



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Next-generation museomics disentangles one of the largest primate radiations

**Citation for published version:**

Guschanski, K, Krause, J, Sawyer, S, Valente, LM, Bailey, S, Finstermeier, K, Sabin, R, Gilissen, E, Sonet, G, Nagy, ZT, Lenglet, G, Mayer, F & Savolainen, V 2013, 'Next-generation museomics disentangles one of the largest primate radiations', *Systematic biology*, vol. 62, no. 4, pp. 539-554.  
<https://doi.org/10.1093/sysbio/syt018>

**Digital Object Identifier (DOI):**

[10.1093/sysbio/syt018](https://doi.org/10.1093/sysbio/syt018)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Systematic biology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Next-Generation Museomics Disentangles One of the Largest Primate Radiations

KATERINA GUSCHANSKI<sup>1</sup>, JOHANNES KRAUSE<sup>2</sup>, SUSANNA SAWYER<sup>2</sup>, LUIS M. VALENTE<sup>1</sup>, SEBASTIAN BAILEY<sup>1</sup>, KNUT FINSTERMEIER<sup>2</sup>, RICHARD SABIN<sup>3</sup>, EMMANUEL GILISSEN<sup>4,5,6</sup>, GONTRAN SONET<sup>7</sup>, ZOLTÁN T. NAGY<sup>7</sup>, GEORGES LENGLET<sup>7</sup>, FRIEDER MAYER<sup>8</sup>, AND VINCENT SAVOLAINEN<sup>1,9,\*</sup>

<sup>1</sup>Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL5 7PY, UK; <sup>2</sup>Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany; <sup>3</sup>Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK; <sup>4</sup>Department of African Zoology, Royal Museum for Central Africa, Leuvensesteenweg 13, Tervuren 3080, Belgium; <sup>5</sup>Laboratory of Histology and Neuropathology CP 620, Université Libre de Bruxelles, 808, route de Lennik, B-1070 Brussels, Belgium; <sup>6</sup>Department of Anthropology, University of Arkansas, Fayetteville, AR 72701, USA; <sup>7</sup>Royal Belgian Institute of Natural Sciences, Joint Experimental Molecular Unit, Rue Vautier 29, 1000 Brussels, Belgium; <sup>8</sup>Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung an der Humboldt-Universität zu Berlin, Invalidenstrasse 43, 10115 Berlin, Germany; and <sup>9</sup>Royal Botanic Gardens, Kew, Richmond TW9 3DS, UK

\*Correspondence to be sent to: Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, SL5 7PY Ascot, Berks, UK; E-mail: v.savolainen@imperial.ac.uk.

Received 4 August 2012; reviews returned 19 November 2012; accepted 8 March 2013  
Associate Editor: Elizabeth Jockusch

**Abstract.**—Guenons (tribe Cercopithecini) are one of the most diverse groups of primates. They occupy all of sub-Saharan Africa and show great variation in ecology, behavior, and morphology. This variation led to the description of over 60 species and subspecies. Here, using next-generation DNA sequencing (NGS) in combination with targeted DNA capture, we sequenced 92 mitochondrial genomes from museum-preserved specimens as old as 117 years. We infer evolutionary relationships and estimate divergence times of almost all guenon taxa based on mitochondrial genome sequences. Using this phylogenetic framework, we infer divergence dates and reconstruct ancestral geographic ranges. We conclude that the extraordinary radiation of guenons has been a complex process driven by, among other factors, localized fluctuations of African forest cover. We find incongruences between phylogenetic trees reconstructed from mitochondrial and nuclear DNA sequences, which can be explained by either incomplete lineage sorting or hybridization. Furthermore, having produced the largest mitochondrial DNA data set from museum specimens, we document how NGS technologies can “unlock” museum collections, thereby helping to unravel the tree-of-life. [Museum collection; next-generation DNA sequencing; primate radiation; speciation; target capture.]

Studying large and taxonomically diverse groups of organisms is a daunting task, particularly when the species in question are elusive, live in inaccessible areas, or are rare and endangered. All these factors impede sample collection and many studies are relegated to focus on a subset of taxa, being unable to explore the entire taxonomic diversity (Moyle et al. 2009; Chan et al. 2010). Museum collections contain specimens gathered over several centuries of natural history explorations and represent a fantastic and underused source of material for biological studies. However, up until recently, effectively exploiting museum specimens as a source of DNA has not been practical on a large scale due to low quantity, poor quality of endogenous DNA, and high levels of contamination (Wandeler et al. 2007).

Here, we focus on guenons (tribe Cercopithecini), the most species-rich group of extant African primates, whose taxonomic diversity is only surpassed by Malagasy lemurs and New World monkeys. Guenons comprise 5 genera (the predominantly arboreal *Cercopithecus*, *Miopithecus*, and *Allenopithecus* and the terrestrial *Erythrocebus* and *Chlorocebus*) and 63 species and subspecies (Grubb et al. 2003; Wilson and Reeder 2005). The genus *Cercopithecus* is further subdivided into 8 species groups, usually comprised of several closely related, allopatric species. However, there is also a number of groups that contain only one species, as they differ substantially from the other species (Grubb et al. 2003). The genus *Chlorocebus*, although

now a separate genus, was traditionally classified as the ninth species group within *Cercopithecus* (Grubb et al. 2003; Wilson and Reeder 2005). To investigate the mechanisms driving guenon diversity, we first set out to resolve their phylogeny. Previous studies included only a limited number of species and phylogenetic inferences were based on relatively short DNA sequences or used Alu elements, which precluded the estimation of well-supported and resolved trees (Tosi et al. 2004, 2005; Xing et al. 2007; Tosi 2008; Chatterjee et al. 2009). These limitations are understandable, given that many taxa are threatened (<http://www.iucnredlist.org/>, last accessed April 2, 2013), making collection in the wild impractical. Samples obtained from captive animals are limited, simply because not all taxa are available. To overcome these problems, we turned to museum collections. Using next-generation DNA sequencing (NGS) technology coupled with target DNA enrichment, we sequenced mitochondrial genomes of 92 museum specimens, obtaining a wealth of genomic information unparalleled by previous studies.

A number of factors might have contributed to the evolutionary radiation of guenons. Guenons occupy a wide geographic range, being distributed over most of sub-Saharan Africa (Butynski 2003). Since the majority of guenons are forest-dwellers, the evolution of new species may have been driven by isolation in forest refugia. Guenon speciation is thought to have taken place during the last 10 myr (Disotell and Raam 2003),

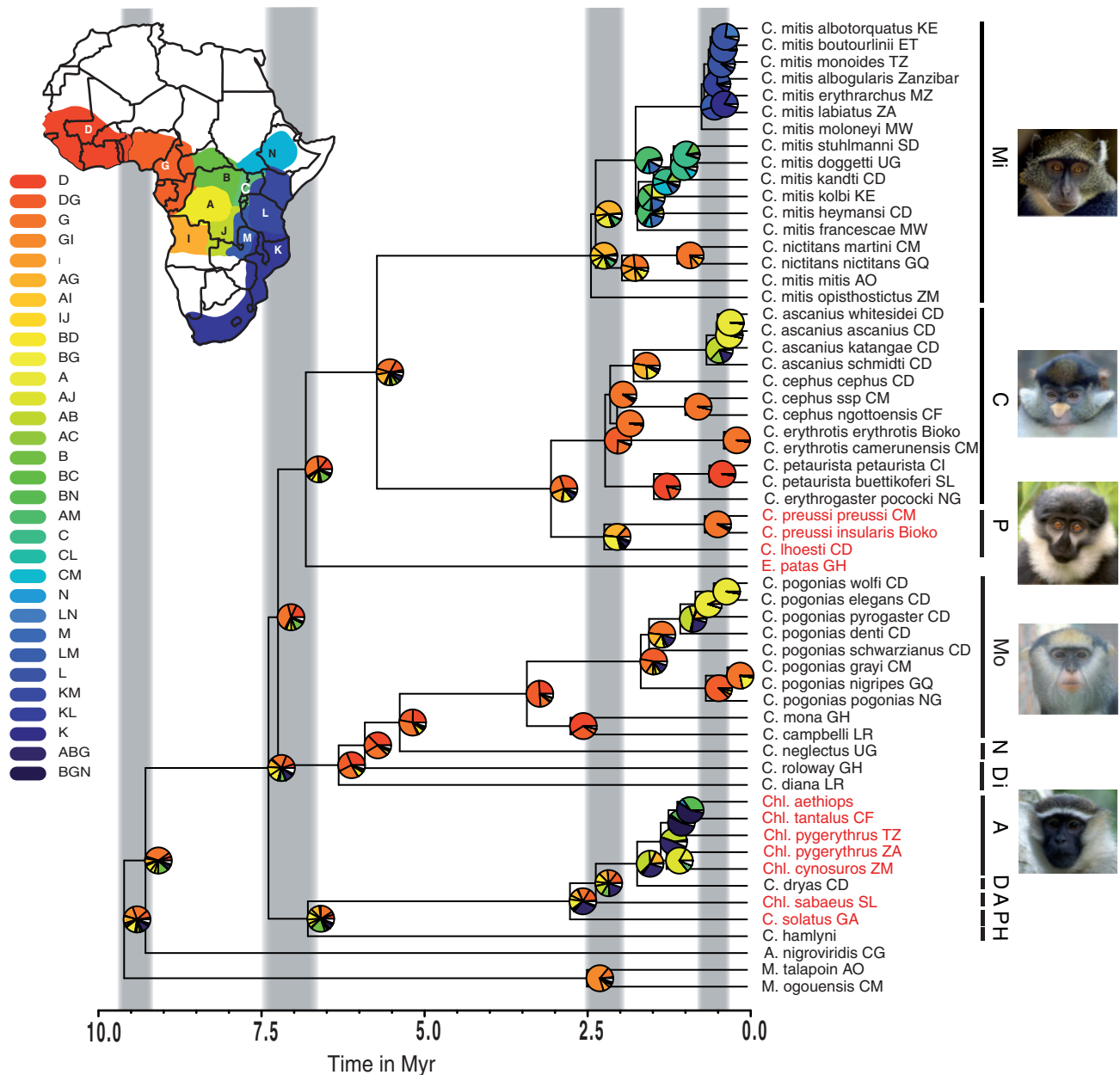


FIGURE 1. Ancestral ranges and the timing of diversification in guenons. Only a single representative for each taxon is shown, for the full tree see Figure 2 and Supplementary Figure S3. Pie charts at the nodes indicate ancestral areas, with color corresponding to the location of these areas on the map of Africa shown in the upper left corner. The relative proportion of each color represents the fraction of the global likelihood for the given geographic area. Gray bars indicate the timing of 4 main speciation events (see “Results” section). Vertical bars next to the species names refer to the species groups with photographs showing one of the species group’ members (Mi: *C. mitis* group, C: *C. cephus* group, P: *C. preussi* group, Mo: *C. mona* group, N: *C. neglectus* group, Di: *C. diana* group, A: *C. aethiops* group, D: *C. dryas* groups, and H: *C. hamlyni* group). Tip labels in red highlight the members of the terrestrial clade, also see legend of Figure 2. Myr = million years. Ancestral ranges: A, Congo basin; B, northern DRC; C, northern Rift Valley; D, Upper Guinea; G, Lower Guinea; I, Angola; J, southeastern DRC; K, southeastern Africa; L, northeastern Africa; M, Zambia; N, Ethiopia/Sudan. Photograph of *C. m. albogularis* by Y.A. de Jong and T.M. Butynski—wildsolutions.nl.

a time period marked by major climatic changes in tropical Africa (Bonnefille 2010). It is thus possible that repeated geographic isolation, leading to reduced gene flow between populations, resulted in the diversity of taxa seen today (i.e., allopatric speciation). Guenons are well known for their morphological (Fig. 1) and acoustic diversity (Kingdon 1980, 1988; Gautier 1988)—traits that

can be subject to mate choice. The evolution of these characteristics together with differences in behavior is hypothesized to contribute to species recognition. Furthermore, guenons exhibit a great cytogenetic diversity, with diploid chromosome numbers ranging from 48 to 72 (Dutrillaux et al. 1988; Moulin et al. 2008), which may convey reproductive isolation. Finally,

hybridization could have been involved in the evolution of guenons, as documented in other mammals (Arnold and Meyer 2006; Larsen et al. 2010). Interspecies mating in guenons is well known from captivity and the wild and is reported to produce viable and fertile offspring (Detwiler et al. 2005). Furthermore, many extant guenon taxa occur in sympatry and frequently form polyspecific troops (Gautier-Hion 1988), which allows for interspecific or even intergeneric hybridization.

In this study, we generate a comprehensive phylogenetic framework based on mitochondrial (mtDNA) genomic sequences from an almost complete sampling of guenons, obtained primarily from museum specimens. The tree was used to infer ancestral distributions, to test for evolutionary constancy of diversification rate, and to evaluate the importance of geographic isolation in guenon radiation. Using phylogenetic, geographic, and climatic information, we assess the possible role of past climatic changes in the speciation of guenons. Finally, comparisons between nuclear and mitochondrial-based phylogenetic trees allow us to highlight important topological discrepancies, which merit future investigations by suggesting that hybridization and introgression could have played a role in shaping guenons' diversity.

## MATERIALS AND METHODS

### *Sample Collection*

The taxonomy of guenons is still relatively unsettled. For our study, we followed the taxonomy proposed by Grubb et al. (2003) and also took into account the taxonomic classification by Wilson and Reeder (2005) (Supplementary Table S1, Dryad doi: 10.5061/dryad.7k14q). We collected samples in 4 European museums: (i) Museum für Naturkunde in Berlin (MfN), Germany; (ii) Royal Museum for Central Africa (RMCA) in Tervuren, Belgium; (iii) Royal Belgian Institute of Natural Sciences (RBINS) in Brussels, Belgium; and (iv) Natural History Museum (NHM) in London, United Kingdom. We collected up to 300 mg of material from skins, skulls, and skeletons. The geographic origin of the specimen and the date of collection were taken from museum records. Sources of sampled material differed in these museums depending on their policies for invasive sample collection (Supplementary Table S2). Most samples represented dried tissue that was still attached to skeletons and skulls. Samples were also collected from skins by taking up to 5 mm<sup>2</sup> of ear cartilage or up to 4 mm<sup>2</sup> of fingertips.

During sample collection we used gloves, which were changed between each specimen, and cleaned the working surfaces with 20% dilution of commercial bleach and distilled water. The collection tools were sterilized either by fire or by incubating them in bleach and subsequent rinsing in distilled water. Where possible, we sampled from secluded parts of the

specimen that were unlikely to have come in direct contact with human handlers. We also attempted to remove the surface layer to avoid sampling highly exposed and potentially contaminated parts of the specimen. In a few cases, we drilled into the broken part of the skulls or skeletons to retrieve bone material or collected the spongy parts of the nasal bone. When drilling, we used sterile one-way coats, facemasks, and hairnets. For some specimens, we collected the root fraction of the teeth. Where possible, the collection and subsequent selection of the samples for extraction were aimed to contain at least 2 representatives of each taxon from across its geographical range, and ideally deposited in different museums. In total, samples from 120 specimens were used for DNA extraction.

### *DNA Extraction and Library Preparation*

DNA from skin, dried tissue, teeth, and bone samples was extracted in a laboratory dedicated to ancient DNA, which implements several precautions against contamination (Green et al. 2008). For the extraction, we used 10–253 mg of material per sample (Supplementary Table S2). Tissue samples were ground with mortar and pestle. The extraction procedure followed a column-based extraction protocol (Rohland et al. 2009) with one modification: the DNA was eluted using TE buffer with 0.05% Tween 20. Siliconized tubes were used in all processing steps and for long-term storage of the extracts. For each 10–20 samples, we processed 2 negative controls to check for possible cross-contamination during the extraction process.

Illumina libraries were prepared from 30 µl of the extract from each sample, under clean room conditions, and following Meyer and Kircher (2010). To allow highly multiplexed sequencing, an Illumina adapter containing a 7-basepair (bp) barcode was attached to the 3'-end of each sample (Meyer and Kircher 2010). We used 96 barcode sequences, which differed from each other by at least 3 substitutions. This dramatically reduces the chance of converting one barcode into another one by sequencing errors. We included additional blanks for every 23 processed samples to check for possible contamination during library preparation. Each indexed extract was amplified with Phusion<sup>TM</sup> High-Fidelity Master Mix, 400 nM of each IS5 and IS6 primers until saturation to ensure high prevalence of DNA fragments.

### *Enrichment for Mitochondrial DNA*

For the targeted capture of mitochondrial fragments, we used a published procedure (Maricic et al. 2010). Briefly, this protocol is based on hybridization of mitochondrial fragments contained in the complex mixture of the sample extract to a bait consisting of fragmented mitochondrial genomes from closely related taxa.



To construct the bait, we amplified the entire mitochondrial genomes of 3 guenon taxa: *Cercopithecus mitis monoides*, *Cercopithecus diana*, and *Erythrocebus patas*. Long-range PCR was performed on high-quality DNA samples resulting in 2 PCR fragments (ca. 8000 and 10 000 bp). The amplification protocol and primate specific mitochondrial primers were as in Finstermeier (2010). The 2 fragments of bait were pooled in equimolar amounts to a total of 2  $\mu$ g for each bait sample and sonicated (Bioruptor, Diagenode, Liege, Belgium) 10 times for 30 s with the output selector set to high. This resulted in fragment size distribution between 250 and 2500 bp. Subsequently, the fragments were biotinylated and immobilized on streptavidin-coated magnetic beads (Maricic et al. 2010).

Indexed DNA extracts from museum samples were combined into 9 different capture pools, each containing 8–20 samples (Supplementary Table S2). In each capture pool, the DNA was made single-stranded and hybridized with blocking oligos that “mask” the Illumina adapters (Maricic et al. 2010). Subsequently, each capture pool was mixed with one or a combination of bait genomes. More specifically, 3 capture pools were hybridized only to the mitochondrial genome of *C. mitis monoides* as bait, 3 more to the *E. patas* mtDNA genome, and the remaining 3 were hybridized to a combination of *C. mitis monoides*, *C. diana*, and *E. patas* (Supplementary Table S2). The choice of samples to be captured in a particular pool was based on current knowledge of taxonomic and phylogenetic relationships within guenons (Tosi 2008), attempting to use a bait from a species that was evolutionarily closest to the sample to be captured. For instance, all representatives of the *C. mitis* species group were captured with *C. mitis monoides*, and species within the proposed terrestrial clade (Tosi et al. 2004) were predominantly captured with *E. patas* (this included all *Chlorocebus* species). Species for which no clear phylogenetic information was available, for example, *Cercopithecus dryas*, or that are supposedly placed equidistant from all bait taxa, for example, *Cercopithecus neglectus*, were captured with the combination of all 3 baits (Supplementary Table S2). We adjusted the amounts of samples to be pooled to match their initial ratio after extraction. This was done to ensure that each DNA fragment present in the initial sample is sequenced to approximately the same depth. Because we amplified each extract until saturation (see above), extracts with low initial DNA content will contain many more copies of the same DNA fragment than extracts with high initial DNA content. The adjustment to the initial ratio thus reduces the chance of sequencing the same PCR-amplified fragment a disproportional number of times. We also captured all negative controls (extraction and library preparation), distributing them randomly among the capture pools. The mixture of sample and bait was rotated in a hybridization oven at 65°C for 48 h. At the end of the incubation period, nonhybridized fragments were washed away and the single-stranded DNA was melted off the beads with 125 mM NaOH. The captured DNA fragments were

purified with MinElute (Qiagen) and eluted with 10  $\mu$ l EB buffer containing 0.05% Tween 20.

### Sequencing and Mapping Analyses

All captured library pools were sequenced directly on 3 lanes on the Illumina GAIIx platform, with no amplification after DNA capture to avoid “jumping PCR.” The pools were combined in such a way that each sample on a given lane had a unique 7-bp barcode. Each lane contained 17–49 samples (19–60 including blanks), following recommendations in Meyer and Kircher (2010). Sequencing was carried out from both ends of the fragments with 76 cycles per read. After standard base calling, the alternative base caller Ibis (Kircher et al. 2009) was used. It relies on a training data set for adjusting quality scores derived from PhiX 174 control reads, which were either spiked into each lane or obtained from a dedicated control lane. A quality filter was applied to the data that removes all reads in which more than 5 bases have quality score below 15 on the PHRED scale (Kircher 2012). Subsequently, reads were separated by sample based on barcode sequences. Perfect matches with the original barcode sequence were required (Meyer and Kircher 2010). Quality scores were applied and sequences from the paired-end reads were merged into single fragments (while simultaneously removing adapters) only if forward and reverse sequences overlapped by a minimum of 11 bp (Kircher 2012). Merging of paired-end reads further decreases the prevalence of sequencing errors, particularly in libraries with fragments shorter or equal to read length. In our case, median fragment size was 58 bp (Supplementary Fig. S1) and thus below the 76-bp read length. For bases in the overlapping stretch, the consensus sequence was called by either summing up the quality scores of identical bases or by calling the base with the higher quality score (Green et al. 2010; Krause et al. 2010). Only merged fragments were retained for subsequent sequence analysis.

To map merged fragments, we used the genomes of *Chlorocebus sabaeus* (accession EF597503 in GenBank), *C. mitis monoides*, and *E. patas* (Finstermeier 2010) as references. The latter 2 reference genomes were derived from the same samples that were also used for mitochondrial capture. As during enrichment, we used the closest possible reference to align the reads. To map reads from species not closely related to any reference genome, 2 or 3 genomes were set as reference. We used the iterative mapping assembler MIA (Green et al. 2008; Briggs et al. 2009), filtered for unique reads by grouping sequences with the same orientation, start, and compatible end coordinates, and only accepted positions with at least 2-fold unique coverage. This ensures that any given position was present in at least 2 independent mtDNA fragments. Ambiguous positions and those with only 1-fold coverage were called N. Their distribution by sample is shown in Supplementary Table S2. In the 92 mtDNA genomes retained for

phylogenetic analyses (see below), the median number of called nucleotides with only 2-fold coverage was 117, while all other called nucleotides had higher coverage. The pooled data set of all samples generated a total of 76 347 651 merged reads, of which 2 904 450 (3.9%) mapped to a reference mtDNA genome (Supplementary Table S2).

Because no museum samples of the *Cercopithecus hamlyni* species group produced a mitochondrial sequence, we used a high-quality DNA sample from *C. hamlyni* of unknown geographic origin (Supplementary Table S1). We first amplified the mitochondrial genome in 2 long-range PCR fragments using the same protocol as for the bait construction. We then sequenced the genome either directly using long-range PCR products as template (Wertheim and Worobey 2007), or after additional amplification with internal primers (Supplementary Table S3). Purified PCR products were sequenced on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) in both directions using a standard protocol. We obtained an almost complete sequence of the genome with 15 093 bp sequenced.

#### Testing for Contamination

Contamination with human DNA is possible with museum specimens and, if present at a high rate, can obscure the phylogenetic signal and lead to erroneous conclusions. To evaluate the prevalence of human contamination, we applied the method detailed in Kircher (2012). It relies on the differences in alignment scores obtained by aligning each read to a set of closely related sequences and to potential homologous contaminant sequences. Reads from each sample were mapped against the 3 Cercopithecini mtDNA reference genomes (as above) using BWA V0.6.1 (Li and Durbin 2009). In a second step, the human “Cambridge reference” mtDNA genome was added to the guenon references and the mapping was repeated. BWA parameters were kept identical in both mappings. The ContTestBWA.py script (Kircher 2012) was used to evaluate the reported alignments for each read and to return the number of endogenous, noninformative and contaminant reads. Confidence levels were calculated in R assuming binomial distribution.

#### Taxon Sampling

**MtDNA sequences.**—In addition to the mitochondrial genomes obtained from museum specimens (Supplementary Table S2), we also included 3 samples (*C. hamlyni*, *C. mitis monoides*, and *E. patas*—the latter 2 used for enrichment and mapping, see above), whose complete mitochondrial sequences were determined from high-quality DNA samples. We downloaded sequences of 6 *Chlorocebus* species from GenBank (Supplementary Table S1) to be included in the phylogenetic analyses. In addition, we generated

mitochondrial sequences for 4 outgroup taxa from museum-preserved samples (gelada baboon, gray-cheeked and collared mangabeys, and drill). We also downloaded 8 additional complete mitochondrial genomes for further outgroup species from GenBank (Supplementary Table S1).

#### Data Partitioning

Alignment of the 110 mitochondrial sequences (Cercopithecini and outgroup) was performed with MAFFT v6.811b (Katoh and Toh 2008; Katoh et al. 2009). Using the annotated genome of *Chlorocebus tantalus* (EF597502) as a reference, we partitioned the aligned genomes into protein-coding genes, tRNAs, rRNAs, and noncoding fragments (including the origin of replication and the hypervariable region). We further partitioned the protein-coding genes into first, second, and third codon positions using Split Codons (Stothard 2000). To evaluate the best partitioning scheme and to determine substitution models for each partition, we used PartitionFinder v. 0.9 (Lanfear et al. 2012). We divided the mtDNA alignment into 42 subsets: all tRNAs (each individual tRNA alignment was too short to be informative), 12s rRNA, 16s rRNA, and the 13 protein-coding genes each subdivided into first, second, and third codon position. Then, because we intended to use RaxML (Stamatakis 2006) and BEAST (Drummond and Rambaut 2007) for phylogenetic reconstructions, we ran PartitionFinder twice: (i) only considering models of nucleotide evolution as available in RaxML and (ii) considering all 56 possible models. The selected partitioning scheme with the highest Bayesian information criterion differed slightly for RaxML and BEAST data (Supplementary Table S4).

#### Phylogenetic Reconstruction and Dating

We performed maximum likelihood (ML) and Bayesian inference (BI) phylogenetic analyses for mitochondrial sequences, adopting the best-fitting model for each partition. ML was performed with RaxML v7.2.8 (Stamatakis et al. 2008) on the CIPRES cluster (Miller et al. 2009). We used BEAST v1.6.1 (Drummond and Rambaut 2007) for BI analyses and coestimated the divergence times from the sequence data. The different partitions shared the same tree topology. A starting tree generated with ML was provided and the prior tree distribution was generated using a Yule process. The final trees were based on 5 independent runs of  $2 \times 10^7$  generations each. Tree sampling was performed every  $1 \times 10^3$  generations. We used Tracer v1.5 (Rambaut and Drummond 2009) to assess convergence among runs and to determine the number of burn-in steps, which were subsequently discarded. The number of trees was thinned to about 10 000 with LogCombiner v1.6.1 and the maximum

credibility trees were estimated with TreeAnnotator v1.6.1.

We specified 2 calibration points relying on undisputed fossils within the outgroup taxa. We used the fossils of *Macaca libyca* (6 Ma) (Stromer 1920) and *Microcolobus tugenensis* (10 Ma) (Benefit and Pickford 1986), which are the oldest African representatives of the respective taxonomic groups (Jablonski 2002). Morphological and stratigraphic data for both species have been reviewed recently in the framework of newly discovered cercopithecoid fossils (Benefit et al. 2008; Gilbert et al. 2010). *Microcolobus tugenensis* provided the minimum age for the Colobinae node, whereas we performed 2 independent analyses to account for the uncertainty in the placement of *M. libyca* as either a stem or crown macaque. For *M. tugenensis* and *M. libyca* as crown macaque, we chose normally distributed priors with a standard deviation (SD) of 1 myr, so that the lower boundary matched the fossil record date. Importantly, all bounds were “soft,” so that the probability of divergence time being outside the bounds was above zero. When *M. libyca* was used to inform the divergence age between macaques and other Papionini, we used instead a lognormal prior with the minimum age of 6 myr, SD of 0.5, and an unconstrained lower bound. This placed the median of the distribution at 9.5 myr. We further put a 25-myrr constraint on the root of the tree, which was derived from previous molecular estimates for Cercopithecidae (Raum et al. 2005; Zinner et al. 2009; Perelman et al. 2011). To account for the uncertainty associated with this age, we applied large and “soft” bounds of 5 myr.

Because prior calibration densities can influence posterior age estimates (Heled and Drummond 2012), we performed an analysis on “empty alignments” with the same priors but no data. We compared the results derived by sampling only from the priors with those obtained for the posterior distributions estimated on sequence data.

Finally, to account for a potential effect of incomplete genome sequences on branch length and divergence time estimates, we repeated ML and BI analyses for a data set of 95 mtDNA sequences that contained only genomes with more than 13 000 sequenced nucleotides. This corresponds to at least 78% completeness.

#### Evaluating Congruence between Nuclear and mtDNA Phylogenetic Trees

We obtained the alignment of nuclear sequences from Perelman et al. (2011). We pruned the published phylogenetic tree so that it contained only the taxa that were also present in our mtDNA data set. The pruned nuclear data set consisted of 17 guenon taxa and 11 outgroup taxa (Supplementary Table S5). Guenons present in the nuclear data set contain representatives of both the arboreal and the terrestrial clades (see below). To evaluate the concordance between phylogenetic trees estimated with mitochondrial and nuclear sequences,

we performed the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa 1999) implemented in the R package phangorn (Schliep 2011). Because we observed differences between the 2 phylogenetic trees within the ingroup and also within the outgroups, we repeated the SH test for a data set with only Colobini as outgroup, but the results did not differ (data not shown).

#### Ancestral Range Reconstruction and Range Overlap Analysis

Since the mtDNA-based phylogenetic trees encompassed the broadest taxon sampling, they formed the basis for our phylogeographic analyses, relying on the chronogram from BEAST. The software Lagrange v 20110117 (Ree et al. 2005; Ree and Smith 2008) was used to perform a ML analysis of dispersal–extinction–cladogenesis for the Cercopithecini. We reduced our sampling to contain a single representative of each taxon, except for *Chlorocebus pygerythrus*, for which we kept 2 specimens from geographically distant locations (South Africa and Tanzania). Taking into account known guenon distribution (Jones et al. 2009), we defined 11 geographical regions that roughly follow the previously classified biogeographical provinces (Udvardy 1975) and ecoregions (Olson et al. 2001) (Fig. 1 and Supplementary Fig. SI). Each taxon was assigned to a maximum of 3 geographical regions—those that covered the largest proportion of its distribution range. The distribution of the ancestor was confined to a maximum of 2 regions. We excluded a number of unrealistic range combinations and only allowed continuous composite ranges in the Lagrange analysis (Supplementary Table S6).

Range overlap analysis was performed following Barraclough and Vogler (2000) by plotting the degree of range overlap against time of divergence. Distribution data for all available guenon taxa were obtained from the PanTHERIA database (Jones et al. 2009). The presence/absence of each taxon was evaluated in 25 km<sup>2</sup> grid cells. We also performed a range overlap analysis using the nuclear data from Perelman et al. (2011), transforming branch lengths with a semiparametric method based on penalized likelihood (Sanderson 2002).

#### Diversification Rates

We used the birth–death likelihood (BDL) method implemented in the R package LASER (Rabosky 2006) to test for different models of diversification. Analyses were conducted on 3 sets of BEAST chronograms that accommodated various taxonomic schemes (Supplementary Information). Akaike information criterion (AIC) scores were calculated for each of the 6 following models—2 rate-constant models: pure birth (constant speciation and no extinction) and birth–death (constant speciation and extinction); 2 density-dependent models: DDL (with logistic



component) and DDX (exponential component); and 2 multivariate variants of the pure birth model: Yule-2-rate and Yule-3-rate. The 2 latter models allow 1 or 2 rate shifts, respectively. The best-fit model was selected by comparing the difference in AIC between the best rate-constant model and each of the rate-variable models for the BEAST maximum credibility trees (SI).

## RESULTS

### *Mitochondrial Genomes from Museum Specimens*

Our sampling comprised all but one of the 63 species and subspecies described in Grubb et al. (2003), with *Chlorocebus djamjamsensis* missing from the museum collections. *Cercopithecus pogonias schwarzi*, which is recognized by Wilson and Reeder (2005) as a separate taxon, was also included here.

Of the 120 sequenced samples, 92 produced complete or partial mitochondrial genome sequences (Supplementary Table S2). We failed to obtain mtDNA sequences from museum-preserved *C. hamlyni*, and therefore included a *C. hamlyni* mitochondrial genome sequence that was derived from a blood sample. Taken together, we obtained mitochondrial genomes for 57 of the 63 guenon taxa—an almost complete representation of the taxonomical variation present within Cercopithecini (Supplementary Table S1). Importantly, each species group was represented by at least one of its members.

The median coverage for each mitochondrial genome was 24-fold (mean=91.4-fold), which is higher than the average coverage obtained from high-quality samples with other NGS protocols (Chan et al. 2010; Gunnarsdottir et al. 2011). We found that coverage along the mtDNA genome was strongly correlated with the local GC content (Pearson's  $r=0.53$ ,  $P < 2.2e^{-16}$ , Supplementary Fig. S2), as has been previously reported (Green et al. 2008; Briggs et al. 2009; Krause et al. 2010).

Samples collected from museum specimens, particularly from exposed surfaces, are prone to contain high level of exogenous DNA from different contamination sources (e.g., bacteria and fungi that colonized the specimen). Despite this complication, the enrichment protocol was highly efficient; an average of 3.9% (but up to 62.4%) of sequenced fragments were mapped to the reference genomes (Supplementary Table S2). Human mitochondrial fragments shed during handling of specimens might hybridize to the probe due to the close phylogenetic relationship between guenons and humans. However, only about 2.3% of all mapped fragments (0.09% of total fragments) were potentially derived from human (Supplementary Table S2). Given that we require at least 2-fold coverage, this frequency amounts to an error rate of  $(0.023 \times 0.023) = 5.3 \times 10^{-4}$  per nucleotide position. Since in a typical genome only about 117 positions had exactly 2-fold coverage, the actual error rate is much smaller and almost negligible.

We tested for the influence of age, weight, and specimen part from which the sample was collected on the quantity of endogenous DNA using a linear regression model in R (adjusted  $R^2=0.1458$ ,  $F=1.642$  on 21 and 58 df,  $P=0.07048$ , Materials and Methods and Supplementary Information, Results). Neither age (spanning more than 100 years) nor weight of the sample (ranging from 10 to 253 mg) had an effect on DNA recovery ( $P=0.89$  and  $0.97$ , respectively). There was also no difference in the performance of the samples from different sources (skin, dried tissue from skeleton and skull, teeth, bone material, etc., pairwise  $t$ -test with adjusted  $P=1$ ).

### *Phylogeny of Guenons*

*MtDNA phylogenetic trees.*—Phylogenetic reconstructions based on mtDNA with ML and BI methods resolved congruent and highly supported tree topologies (Figs. 1 and 2). Overall, the topologies agree with an earlier study based on limited sampling (Chatterjee et al. 2009). Species groups are monophyletic, with the notable exception of the *C. diana* and *Cercopithecus preussi* species groups (Fig. 1 and Supplementary Fig. S3, see below). *Cercopithecus dryas* and *C. hamlyni* species groups are represented by a single sample, and thus their monophyly could not be ascertained (Figs. 1 and 2 and Supplementary Fig. S3). The genus *Cercopithecus* is paraphyletic (see also Chatterjee et al. 2009); *Erythrocebus* is placed within *Cercopithecus*, whereas *C. hamlyni* and *Cercopithecus solatus* are more closely related to the genus *Chlorocebus* (Figs. 1 and 2 and Supplementary Fig. S3).

We observed some deep within-species splits, which merit further investigation (Fig. 2 and Supplementary Fig. S3). For example, in *E. patas*, the lineage from Sudan shows an early separation from that from Ghana. The difference between eastern and western populations of *E. patas* is also supported by morphology (Kingdon 2005). Similarly deep divergences are also found within *Cercopithecus cephus*. Within the *C. mitis* group, we found *C. mitis mitis* to be sister to *Cercopithecus nictitans* instead of clustering with other *C. mitis*. It is noteworthy that *C. mitis mitis* is the westernmost representative of *C. mitis* species and occurs in geographic proximity to *C. nictitans*. Consistent with previous finding (Wertheim and Worobey 2007), we also identified deep divergence within *Chl. pygerythrus*, separating lineages from South Africa and Tanzania.

The almost complete taxonomic sampling allowed us to infer the phylogenetic position of the enigmatic *C. dryas*, whose evolutionary relationships have been disputed (Grubb et al. 2003). Based on mitochondrial sequences, we found *C. dryas* nested within the *Chlorocebus* (African green monkeys), with *Chl. sabaeus* as sister species.

*MtDNA versus nuclear DNA data.*—We compared the phylogenetic trees derived from mitochondrial and





nuclear data. The SH test indicated that there were significant differences between these trees (Table 1). Most of these differences refer to the placement of taxa belonging to the terrestrial group: *Cercopithecus lhoesti*, *E. patas*, and the genus *Chlorocebus* (Fig. 3). For example, *C. lhoesti* and *E. patas* are sisters to each other and to *Chlorocebus* in the nuclear phylogenetic tree, whereas they are placed on different branches within the arboreal group in the mtDNA-based phylogenetic tree. Thus, the mtDNA phylogenetic tree does not support a single terrestrial group; instead, it favors multiple transitions to terrestriality (Figs. 1 and 3).

TABLE 1. SH test results on nuclear and mitochondrial phylogenetic trees

Sequence data	Tree	−ln L	SH test P-value
mtDNA	mtDNA	−109219.5	0.5
mtDNA	Nuclear	−61720.8	< 0.00001
Nuclear	mtDNA	−113033.1	< 0.00001
Nuclear	Nuclear	−60699.9	0.5

### Timing of Diversification

**MtDNA phylogeny.**—The posterior distributions of the calibration nodes for *Macaca* and Cercopithecidae obtained in the BEAST analyses with sequence data were distinct from the respective assigned prior distributions, retrieved from the analyses on empty alignments (Supplementary Fig. S4). This observation was independent of the calibration scheme (*M. libyca* as crown or stem macaque) and indicated that the inferred dates were indeed informed by the data. The posterior distribution of the Colobinae node was not different from the prior and was therefore uninformative (Supplementary Fig. S4). The divergence times estimated from the full mtDNA data set or the reduced data set (samples with more than 13 kb sequences) were nearly identical (Supplementary Fig. S5); we therefore present here the results of the larger data set. Also, whether we used the *M. libyca* fossil as stem or as crown macaque had a minor effect on the divergence time estimates (Supplementary Fig. S6). When the fossil was used to inform the split between macaques and other Papionini, the inferred divergence dates were slightly older than

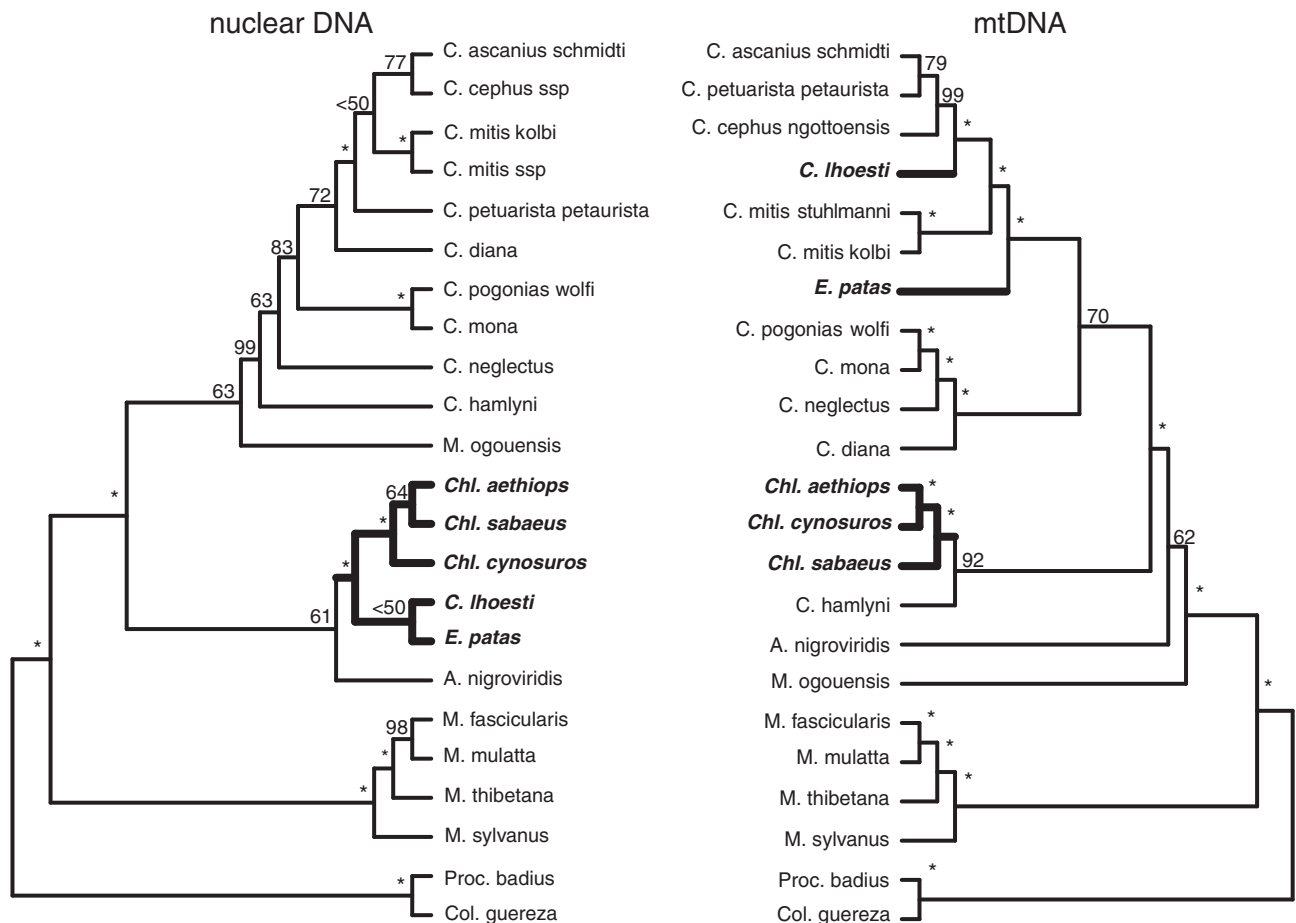


FIGURE 3. Nuclear (Perelman et al. 2011) versus mitochondrial trees for 17 guenons species belonging to the arboreal and the terrestrial clades. Numbers at the nodes indicate ML bootstrap support, nodes with support of 100 are labeled with asterisks. Taxa from the “terrestrial clade” (Tosi et al. 2004; Tosi 2008) are highlighted in bold.

when it was used as the crown node of macaques. The strongest effect was on the age of the root (difference of ca. 0.5 myr) and the relatively older nodes. Differences diminished toward the tips, so that the age of the ingroup differed by 0.3 myr, whereas the ages of species group radiations differed by approximately 0.06 myr.

The split between Colobinae and Cercopithecinae was dated to approximately 19 Ma (22.9–15.1; 95% highest posterior density [HPD] interval), the split between Papionini and Cercopithecini to approximately 12.3 Ma (15.0–9.6 95% HPD). The first radiation within Cercopithecini was estimated to have occurred approximately 9.6 Ma (11.7–7.5 95% HPD). We identified 4 radiation events at different taxonomic levels (Fig. 1). First, between 9.6 and 9.3 Ma (11.7–7.3 95% HPD), major genera were formed: *Miopithecus* and *Allenopithecus* split from all other guenons. Second, between 7.4 and 6.8 Ma (9.1–5.3 95% HPD), most species groups originated: The *C. hamlyni* species group with the representative *C. hamlyni*; the *C. aethiops* species group—*Chlorocebus*; the clade with *C. diana*, *Cercopithecus roloway*, *C. neglectus*, and the *Cercopithecus mona* species group; the genus *Erythrocebus*; and finally, the clade containing *C. mitis* and *C. cephus* species groups (Fig. 1). Third, between 2.4 and 2.1 Ma (3.0–1.7 95% HPD), diversification occurred within the species groups. Finally, within the past 1 myr (1.9–0.1 95% HPD), most subspecies evolved, notably within *Cercopithecus ascanius*, the eastern radiation of *C. mitis*, and within *C. pogonias*.

**Differential diversification through time.**—We used mtDNA phylogenetic trees to estimate changes in diversification rates during guenon evolution. Because the results were consistent for all 3 taxonomic scenarios, we present here only the results for the complete sampling at the subspecies level (Supplementary Information, Results; taxonomic scenario C). The BDL model with the lowest AIC score was the Yule-3-rate model, with 2 shifts in diversification rate (Supplementary Table S7 and Supplementary Fig. S7). This model also consistently produced the lowest AIC scores and provided a significantly better fit than the best rate-constant model ( $\Delta\text{AIC}_{\text{RC}} = 13.75$ ,  $P < 0.01$ ). This result was not affected by incomplete taxon sampling (Supplementary Table S7). The ML estimate of the first shift in diversification rate was 2.77 Ma, which roughly corresponds to the radiation event within the species groups. At this point in time, there was a significant increase of 2.8 times ( $r_2/r_1$ , Supplementary Table S7) in the rate of diversification. Later on, approximately 0.44 Ma, a strong decrease of over 7-fold was identified in diversification rate. This apparent slowdown may be an artifact and due to the presence of young lineages that are still too similar to be recognized as separate taxa (Etienne and Rosindell 2012).

#### Geography of Speciation

Range overlap increased with time for both mitochondrial and nuclear data, which is consistent

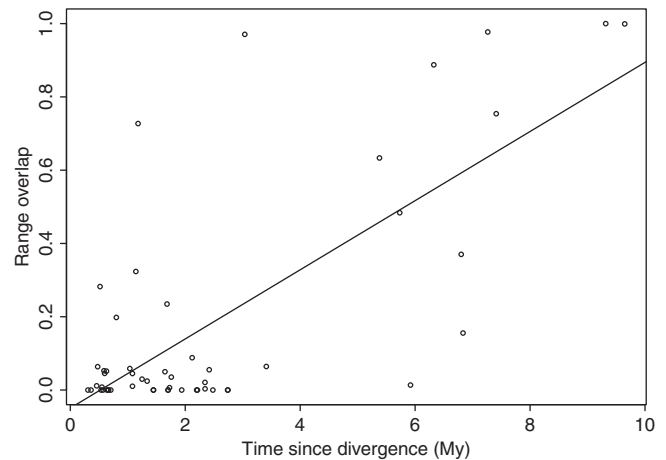


FIGURE 4. Geography of speciation based on mtDNA data. Range overlap plotted against divergence time shows the pattern predicted when species splits are driven by geographic isolation (see text for details). The regression line is plotted in black. See Supplementary Figure S8 for the corresponding analyses on nuclear data.

with a scenario of allopatric speciation (Fig. 4 and Supplementary Fig. S8). Based on mtDNA, the Y-intercept of the regression between range overlap and node age for the entire group of Cercopithecini was  $-0.05$  and not significantly different from 0 ( $P = 0.25$ ), while the slope was positive and significantly different from 0 (adjusted  $R^2 = 0.53$ ,  $P = 4.7 \times 10^{-10}$ ).

The reconstruction of ancestral geographic ranges along the phylogenetic tree indicated that guenons most likely originated in western Africa (Fig. 1 and Supplementary Information, Results). Although several lineages, notably the *C. cephus* group, remained in this region, others dispersed eastward toward the Albertine Rift Valley and westward to the upper Guinean forests (SI Results). The Lower and Upper Guinean regions repeatedly played a role in guenon diversification at different points in time. Our reconstructions also indicated that most dispersal events to the east went around the Congo basin. This supports the idea that the Congo basin served as a barrier for forest-dwelling taxa during the past 10 myr.

The species groups show group-specific dispersal patterns across the African continent. The dispersal pattern of the *C. cephus* and *C. mona* groups are similar. Both originated between Upper and Lower Guinea (Fig. 1). In each case, some taxa went westward across the Dahomey Gap and dispersed further west in a stepwise manner. Other taxa went eastward, first taking the northern route around the Congo River. Upon entering the Congo basin they diversified within it (e.g., subspecies of *C. ascanius* less than 1 Ma).

The *C. mitis* group originated further east compared with the *C. cephus* and *C. mona* groups. Some taxa dispersed to the west but not further than Lower Guinea. Another lineage went south around the Congo basin toward Zambia. The Albertine Rift Valley split this lineage into 2, one confined to the west of the Rift and the

other to the east African coast (Fig. 1). The eastern lineage experienced rapid diversification within the past 1 myr.

Finally, we found that *C. preussi*/*C. lhoesti* originated in western Africa and dispersed northeast around the Congo basin, as suggested by Tosi (2008). A similar dispersal route was taken by the widespread *Cercopithecus aethiops* group (now genus *Chlorocebus*).

## DISCUSSION

### *Guenon mtDNA Phylogeny*

Although many features of the tree topologies, including the divergence times of the major groups, were comparable between the mtDNA-based phylogenetic tree and previously published nuclear and sex-chromosomal phylogenetic trees (Raaum et al. 2005; Tosi 2008; Fabre et al. 2009; Zinner et al. 2009; Perelman et al. 2011), there were also significant discrepancies. We discuss these in detail below.

The almost complete taxonomic sampling allowed us to evaluate the so far unknown aspects of guenon phylogenetic relationships. For instance, mtDNA sequences place *C. dryas* as a member of the *C. aethiops* species group. This goes against previous ideas that *C. dryas* is a representative of the *C. diana* species group. Although no nuclear data are available to corroborate our mtDNA-based phylogenetic inference, a close relationship of *C. dryas* with the members of the African green monkeys is also supported by similarities in feeding behavior, locomotion, cranial, and dental characteristics (Kuroda et al. 1985; Groves 1989).

### *Guenon Radiation as Inferred by mtDNA Analysis*

By applying the range overlap test, we showed that guenons predominantly speciated in allopatry and that geographic isolation played a pivotal role in their diversification. Dating the mtDNA phylogenetic trees, we inferred that guenons radiated during the Late Miocene and Pleistocene—a time period known for increased rates of speciation in mammals (Johnson et al. 2006; Krause et al. 2008; Chan et al. 2010; Fulton and Strobeck 2010) and pronounced climatic fluctuations (Bonnefille 2010). Since the majority of guenons are forest-dependent, the dynamic history of African forests ultimately had a dramatic effect on their evolution. Because gene divergence usually precedes population divergence (e.g., Jennings and Edwards 2005), the actual speciation events might be younger than the dates inferred for mtDNA data. Although it is difficult to pinpoint the specific climatic events, the past 10 myr were characterized by dynamic and repeated fluctuations in local and global forest cover in Africa (Bonnefille 2010). From around 16 Ma, forest cover diminished throughout Africa and savannah expanded by 8 Ma in western and eastern Africa (Jacobs 2004). Furthermore, 3 peaks of grass pollen prevalence were observed in marine core at 10, 7, and 2.8 Ma (Bonnefille 2010)—loosely coinciding

with the time periods of 3 important speciation events in guenons (Fig. 1). The latter aridification is linked to the onset of the “Northern Hemisphere Glaciation” (Haug et al. 2005) and coincides with the major increase in guenon diversification rate, as indicated by the LASER analyses.

Within the Congo basin, changing courses of rivers might have played a more pronounced role than the forest cover fluctuations. Many of the subspecies in *C. ascanius* and *C. pogonias* are indeed confined to interfluvial areas. The Congo basin is a region of low altitudinal profile and was suggested to have been occupied by a large water body as recently as the Pliocene (Goudie 2005). This might have made it impenetrable to guenons and other species and forced dispersal around the basin. If the course of the Congo River and its tributaries was established only recently, it would explain the young separation ages of the aforementioned taxa.

### *Possible Influence of Other Factors on Guenon Radiation*

In addition to geographic isolation, other factors might have contributed to guenon diversity. For example, guenons are extremely colorful and their species-specific facial patterns are thought to contribute to species recognition (Kingdon 1980, 1988). Guenons are also well known for their diverse acoustic repertoire, which was partly shaped by sexual selection (Zuberbühler 2003) and hence might enforce species boundaries. However, morphological affinities (Kingdon 1980) and species relationships based on vocalization (Gautier 1988) stand in stark contrast to both mtDNA and nuclear DNA-based phylogenetic trees. Although acoustic affinities in general support the separation into arboreal and terrestrial forms, the relationships within the terrestrial group are not well supported (Gautier 1988). In addition, greater call diversity was observed within a *Chl. aethiops* group in different habitats than between *Chl. aethiops* and *E. patas* in the same habitat (Enstam and Isbell 2002).

Guenons also display great differences in behavior and social structure. The ecology of guenons with their diverse choice of habitat and dietary preferences (Butynski 2003) could have played an important role in further enhancing differentiation between populations.

Furthermore, guenons are well known for their large cytogenetic differences (Dutrillaux et al. 1988). Frequent chromosomal rearrangements were held responsible for the taxonomic diversity in this primate group. However, molecular trees and chromosomal trees (Moulin et al. 2008) show multiple discrepancies to each other and hybrids reported from the wild clearly disrespect chromosomal boundaries (e.g., *C. ascanius* with 66 chromosomes and *C. mitis* with 70 chromosomes). Discordance between molecular and cytogenetic phylogenetic trees is not infrequent, even within primates (Rumpler et al. 2011). Overall, speciation in guenons is bound to be a complex process influenced by many factors.



*Phylogenetic Conflicts between Mitochondrial and Nuclear Markers*

**Divergence dates.**—The divergence dates for major phylogenetic groups derived from nuclear (Perelman et al. 2011) and mtDNA data (Results and Supplementary Fig. S3) were similar, with overlapping confidence intervals (e.g., split between Colobinae and Cercopitheciinae ca. 17.6 Ma [21.5–13.9 95% HPD], emergence of Cercopitheciini ca. 11.5 Ma [13.9–9.2 95% HPD], and the first radiation within Cercopitheciini 8.2 Ma [10.0–6.6 95% HPD]). Some ingroup taxa also showed similar separation dates (e.g., the split between *C. neglectus* and the *C. mona* group ca. 5.2 Ma [6.6–3.9 95% HPD] for nuclear data and 5.4 Ma [6.6–4.2 95% HPD] for mitochondrial data). For others, however, the dates estimated with nuclear data were younger than those estimated with mtDNA data and the confidence intervals did not overlap: for example, diversification of *C. cephus* and *C. mitis* species groups less than 3 Ma (3.9–1.9 95% HPD) with nuclear data compared with more than 5.7 Ma (7.0–4.5 95% HPD) with mtDNA data. Due to the differences in mtDNA and nuclear tree topologies, many nodes could not be compared.

**Topology.**—We uncovered strong discrepancies between nuclear and mitochondrial phylogenetic trees. The mitochondrial phylogenetic trees indicated that the genus *Cercopithecus* is paraphyletic (Results and Figs. 1 and 2 and Supplementary Fig. S3), whereas nuclear data supported the monophyly of this group (Perelman et al. 2011). The direct comparison of taxa for which both nuclear and mitochondrial data were available further indicated that most discrepancies involved the placement of species from the “terrestrial group” (Tosi et al. 2004) (Fig. 3).

Several factors can be invoked to explain the discrepancy between mitochondrial and nuclear phylogenetic trees: paralogous relationships of studied genes, incomplete lineage sorting (ILS), and hybridization/introgression (Funk and Omland 2003).

It is rather unlikely that gene duplications had a major effect on the nuclear phylogeny. This is because congruent phylogenetic patterns were derived for autosomal, X- and Y-chromosomal loci independently (Perelman et al. 2011). Being a single locus, mitochondrial genomes are not expected to have paralogous relationships, with the notable exception of nuclear copies of mitochondrial DNA (numts). Given our enrichment, sequencing, and mapping protocol, it is highly unlikely that numts could obscure the phylogenetic relationships inferred from mitochondrial genomes. First, each mammalian cell contains multiple mtDNA genomes (estimates range from 220 to a few thousand copies per cell depending on the cell type [Robin and Wong 1988; Legros et al. 2004]). Second, numt enrichment was observed in PCR-based studies that used universal primers across distantly related species (Thalmann et al. 2004). The preferential numt amplification was obtained because of the slower rate

of sequence evolution in the nucleus compared with the mitochondria. This problem is overcome by sequencing all mitochondrion-like fragments in our study in an unbiased manner. Third, mtDNA was shown to be better preserved in old DNA samples than nuclear DNA (Green et al. 2008, 2010) as indicated by longer fragment size of mitochondrial (70 nt) than nuclear (48 nt) DNA. Thus, we expect to predominantly capture mtDNA. Finally, let us evaluate the probability of sequencing a nuclear copy. To do that, we need to estimate the prevalence of numts in a typical guenon genome. No such information is available to date. However, a recent study investigated the accumulation of numts on the macaque lineage since the divergence from human (note here that the divergence time between humans and guenons will be the same). It found 101 numts private to the macaque and a total of 434 numts in the macaque genome (Hazkani-Covo 2009). The total length of macaque numts is 261.6 kb (Hazkani-Covo et al. 2010), resulting in the average numt length of 602.8 bp. Since the divergence from human, we would thus expect about 60.9 kb of macaque-specific numts in the macaque genome. The macaque genome size is about 3 Gb and a typical ratio of nuclear to mtDNA fragments in old DNA samples is about 200 (Green et al. 2009). Thus, the number of numt-derived fragments for each mtDNA fragment will be  $200 \times 60.9 \times 10^3 / 3 \times 10^9 = 0.00406$ . This corresponds to the ratio of numts to the total pool of numts and mtDNA of  $0.00406 / (0.00406 + 1)$ . Given that we require at least 2-fold coverage, this translates into a numt-derived error rate of  $(0.00404)^2 = 1.6 \times 10^{-5}$  and is therefore negligible.

With the data at hand, we cannot distinguish whether hybridization or ILS is the cause of the observed incongruence between nuclear and mitochondrial phylogenetic trees. More data and a targeted study design are needed to address this question in full. However, given the wealth of information on guenon hybrids in the wild—even between lineages that show old divergence dates—it is tempting to speculate that hybridization has played a role in guenon speciation. For instance, *C. mitis* and *C. ascanius* are known to hybridize (Aldrich-Blake 1968; Struhsaker et al. 1988). They show a mitochondrial divergence time of more than 5 myr, although the nuclear divergence time is smaller (>3 myr). The recently reported intergeneric hybrid between *C. mitis* and *Chl. pygerythrus* (de Jong and Butynski 2010) extends the divergence time between hybridizing parental taxa to more than 7 myr for mtDNA genomes and more than 8 myr for nuclear genomes. Given that hybridization between these distantly related taxa can still occur today, we may speculate that repeated hybridization and extensive backcrossing among their ancestors might have contributed to the observed discordance between nuclear and mitochondrial phylogenetic trees.

It is noteworthy that divergence dates inferred with nuclear genomes are frequently younger than those based on mtDNA genomes ([Perelman et al. 2011, p. 591; Tosi et al. 2005, p. 500] and Supplementary Fig. S3). This is particularly evident for the inferred

separation between *C. solatus* and *C. lhoesti*/*C. preussi* (>7 myr based on mtDNA genomes and only ca. 2 myr based on sex-chromosomal sequences [Tosi 2008]). Furthermore, nuclear genomes provide a clear signal for grouping species that have similar habitat preferences (terrestrial vs. arboreal). A possible explanation for this observation is a scenario in which an initial separation driven by geographic isolation (reflected by mtDNA divergence dates) was followed by secondary contact with male-mediated gene flow. Most guenon taxa for which dispersal behavior was studied show male-biased dispersal (Schoof et al. 2009). If intensive backcrossing with more invading males occurred, this social system might have promoted nuclear swamping, as was recently proposed for colobine monkeys (Roos et al. 2011). Having similar habitat preferences could have facilitated this process. If this is true, the transition to terrestriality could have occurred repeatedly. In concordance with mtDNA phylogenetic trees, 3 independent origins of terrestriality were suggested for *C. lhoesti*, *E. patas*, and *Chl. aethiops* based on postcranial adaptations (Gebo and Sargis 1994).

It is now widely recognized that hybridization has contributed to the evolution of mammals, including primates (Arnold and Meyer 2006). For example, in colobus monkeys, reticulate evolution has been invoked to explain differences in nuclear versus mitochondrial-based phylogenetic trees (Roos et al. 2011). In baboons, fertile hybrids are known from the wild (Nagel 1973; Bergman et al. 2008), and hybridization was put forward to explain discordances between mitochondrial phylogenetic trees and morphology (Zinner et al. 2009; Keller et al. 2010). Anatomically modern humans show signature of ancient admixture with Neanderthals and Denisovans (Green et al. 2010; Meyer et al. 2012). We suggest the same applies to guenons.

#### Museomics

We show here that museum-preserved samples constitute an extraordinarily rich resource for DNA studies. We have been successful in recovering mitochondrial genome sequences from as little as 10 mg of biological material. The high success rate of approximately 77% did not depend on the source of the sample, indicating that any part of the specimen can be sampled. This is particularly important for museum scientists, as it removes the need for destructive tooth or bone sampling.

NGS coupled with target enrichment methods (Maricic et al. 2010; Mason et al. 2011) is a reliable and cost-effective method, which makes it possible to obtain complete mitochondrial genome sequences even from difficult source materials. Particularly given the rapid development in this field and the availability of new methods with increased throughput and lower sequencing costs (Glenn 2011), this approach is likely to become standard in the near future. It helps to overcome many biases and limitations inherent to museum-preserved specimens, such as low DNA content and

high levels of contamination. Furthermore, because of an unbiased sequencing of all mitochondrion-like fragments, numts will be swamped out by the overwhelming majority of genuine mitochondrial sequences. In this study, we have illustrated the power of NGS for retrieving genomic data from museum-preserved specimens and expect that museum collections will be “unlocked” and millions of specimens studied in the near future.

#### SUPPLEMENTARY MATERIAL

Supplementary material, including data files and online-only appendices, can be found in the Dryad data repository at <http://datadryad.org/>, doi:10.5061/dryad.7k14q. Tree files and alignments have been submitted to TreeBASE under preliminary Study Accession URL: <http://purl.org/phylo/treebase/phylows/study/TB2:S13808?x-access-code=a8ab934471238214b271aa112d8df166&format=html>.

#### FUNDING

This study was funded by SYNTHESYS-BE-TAF-150, the European Research Council, the Leverhulme Trust, the Royal Society, Natural Environment Research Council, and FP7 Marie Curie Actions grants.

#### ACKNOWLEDGMENTS

We are grateful to following people for providing samples, support, and advice during the course of the project: Svante Pääbo, Tomislav Maricic, Martin Kircher, Matthias Meyer, Qiaomei Fu, Robert Lanfear, Damien Caillaud, Ed Green, Paula Jenkins, Alban Lemasson, Christian Roos, Wim Wendelen, Kowiyou Yessoufou, Anamaria Necșulea, and David Brawand. We thank Ronald DeBry, Elizabeth Jockusch, and anonymous reviewers for their helpful comments.

#### REFERENCES

- Aldrich-Blake F.P.G. 1968. A fertile hybrid between two *Cercopithecus* species in the Budongo Forest, Uganda. *Folia Primatol.* 9:15–21.
- Arnold M.L., Meyer A. 2006. Natural hybridization in primates: one evolutionary mechanism. *Zoology* 109:261–276.
- Barraclough T.G., Vogler A.P. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* 155:419–434.
- Benefit B.R., McCrossin M.L., Boaz N.T., Pavlakis P. 2008. New fossil *Cercopithecoids* from the Late Miocene of As Sahabi, Libya. *Garyounis Sci. Bull. Special Issue, No. 5*:265–282.
- Benefit B.R., Pickford M. 1986. Miocene fossil *Cercopithecoids* from Kenya. *Am. J. Phys. Anthropol.* 69:441–464.
- Bergman T.J., Phillips-Conroy J.E., Jolly C.J. 2008. Behavioral variation and reproductive success of male baboons (*Papio anubis* × *Papio hamadryas*) in a hybrid social group. *Am. J. Primatol.* 70:136–147.
- Bonnefille R. 2010. Cenozoic vegetation, climate changes and hominid evolution in tropical Africa. *Glob. Planet. Change* 72:390–411.

- Briggs A.W., Good J.M., Green R.E., Krause J., Maricic T., Stenzel U., Lalueza-Fox C., Rudan P., Brajkovic D., Kucan Z., Gusic I., Schmitz R., Doronichev V.B., Golovanova L.V., de la Rasilla M., Fordea J., Rosas A., Pääbo S. 2009. Targeted retrieval and analysis of five neandertal mtDNA genomes. *Science* 325:318–321.
- Butynski T.M. 2003. The guenons: an overview of diversity and taxonomy. In: Glenn M.E., Cords M., editors. *The guenons: diversity and adaptation in African monkeys*. New York: Kluwer Academic/Plenum Publishers. p. 3–13.
- Chan Y.C., Roos C., Inoue-Murayama M., Inoue E., Shih C.C., Pei K.J., Vigilant L. 2010. Mitochondrial genome sequences effectively reveal the phylogeny of *Hylobates gibbons*. *PLoS One* 5:e14419.
- Chatterjee H.J., Ho S.Y., Barnes I., Groves C. 2009. Estimating the phylogeny and divergence times of primates using a supermatrix approach. *BMC Evol. Biol.* 9:259.
- de Jong Y.A., Butynski T.M. 2010. Three Sykes's monkey *Cercopithecus mitis* x Vervet monkey *Chlorocebus pygerythrus* hybrids in Kenya. *Primate. Conserv.* 25:43–56.
- Detwiler K.M., Burrell A.S., Jolly C.J. 2005. Conservation implications of hybridization in African Cercopithecine monkeys. *Int. J. Primatol.* 26:661–684.
- Disotell T.R., Raam R.L. 2003. Molecular timescale and gene tree incongruence in the guenons. In: Glenn M.E., Cords M., editors. *The guenons: diversity and adaptation in African monkeys*. New York: Kluwer Academic/Plenum Publishers. p. 25–34.
- Drummond A., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Dutrillaux B., Muleris M., Couturier J. 1988. Chromosomal evolution of Cercopithecinae. In: Gautier-Hion A., Bourliere F., Gautier J.-P., Kingdon J., editors. *A primate radiation: evolutionary biology of the African guenons*. New York: Cambridge University Press. p. 150–159.
- Enstam K.L., Isbell L.A. 2002. Comparison of responses to alarm calls by patas (*Erythrocebus patas*) and vervet (*Cercopithecus aethiops*) monkeys in relation to habitat structure. *Am. J. Phys. Anthropol.* 119:3–14.
- Etienne R.S., Rosindell J. 2012. Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Syst. Biol.* 61:204–213.
- Fabre P.H., Rodrigues A., Douzery E.J.P. 2009. Patterns of macroevolution among Primates inferred from a supermatrix of mitochondrial and nuclear DNA. *Mol. Phylogenet. Evol.* 53: 808–825.
- Finstermeier K. 2010. Revision of primate phylogeny [dissertation]. University of Leipzig.
- Fulton T.L., Strobeck C. 2010. Multiple fossil calibrations, nuclear loci and mitochondrial genomes provide new insight into biogeography and divergence timing for true seals (Phocidae, Pinnipedia). *J. Biogeogr.* 37:814–829.
- Funk D.J., Omland K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Syst.* 34:397–423.
- Gautier J.-P. 1988. Interspecific affinities among guenons as deduced from vocalization. In: Gautier-Hion A., Bourliere F., Gautier J.-P., Kingdon J., editors. *A primate radiation: evolutionary biology of the African guenons*. New York: Cambridge University Press. p. 194–226.
- Gautier-Hion A. 1988. Polyspecific associations among forest guenons: ecological, behavioural and evolutionary aspects. In: Gautier-Hion A., Bourliere F., Gautier J.-P., Kingdon J., editors. *A primate radiation: evolutionary biology of the African guenons*. New York: Cambridge University Press. p. 452–476.
- Gebo D.L., Sargis E.J. 1994. Terrestrial adaptations in the postcranial skeletons of guenons. *Am. J. Phys. Anthropol.* 93:341–371.
- Gilbert C.C., Goble E.D., Hill A. 2010. Miocene Cercopithecoidea from the Tugen Hills, Kenya. *J. Hum. Evol.* 59:465–483.
- Glenn T.C. 2011. Field guide to next-generation DNA sequencers. *Mol. Ecol. Resour.* 11:759–769.
- Goudie A.S. 2005. The drainage of Africa since the cretaceous. *Geomorphology* 67:437–456.
- Green R.E., Briggs A.W., Krause J., Prufer K., Burbano H.A., Siebauer M., Lachmann M., Pääbo S. 2009. The Neandertal genome and ancient DNA authenticity. *EMBO J.* 28:2494–2502.
- Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W.W., Fritz M.H.Y., Hansen N.F., Durand E.Y., Malaspina A.S., Jensen J.D., Marques-Bonet T., Alkan C., Prufer K., Meyer M., Burbano H.A., Good J.M., Schultz R., Aximu-Petri A., Butthof A., Hober B., Hoffner B., Siegemund M., Weihmann A., Nusbaum C., Lander E.S., Russ C., Novod N., Affourtit J., Egholm M., Verna C., Rudan P., Brajkovic D., Kucan Z., Gusic I., Doronichev V.B., Golovanova L.V., Lalueza-Fox C., de la Rasilla M., Fordea J., Rosas A., Schmitz R.W., Johnson P.L.F., Eichler E.E., Falush D., Birney E., Mullikin J.C., Slatkin M., Nielsen R., Kelso J., Lachmann M., Reich D., Pääbo S. 2010. A draft sequence of the neandertal genome. *Science* 328:710–722.
- Green R.E., Malaspina A.S., Krause J., Briggs A.W., Johnson P.L.F., Uhler C., Meyer M., Good J.M., Maricic T., Stenzel U., Prufer K., Siebauer M., Burbano H.A., Ronan M., Rothberg J.M., Egholm M., Rudan P., Brajkovic D., Kucan Z., Gusic I., Wikstrom M., Laakkonen L., Kelso J., Slatkin M., Pääbo S. 2008. A complete neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* 134:416–426.
- Groves C.P. 1989. *A theory of human and primate evolution*. New York: Oxford University Press.
- Grubb P., Butynski T.M., Oates J.F., Bearder S.K., Disotell T.R., Groves C.P., Struhsaker T.T. 2003. Assessment of the diversity of African primates. *Int. J. Primatol.* 24:1301–1357.
- Gunnarsdottir E.D., Li M., Bauchet M., Finstermeier K., Stoneking M. 2011. High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Res.* 21:1–11.
- Haug G.H., Ganopolski A., Sigman D.M., Rosell-Mele A., Swann G.E., Tiedemann R., Jaccard S.L., Bollmann J., Maslin M.A., Leng M.J., Eglinton G. 2005. North Pacific seasonality and the glaciation of North America 2.7 million years ago. *Nature* 433: 821–825.
- Hazkani-Covo E. 2009. Mitochondrial insertions into primate nuclear genomes suggest the use of numts as a tool for phylogeny. *Mol. Biol. Evol.* 26:2175–2179.
- Hazkani-Covo E., Zeller R.M., Martin W. 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet.* 6:e1000834.
- Heled J., Drummond A.J. 2012. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Syst. Biol.* 61: 138–149.
- Jablonski N.G. 2002. Fossil old world monkeys: the Late Neogene radiation. In: Hartwig W., editor. *The primate fossil record*. Cambridge: Cambridge University Press. p. 255–300.
- Jacobs B.F. 2004. Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359:1573–1583.
- Jennings W.B., Edwards S.V. 2005. Speciation history of Australian grass finches (Poephila) inferred from thirty gene trees. *Evolution.* 59:2033–2047.
- Johnson W.E., Eizirik E., Pecon-Slattery J., Murphy W.J., Antunes A., Teeling E., O'Brien S.J. 2006. The late Miocene radiation of modern Felidae: a genetic assessment. *Science* 311:73–77.
- Jones K.E., Bielby J., Cardillo M., Fritz S.A., O'Dell J., Orme C.D.L., Safi K., Sechrest W., Boakes E.H., Carbone C., Connolly C., Cutts M.J., Foster J.K., Grenyer R., Habib M., Plaster C.A., Price S.A., Rigby E.A., Rist J., Teacher A., Bininda-Emonds O.R.P., Gittleman J.L., Mace G.M., Purvis A., Michener W.K. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90:2648–2648.
- Katoh K., Asimenos G., Toh H. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537:39–64.
- Katoh K., Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* 9:286–298.
- Keller C., Roos C., Groeneveld L.F., Fischer J., Zinner D. 2010. Introgressive hybridization in southern African baboons shapes patterns of mtDNA variation. *Am. J. Phys. Anthropol.* 142: 125–136.
- Kingdon J. 1988. What are face patterns and do they contribute to reproductive isolation in guenons? In: Gautier-Hion A., Bourliere F., Gautier J.-P., Kingdon J., editors. *A primate radiation: evolutionary biology of the African guenons*. New York: Cambridge University Press. p. 228–245.

- Kingdon J. 2005. The Kingdon pocket guide to African mammals. Princeton (NJ): Princeton University Press.
- Kingdon J.S. 1980. The role of visual signals and face patterns in African forest monkeys (guenons) of the genus *Cercopithecus*. *Trans. Zool. Soc. Lond.* 35:425–475.
- Kircher M. 2012. Analysis of high-throughput ancient DNA sequencing data. *Methods Mol. Biol.* 840:197–228.
- Kircher M., Stenzel U., Kelso J. 2009. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol.* 10:R83.
- Krause J., Briggs A.W., Kircher M., Maricic T., Zwyns N., Derevianko A., Pääbo S. 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr. Biol.* 20:231–236.
- Krause J., Unger T., Nocon A., Malaspina A.S., Kolokotronis S.O., Stiller M., Soibelzon L., Spriggs H., Dear P.H., Briggs A.W., Bray S.C., O'Brien S.J., Rabeder G., Matheus P., Cooper A., Slatkin M., Pääbo S., Hofreiter M. 2008. Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene–Pliocene boundary. *BMC Evol. Biol.* 8:220.
- Kuroda S., Kano T., Muhindo K. 1985. Further information on the new monkey species, *Cercopithecus salongo* THYS VAN DEN AUDENAERDE, 1977. *Primates* 26:325–333.
- Lanfear R., Calcott B., Ho S.Y., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29:1695–1701.
- Larsen P.A., Marchan-Rivadeneira M.R., Baker R.J. 2010. Natural hybridization generates mammalian lineage with species characteristics. *Proc. Natl Acad. Sci. U. S. A.* 107:11447–11452.
- Legros F., Malka F., Frachon P., Lombs A., Rojo M. 2004. Organization and dynamics of human mitochondrial DNA. *J. Cell Sci.* 117: 2653–2662.
- Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760.
- Maricic T., Whitten M., Pääbo S. 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 5:e14004.
- Mason V.C., Li G., Helgen K.M., Murphy W.J. 2011. Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Res.* 21:1695–1704.
- Meyer M., Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* 2010:prot5448.
- Meyer M., Kircher M., Gansauge M.T., Li H., Racimo F., Mallick S., Schraiber J.G., Jay F., Prufer K., de Filippo C., Sudmant P.H., Alkan C., Fu Q.M., Do R., Rohland N., Tandon A., Siebauer M., Green R.E., Bryc K., Briggs A.W., Stenzel U., Dabney J., Shendure J., Kitzman J., Hammer M.F., Shunkov M.V., Derevianko A.P., Patterson N., Andres A.M., Eichler E.E., Slatkin M., Reich D., Kelso J., Pääbo S. 2012. A high-coverage genome sequence from an Archaic Denisovan individual. *Science* 338:222–226.
- Miller M.A., Holder M.T., Vos R., Midford P.E., Liebowitz T., Chan L., Hoover P., Warnow T. 2009. The CIPRES portals, CIPRES. 2009–08–04. Available from: URL <http://www.webcitation.org/5imQJJeQa> last accessed April 02, 2013.
- Moulin S., Gerbault-Seureau M., Dutrillaux B., Richard F.A. 2008. Phylogenomics of African guenons. *Chromosome Res.* 16:783–799.
- Moyle R.G., Filardi C.E., Smith C.E., Diamond J. 2009. Explosive Pleistocene diversification and hemispheric expansion of a “great speciator”. *Proc. Natl Acad. Sci. U. S. A.* 106:1863–1868.
- Nagel U. 1973. A comparison of anubis baboons, hamadryas baboons and their hybrids at a species border in Ethiopia. *Folia Primatol.* 19:104–165.
- Olson D.M., Dinerstein E., Wikramanayake E.D., Burgess N.D., Powell G.V.N., Underwood E.C., D'Amico J.A., Itoua I., Strand H.E., Morrison J.C., Loucks C.J., Allnutt T.F., Ricketts T.H., Kura Y., Lamoreux J.F., Wettengel W.W., Hedao P., Kassem K.R. 2001. Terrestrial ecoregions of the worlds: a new map of life on Earth. *Bioscience* 51:933–938.
- Perelman P., Johnson W.E., Roos C., Seuanez H.N., Horvath J.E., Moreira M.A., Kessing B., Pontius J., Roelke M., Rumpler Y., Schneider M.P., Silva A., O'Brien S.J., Pecon-Slatery J. 2011. A molecular phylogeny of living primates. *PLoS Genet.* 7:e1001342.
- Raaum R.L., Sterner K.N., Noviello C.M., Stewart C.B., 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.
- Rabosky D.L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60:1152–1164.
- Rambaut A., Drummond A.J. 2009. Tracer v1.5. Available from: URL <http://tree.bio.ed.ac.uk/software/tracer/> last accessed April 02, 2013.
- Ree R.H., Moore B.R., Webb C.O., Donoghue M.J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.
- Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14.
- Robin E.D., Wong R. 1988. Mitochondrial-DNA molecules and virtual number of mitochondria per cell in mammalian-cells. *J. Cell. Physiol.* 136:507–513.
- Rohland N., Siedel H., Hofreiter M. 2009. A rapid column-based ancient DNA extraction method for increased sample throughput. *Mol. Ecol. Resour.* 10:677–683.
- Roos C., Zinner D., Kubatko L.S., Schwarz C., Yang M., Meyer D., Nash S.D., Xing J., Batzer M.A., Brameier M., Leendertz F.H., Ziegler T., Perwitasari-Farajallah D., Nadler T., Walter L., Osterholz M. 2011. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. *BMC Evol. Biol.* 11:77.
- Rumpler Y., Hauwy M., Fausser J.L., Roos C., Zaramody A., Andriaholinirina N., Zinner D. 2011. Comparing chromosomal and mitochondrial phylogenies of the Indriidae (Primates, Lemuriformes). *Chromosome Res.* 19:209–224.
- Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.
- Schliep K.P. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27:592–593.
- Schoof V.A.M., Jack K.M., Isbell L.A. 2009. What traits promote male parallel dispersal in primates? *Behaviour* 146:701–726.
- Shimodaira H., Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114–1116.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
- Stamatakis A., Hoover P., Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.* 57: 758–771.
- Stothard P. 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 28:1102, 1104.
- Stromer E. 1920. Mitteilungen über die Wirbeltierreste aus dem Mittelpliocän des Natrontales (Ägypten). 5: Nachtrag zur 1: Affen. *Sitzungsberichte der bayerische Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse.* 1:345–370.
- Struhsaker T.T., Butynski T.M., Lwanga J.S. 1988. Hybridization between redbellied (*Cercopithecus ascanius schmidti*) and blue (*C. mitis stuhlmanni*) monkeys in the Kibale Forest, Uganda. In: Gautier-Hion A., Bourliere F., Gautier J.-P., Kingdon J., editors. *A primate radiation: evolutionary biology of the African guenons*. New York: Cambridge University Press. p. 54–78.
- Thalmann O., Hebler J., Poinar H.N., Pääbo S., 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.
- Tosi A.J. 2008. Forest monkeys and *Pleistocene refugia*: a phylogeographic window onto the disjunct distribution of the *Chlorocebus lhoesti* species group. *Zool. J. Linn. Soc.* 154: 408–418.
- Tosi A.J., Detwiler K.M., Disotell T.R. 2005. X-chromosomal window into the evolutionary history of the guenons (Primates : Cercopithecini). *Mol. Phylogenet. Evol.* 36:58–66.
- Tosi A.J., Melnick D.J., Disotell T.R. 2004. Sex chromosome phylogenetics indicate a single transition to terrestriality in the guenons (tribe Cercopithecini). *J. Hum. Evol.* 46:223–237.



- Udvardy M.D.F. 1975. A classification of the biogeographical provinces of the world. Morges, Switzerland: IUCN.
- Wandeler P., Hoeck P.E., Keller L.F. 2007. Back to the future: museum specimens in population genetics. *Trends Ecol. Evol.* 22:634–642.
- Wertheim J.O., Worobey M. 2007. A challenge to the ancient origin of SIVagm based on African green monkey mitochondrial genomes. *PLoS Pathog.* 3:866–873.
- Wilson D.E., Reeder D.M. 2005. *Mammal species of the world: a taxonomic and geographic reference*. Baltimore (MD): The Johns Hopkins University Press.
- Xing J., Wang H., Zhang Y., Ray D.A., Tosi A.J., Disotell T.R., Batzer M.A. 2007. A mobile element-based evolutionary history of guenons (tribe Cercopithecini). *BMC Biol.* 5:5.
- Zinner D., Groeneveld L.F., Keller C., Roos C. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.)—indication for introgressive hybridization? *BMC Evol. Biol.* 9:83.
- Zuberbühler K. 2003. Effects of natural and sexual selection on the evolution of guenon loud calls. In: Glenn M.E., Cords M., editors. *The guenons: diversity and adaptation in African monkeys*. New York: Kluwer Academic/Plenum Publishers. p. 289–306.