAGTCA	600	300	58	chr1:217214420
3377	0			0*	0	1		1		1	1	1	CTCTTGGTTGGGGTGAAGC	GATGGTTGAACAGTGAGACTTGA	500	200	58	chr10:119562900
3379	0			0*	0	1		1		1	1	1	AGCACCATCAGGCACTCACT	GGGAGATTGGGAAATGGAGT	700	400	58	chr6:104107889
PM/odd-nosed																		
Douc_PY2_22	0	0	0	0		0	0		0	1	1	1	TGGTGCTCACTATCAGCAATTT	TGCTTAAAGTCCATCAACATGC	524	214	55	chr12:18654883
3277				0		0		0		1	1	1	ACCTTGATCTCAGGGATCCT	GGTCAAAGTCCTACTTAAGGA	410	110	58	chr7:20757072
odd-nosed																		
DoucL_PY2_13RE	0	0	0	0		0	0	0	0	0	1	1	CAATGGGGACAGCTATTTTTCT	CTTTGGGTGTGTAGTTCCCATT	451	141	60	chr13:104854571
DoucL_Yd_36	0	0	0	0		0	0	0	0	0	1	1	CTCAaGCTTCTCCCCTCCTTA	AAGGCACAGGATTCTGCTTTT	434	124	60	chr11:20035391
Nasalis_PY2_23	0	0	0	0			0	0	0	0	1	1	GGGAAAAGGATGAGTGAATCTG	TCTTCTCATAGGATGCCAGTCA	700	351	60	chr4:86357817
N6	0	0	0	?		0	0		0	0	1	1	TGCCAGGAGAAAACTTAAATGA	TTAAGGGGAGGAAGTAGGGAAG	550	240	60	chr7:23888495
DoucL_Yd_20	0	0	0	0		1	0	1	0	1	1	1	TCAAAACTTGCATTTCTTCACAA	CACTTTCAAAATTGCATGAACA	680	370	51	chr11:24290034
3143	0			0		0		0		0	1	1	AGAAAGTCCCTCCCCAACAC	AAGTTGGCAAAGTGGATTGC	550	250	58	chr1:201638132
3381	0			?		0		0		0	1	1	GCATGATAAGAGTGGAAATCTGTG	TCAACTGATGCAGAAAATGC	500	200	58	chr18:3757704
T/S																		
PFL_PY2_2	0	0	0	0		1	1	1	1	0	0	0	CATGCTCACCTTGTATTCCTTG	TGAGGTACTGCTCTGGTGAGTT	664	354	55	chr11:111606943
PFL_PY2_14RE	0	0	0	0		1	1	1	1	0	0	0	TGCCAAAAACTCAGGTTAAGAGA	ATTTTGGGGGAAAACTGCTATC	563	253	60	chrX:14639350
Sll_Yd_18	0	0	0	0		1	1	1	1	0	0	0	GCATTCTGGACAGTGGTGATT	GACGAGATAcaATGCTTCtGAAA	897	587	60	chr5:50539061
SLL_PY2_17	0	0	0	0		1	1	1	1	0	0	0	TGATCCATCCCTCTTAGGAGTC	AGATCTCGGTGCCACAAATAGT	694	384	50	chr12:118197378
SLL_PY2_23	0	0	0	0		1	1	1	1	0	0	0	ATAGCCTGCACGAAAAGACCTA	CAGGAGTgTTTCTCATTGACCA	513	203	60	chr22:16306635

Name	MM	PC	CGK	CG	PB	SE	SV	TO	TC	PM	PN	NL	Forward	Reverse	Filled	Empty	AT	Locus position (Homo)
2457				**0***		**1**		**1**		**0**	**0**	**0**	TGATTAAAGTCAGATGAACACC	GTGTAATGGGATGAAGAACAC	540	240	58	chr15:52962593
2652				**0***		**1**		**1**		**0**	**0**	**0**	ATACATAGCATTGACTTAACTCT	GATCCTGAGCCCACTATTCT	520	220	58	chr5:25601752
2668				**0***		**1**		**1**		**0**	**0**	**0**	ACATCAGTGACATCAAATAAGG	GAGGAAAAGATACTTTCTCATG	470	170	58	chr8:95826066
3257				**0**		**1**		**1**		**0**	**0**	**0**	GGATTGAGAGCAATTTTAAAAGGA	GTTCACTCCCAAATCATACTTC	940	640	58	chr15:63693377
3269				**0**		**1**		**1**		**0**	**0**	**0**	TGTAGCCAGGGAAGCCTCT	TGGGATTTCTAATACTATGCCTTTG	800	500/350	58	chr11:12743945
Genus specific																		
AC123963_7_2b	0	0	1	1	0		0		0		0	?	AACTTGGAGGGGAATGTAATCTTT	TGAATCAAGAgTTTAAaAGATaGAGGT	597	277	55	chr7:116755036
Kirk_Yd_1	0	0	1	1	0		0		0		0	0	TTTGAGCAAGCTATTGGCATT	CATTGTGATGCTGAAGCTGAA	640	330	60	chr2:21489425
Kirk_Yd_14	0	0	1	1	0		0		0		0	0	GCCATGGTAATAAAACACCTTTC	CCAAGATTCCgGTGTTCTTTT	897	587	50	chr11:74212267
Kirk_Yd_27	0	0	1	1	0		0		0		0	0	TCTCCCAACATCCTCAAACAC	CCACTTCCCAAACCATGATAA	562	252	50	chr2:116577719
B5	0	0	1	1	0		0		0		0	0	CGACCAGGAGAAAAATCCTAAG	CCTGTACCATGGCTTTTTATGg	768	458	60	chr6:144730945
B9	0	0	1	1	0		0		0		0	0	TCCTTAAGGCTAATGCAAGGAC	CCTcCATACACAGTGGTaAGAAAA	490	180	55	chr3:193811839
B11	0	0	1	1	0		0		0		0	0	cGTgTACAGGATTTATGgTGGt	ATCCAGTGATTTCCTGTTGTATCA	555	245	50	chr5:114584227
B12	0	0	1	1	0		0		?		?	0	TGAGATAGGAAATGGAAGAGAAGC	CCTGAGTCAATGGAAcTCCTTc	550	240	50	chr14:32017959
B13	0	?	1	1	0		0		0		0	0	CCATTTAACTTCCTCGTTTCCA	GGTGACTTTTTGATCAGAGTTGC	570	260	60	chr16:31815841
B14	?	0	1	1	0		0		0		?	0	AGGTAGACTTCCGTATGATAAAACA	CCAAATGTAATGGtTTTTcTTGTGT	560	250	50	chr8:127080357
B15	0	0	1	1	0*		0		0		0	0	TATGTCTTAGGGCTGCAAATCC	TGTTAAAAGtGCgGTTCTCAAA	520	210	50	chr7:79330661
2465				0*		0		0		1	0	0	ACTAGTTGCAGAGGTGTTTCT	AAGATGAGCAGGTAGAATCCA	560	260	58	chr14:41243669
2476				0*		0		0		1	0	0	CAAACACACACTGAATCTAAGT	ATAGACAGTTTGGAAAATGTTGA	600	300	58	chr7:103615121
3243				0		0		0		1	0	0	GTGGTGAGTTGAAATTCATTTGTT	AGGGATTATATTCCTCTTGTTATGAA	430	130	58	chr11:89502226
3255				0		0		0		1	0	0	CTTAGACAAATATCACTAGTGC	GCAAAACAGATGGTAGGATGC	380	80	58	chr12:85138468
3259				?		0		0		1	0	0	ATGCAGGTAGGGTCCTCTC	CTTCTTGGTGCTGAATTGGC	500	200	58	chr12:29484911
3387				0		0		0		1	0	0	TCCCACTAACTCAAGCTCTATC	GTGGAGATTTTCTCAACCTTGG	530	230	58	chrx:24886003

0 and 1: absence and presence of integrations
Bold: loci embedding in Figure 5.1
Asterisks: not sequenced loci
Boxes: genera tested by Xing et al. (2005)
Yellow: loci from Xing et al. (2005)
Green: loci from Xing et al. (2005), but not embedded in their cladogram
Orange: loci from this study.
Abbreviations: AT (Annealing temperature), MM (*Macaca mulatta*), PC (*Papio cynocephalus*), CGK (*Colobus guereza kikuyuensis*), CG (*Colobus guereza*), PB (*Piliocolobus badius*), SE (*Semnopithecus entellus*), SV (*Semnopithecus vetulus*), TO (*Trachypithecus obscurus*), TC (*Trachypithecus cristatus*), PM (*Presbytis melalophos*), PN (*Pygathrix nemaeus*), NL (*Nasalis larvatus*).

Supplementary Table 5.3: Accession numbers from GenBank of complete mitochondrial genome sequences

Species	Accession number
Homo sapiens	AC000021
Pan troglodytes	EU095335
Papio hamadryas	NC001992
Chlorocebus aethiops	NC007009
Colobus guereza	NC006901
Piliocolobus badius	NC008219
Semnopithecus entellus	NC008215
Trachypithecus obscurus	NC006900
Presbytis melalophos	NC008217
Pygathrix nemaeus	NC008220
Nasalis larvatus	NC008216

Danksagung

Ich bedanke mich herzlich bei PD Dr. Lutz Walter, der mir die Möglichkeit gegeben hat in seiner Forschergruppe im Rahmen einer Doktorarbeit weiter an diesem interessanten Thema zu arbeiten.

Ein grosses Dankeschön geht natürlich an Christian Roos mit seinem unfassbaren Schatz an Primatenmaterial. Es ist wirklich spannend mit so wertvollen Proben zu arbeiten, und sei es auch nur der Yeti. Er hatte immer ein offenes Ohr für Probleme oder Anregungen jeglicher Art, ob privat oder beruflich. Die Abende mit Augustiner, Weisswürsten und den Schnecken sind unvergesslich. Ausserdem hat er viel Geduld mit mir bewiesen… aber hey, ich auch mit ihm. Ohne Ihn wäre diese Arbeit nicht zustande gekommen.

Ein weiterer Dank geht selbstverständlich auch an meinen Referenten Prof. Dr. Peter Kappeler und Korreferenten Prof. Dr. Ernst Wimmer.

Ich möchte mich bei allen derzeitigen und ehemaligen Primatengenetikern für das gute Arbeitsklima, für das gemeinsame Speisen in der Mensa und sonstige Freizeit- und Laboraktivitäten bedanken.

Ein riesengroßer Dank geht an Christiane Schwarz, die mir unglaublich viel dabei geholfen hat die Mengen an Laborarbeit zu bewältigen.

Für die schönen Abende ausserhalb des Labors, die natürlich nie etwas mit der Arbeit zu tun hatten, möchte ich mich bei Anna, Alex, Philip, Conny, Wolf und Meike bedanken. Das Gleiche gilt für meine Freunde Andy, Kristof und Lu sowie für die über Deutschland verstreuten Reste.

Schließlich möchte ich mich noch bei meinen Eltern bedanken, die irgendwie ja hauptverantwortlich für mich sind, und bei Laura, die mich ertragen hat, mich aufgebaut hat, mir Ratschläge gegeben hat und immer für mich da war.

Lebenslauf

Martin Peter Osterholz
Diplom-Biologe
geboren am 26.02.1980 in Brunsbüttel
Staatsangehörigkeit: deutsch

seit 2005	Anfertigung der vorliegenden Dissertation in der Forschergruppe Primatengenetik bei PD Dr. Lutz Walter am Deutschen Primatenzentrum in Göttingen
2005	Abschluss des Studiums zum Diplom-Biologen
2004 – 2005	Anfertigung der Diplomarbeit „Untersuchungen von SINE-Integrationen im Genom der Colobidae„ in der Forschergruppe Primatengenetik bei PD Dr. Lutz Walter am Deutschen Primatenzentrum in Göttingen
1999 – 2004	Studium im Fachbereich Biologie an der Georg-August Universität in Göttingen
1999	Abitur
1990 – 1999	Gymnasium Marne
1986 – 1990	Grundschule Helse