**Inverted intergeneric introgression between critically endangered kipunjis and yellow baboons in two disjunct populations**

Dietmar Zinner, Idrissa S. Chuma, Sascha Knauf, Christian Roos

**Electronic supplementary material:**

**Ethical statement**

No animals were captured specifically of this study. Tissue samples of Udzungwa baboons were originally collected for a German Research Foundation (DFG) funded project (KN1097/3-1 and RO3055/2-1) on screening of *Treponema pallidum* infection in non-human primates in Tanzania. In this study, baboons were chemically immobilized and sampled in accordance with the requirements of the relevant guidelines and regulations. The study was in line with the Veterinary Act of 2003 and Tanzania Wildlife Research Institute’s (TAWIRI) Guideline for Conducting Wildlife Research (2001). Respective permits for wildlife-protected areas were issued by the Commission for Science and Technology in Tanzania (2015-89-NA-2014-228), Ministry for Natural Resources and Tourism (Wildlife Division, HA.403/563/01B/90, 178/606/01/115 and HA.178/606/01/6), Tanzania National Parks (TNP/HQ/C.10/13), and Ngorongoro Conservation Area Authority (NCAA/D/240/Vol.XXV/130) as well as the Revolutionary Government of Zanzibar through the second Vice-President’s Office (Zanzibar Research Committee OMPR/M.95/C.6/2/Vol.IV/60). The study methodology including the animal handling protocols were reviewed and approved by the Vice Chancellor of Sokoine University of Agriculture (SUA/ADM/R.1/8). ‘Good Veterinary Practice’ rules were applied to all procedures where animals were handled. Registered veterinarians did the immobilization of baboons and closely monitored anesthetized animals until they fully recovered. Additional DNA sequences for our study were downloaded from GenBank.

**Methods**

*Sample Collection*

Sampling procedures followed a standardized protocol previously used for baboon immobilization (Knauf et al. 2015). Briefly, chemical immobilization was achieved by remote distance injection of 10.0 mg ketamine/kg body mass (bm) (Kyron Prescriptions) in combination with 0.05 mg/kg bm medetomidine (Domitor, Pfizer). Anesthetics were intramuscularly injected using a cold-gas immobilization rifle (MOD JM, Dan-Inject ApS) and appropriate projectiles. Recommended distance-adjusted bar levels were applied to minimize tissue damage. Immobilized baboons were continuously observed for vital parameters such as breath, pulse frequency, and internal body temperature. Pulse frequency and blood oxygen saturation were monitored utilizing a Nellcor OxiMax N65 Pulse Oximeter. A 6-mm biopsy from the skin was taken using a sterile dermal biopsy punch and preserved in lysis buffer (10 mM Tris [pH 8.0], 0.1 EDTA, 0.5 % SDS). Samples were incubated at environmental temperature for 12-24 hours and subsequently frozen and stored in liquid nitrogen.

Biopsy wounds were treated with Silverspray (Silver Alluminium Aerosol, Henry Schein) and animals were allowed to recover under close supervision until they were able to return to their group on their own volition. Samples were processed and aliquoted at the end of each day and temporarily stored at -80°C at the TAWIRI headquarter in Arusha, Tanzania. Aliquots were exported to the German Primate Center in Göttingen, Germany for further analysis.

*Laboratory methods*

DNA extraction was performed following the standard protocol of the QiAmp DNA Mini Kit (Qiagen), with some minor modifications. All working steps related to the handling of tissue and DNA extraction were carried out under a lateral-flow workbench. Briefly, approximately 25 mg tissue was cut into small pieces and incubated in 180 µl lysis buffer in which the sample was stored since collection. After adding 20 µl proteinase K, samples were digested over night at 56°C and 900 rpm (Thermomixer comfort, Eppendorf). We further added an additional washing step with 300 µl AW1 buffer and eluted the DNA twice with 100 µl AE buffer.Extracted DNA was further cleaned using glycogen precipitation according to the protocol published in Knauf et al. (2016). DNA concentration was quantified using the NanoDrop spectrophotometer.

We amplified and sequenced the Brown region (Brown et al. 1982) of the mtDNA genome of eleven northern Udzungwa and six southern Udzungwa yellow baboons using methods described in Newman et al. (2004) and Zinner et al. (2009b). In baboons, the region has a length of 896 bp and contains 457 bp of the 3' end of the NADH dehydrogenase subunit IV (ND4) gene, the tRNA genes for histidine (His), serine (Ser), and leucine (Leu), and 239 of the 5' end of the NADH dehydrogenase subunit V (ND5) gene (Wildman et al. 2004). PCR reactions were carried out in a total volume of 30 µl containing 1x reaction buffer, 0.16 mM for each dNTP, 0.33 μM for each primer (5'-CTAGTAATTGTAGCCTCCCTC-3', 5'-TAGACCAGGTAATGAATAGTGC-3', Newman et al. 2004), 0.6 mg/ml BSA, 1 U BiothermTaq 5000 (Genecraft, Germany) and 10-20 ng genomic DNA. Amplification was performed in a Sensoquest Labcycler (Sensoquest) under following conditions: pre-denaturation at 95°C for 2 min, 40 cycles each with denaturation at 95°C for 1 min, annealing at 58°C for 1 min and extension at 72°C for 1 min, followed by a final extension step at 72°C for 5 min. PCR performance was checked on 1% agarose gels and PCR products with correct size were excised from the gel. After purification with the Qiagen Gel Extraction Kit, PCR products were sequenced on an ABI 3130 xL sequencer using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). To avoid and monitor cross-sample contamination, we conducted all working steps (DNA extraction, PCR setup, gel electrophoresis, PCR purification, etc.) in separate laboratories. Further, we regularly changed gloves, used dual-filter tips, and ran various negative controls. We checked sequence electropherograms with 4Peaks 1.8 (www.nucleobytes.com) and edited sequences in SeaView 4.5.4 (Gouy et al. 2010). SeaView was also used to check for premature stop codons in protein-coding gene fragments. Among the eleven northern Udzungwa samples we found two haplotypes (UN02 and UN03) and among the six southern Udzungwa samples we found just one haplotype (UN18). All newly generated haplotypes were deposited at GenBank (for accession numbers see suppl. Table 1).

We minimized the risk to amplify nuclear integrations of mitochondrial fragments (*numts*) by using baboon-specific primers that were used in previous baboon studies (Newman et al. 2004; Zinner et al. 2009a, 2009b). We observed no multiple peaks in sequence electropherograms as indication for amplification of multiple gene copies and both protein-coding fragments were correctly transcribed. Furthermore, we sequenced the complete mtDNA genome of a southern Udzungwa yellow baboon (UN02, Roos et al. unpublished) by methods described in Zinner et al. (2013). The sequence of the Brown region in the mtDNA genome was identical with that generated by the Brown region-specific primers.

*Phylogenetic analysis*

Newly generated haplotypes were aligned against orthologous sequences derived from GenBank of other baboon species, kipunjis, and related taxa that were used as outgroups (suppl. Table 1). The alignment subjected to phylogenetic tree reconstruction contained 50 unique sequences. We reconstructed phylogenetic trees using maximum-likelihood (ML) and Bayesian inference. For both methods, we applied the optimal substitution models for each partition (NADH4, tRNAs, NADH5) as determined with ModelFinder (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017) in IQ-TREE 1.5.2 (Nguyen et al. 2015) under the Bayesian Information Criterion (BIC). The ML tree was generated with IQ-Tree using 10,000 ultrafast bootstrap replications (Minh et al. 2013). Bayesian trees were reconstructed in MrBayes 3.2.6 (Ronquist et al. 2012) with four independent Markov Chain Monte Carlo (MCMC) runs. All repetitions were run for 10 million generations with tree and parameter sampling occurring every 100 generations and burn-in of 10%. To check convergence of all parameters and the adequacy of the burn-in, we assessed the uncorrected potential scale reduction factor (PSRF) (Gelman and Rubin 1992) as calculated by MrBayes. We calculated posterior probabilities and a phylogram with mean branch lengths from the posterior density of trees using MrBayes.

We estimated divergence times with BEAST 2.4.7 (Bouckaert et al. 2014) applying a Bayesian MCMC method with a relaxed molecular clock approach (Drummond et al. 2006). Therefore, we assumed a relaxed lognormal model of lineage variation and a Birth-Death Process prior for branching rates. For each of the three partitions, the best-fit substitution model was applied. To calibrate the molecular clock, we used the split between *Theropithecus* and *Papio* based on fossil evidence (Jablonski and Frost 2010). We constrained this node using a lognormal distribution with mean = 0.0, standard deviation = 0.5 and offset = 4.0 resulting in a 95% HPD of 4.38-6.66 Ma. Four replicates were run in BEAST for 25 million generations with tree and parameter sampling occurring every 100 generations. TRACER 1.6.1 (http://beast.bio.ed.ac.uk/Tracer) was used to assess the adequacy of a 10% burn-in and the convergence of all parameters via visual inspection of the trace of the parameter across generations. Sampling distributions were combined (10% burn-in) using the software LogCombiner 2.4.7. A consensus chronogram with node height distribution was generated with TreeAnnotator 2.4.7. All trees were visualized with FigTree 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Suppl. Table 1 List of samples. Accession numbers, taxa, clades and geographic provenances. Samples without accession numbers are identical with the respective previous sample with a number.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Acc. Num. | Taxon | Clade | Location | Country | Latitude | Longitude |
| JQ068264 | *M.leucophaeus* |  |  |  |  |  |
| KC757401 | *L.aterrimus* |  |  |  |  |  |
| KJ434960 | *L.aterrimus* |  |  |  |  |  |
| KM267477 | *T.gelada* |  |  |  |  |  |
| KC757412 | *T.gelada* |  |  |  |  |  |
| EU885833 | *P.ursinus* | pus | De Hoop NR | South Africa | -34.45621 | 20.40658 |
| EU885827 | *P.ursinus* | pus | Hakos Guest Farm | Namibia | -23.23708 | 16.36463 |
| MG569925 | *P.anubis* | woe | Mbam et Djerem NP | Cameroon | 12.84675 | 6.09046 |
| EU885769 | *P.anubis* | woe | Lum(m)a | Nigeria | 10.31810 | 4.26420 |
| JX946198 | *P.anubis* | wow | Komoe NP | Ivory Coast | 8.80000 | -3.79000 |
| EU885768 | *P.anubis* | wow | Bwari (Barari) | Nigeria | 8.81140 | 7.22030 |
| EU885885 | *P.papio* | pp | Haute Niger NP | Guinea | 10.50143 | -10.33702 |
| EU885807 | *P.papio* | pp | Niokolo Koba NP | Senegal | 13.07467 | -12.72090 |
| EU885798 | *P.kindae* | pk | Kasanka NP | Zambia | -12.59059 | 30.25202 |
| MG569932 | *P.kindae* | pk | North Kafue (Mokambi Lodge) | Zambia | -14.98034 | 25.99480 |
| EU885851 | *P.anubis* | ph | Haykota, R. Gash | Eritrea | 15.15695 | 37.06600 |
| EU885800 | *P.hamadryas* | ph | Abdur, R. Dekano | Eritrea | 15.12857 | 39.84585 |
| EU885783 | *P.anubis* | ph | Managasha NP | Ethiopia | 8.96838 | 38.57125 |
| KM267422 | *P.hamadryas* | ph | Awash Station | Ethiopia | 8.99268 | 40.17775 |
| EU885817 | *P.ursinus* | pun | Pilanesberg GR | South Africa | -25.11111 | 26.87805 |
| EU885814 | *P.ursinus* | pun | Gorongosa NP | Mozambique | -18.97833 | 34.36111 |
| MG569933 | *P.cynocephalus* | pcl | Senga Hills FR | Malawi | -13.70677 | 34.61853 |
| MG569934 | *P.cynocephalus* | pcl | Senga Hills FR | Malawi | -13.70967 | 34.62217 |
| EU885794 | *P.cynocephalus* | pcl | South Luangwa NP | Zambia | -13.26840 | 31.63793 |
| EU885795 | *P.cynocephalus* | pcl | South Luangwa NP | Zambia | -13.26840 | 31.63793 |
| EU885796 | *P.cynocephalus* | pcl | Luambe NP | Zambia | -12.45780 | 32.14550 |
| MG569926 | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.52151 | 36.60478 |
|  | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.52020 | 36.59877 |
|  | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.49790 | 36.56256 |
|  | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.52070 | 36.62044 |
|  | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.51679 | 36.62833 |
|  | *P.cynocephalus* | pcn | Udzungwa NP north | Tanzania | -7.49594 | 36.56725 |
| EU885787 | *P.cynocephalus* | pcn | Webi Shebelli | Somalia | 2.42083 | 45.43333 |
| EU885790 | *P.cynocephalus* | pcn | Amboseli | Kenya | -2.29000 | 37.39000 |
| MG569927 | *P.cynocephalus* | pcn | 30km W of Dodoma | Tanzania | -6.05992 | 35.57563 |
| MG569928 | *P.cynocephalus* | pcn | Dodoma->Iringa | Tanzania | -7.13016 | 35.98829 |
| MG569929 | *P.cynocephalus* | pcn | Mikumi NP | Tanzania | -7.26596 | 37.18905 |
| MG569930 | *P.cynocephalus* | pcn | N of Morogoro | Tanzania | -6.62585 | 37.57897 |
| MG569931 | *P.cynocephalus* | pcn | Mikumi NP | Tanzania | -7.23480 | 37.20976 |
| GU068081 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| GU068082 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| GU068083 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| GU068084 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| GU068085 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| GU068086 | *R.kipunji* | Rk | Ndundulu FR | Tanzania | -7.79730 | 36.51195 |
| MG569923 | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84488 | 36.88652 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84385 | 36.88650 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.85745 | 36.89092 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.85807 | 36.89122 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84805 | 36.88613 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84347 | 36.88682 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.81512 | 36.89523 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.80575 | 36.89683 |
| MG569924 | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84472 | 36.88388 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.85593 | 36.89015 |
|  | *P.cynocephalus* | Rk | Udzungwa NP south | Tanzania | -7.84988 | 36.88511 |
| GU068078 | *R.kipunji* | pcs | Mt Rungwe | Tanzania | -9.16412 | 33.63200 |
| GU068079 | *R.kipunji* | pcs | Livingstone Forest | Tanzania | -9.20483 | 33.89046 |
| GU068080 | *R.kipunji* | pcs | Mt Rungwe, Syukula | Tanzania | -9.16651 | 33.63304 |
| MG569935 | *P.cynocephalus* | pcs | Nanguruwe | Tanzania | -10.49506 | 39.99297 |
| MG569936 | *P.cynocephalus* | pcs | Mingoyo | Tanzania | -10.10855 | 39.65751 |
| MG569937 | *P.cynocephalus* | pcs | north of Kitaya | Tanzania | -10.59522 | 40.13528 |
| MG569938 | *P.cynocephalus* | pcs | Chifunde (North Luangwa) | Zambia | -11.86058 | 32.43357 |
| JX946200 | *P.cynocephalus* | pcs | Amani (S of Tunduru) | Tanzania | -11.26054 | 37.51363 |
| MG569939 | *P.cynocephalus* | pcs | Amani (S of Tunduru) | Tanzania | -11.26054 | 37.51363 |
| MG569940 | *P.cynocephalus* | pcs | W of Masasi | Tanzania | -10.86915 | 38.60093 |
| MG569941 | *P.cynocephalus* | pcs | Chiwata | Tanzania | -10.59470 | 38.98838 |
| MG569942 | *P.cynocephalus* | pcs | E of Sinza | Tanzania | -8.63751 | 39.30755 |
| MG569943 | *P.cynocephalus* | pcs | 100km N of Songea | Tanzania | -9.90943 | 35.50796 |
| EU885792 | *P.cynocephalus* | pcs | Michiru Mountains CA | Malawi | -15.72307 | 34.98668 |
| MG569944 | *P.cynocephalus* | pcs | Kamani | Tanzania | -8.90926 | 34.19086 |

clades: pcn = northern yellow + eastern olive baboon; pcs = southern yellow baboon; pcl = Luangwa Valley yellow baboon; pun = northern chacma baboon; pus = southern chacma baboon; pp = Guinea baboon; pk = Kinda baboon; ph = hamadryas + north-eastern olive baboon; wow = western olive baboon west; woe = western olive baboon east; Rk = kipunji.

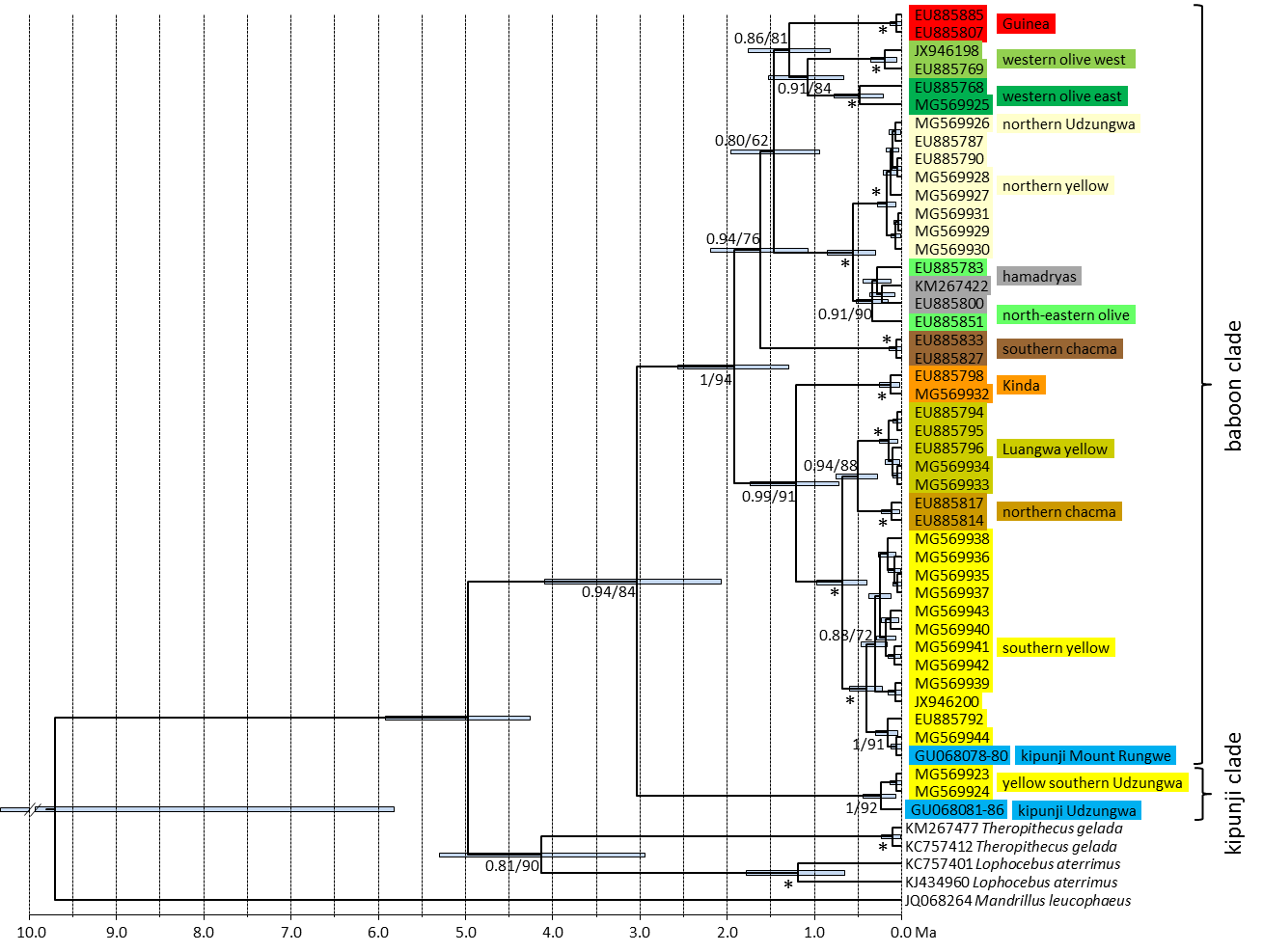
CA = Conservation Area; FR = Forest Reserve; GR= Game Reserve; NP = National Park; NR = Nature Reserve

Original long version IDs for analysed baboon samples from Udzungwa NP: UN02 = 02UNM1220616; UN03 = 03UNF1220616; UN18 = 18UNM1100317

Suppl. Table 2. Estimated divergence ages of mtDNA lineages in million years ago (Ma)

|  |  |
| --- | --- |
| Split | divergence ages (95% HPD) |
| *Mandrillus leucophaeus* vs. other Papionini | 9.70 (5.82-14.12) |
| *Theropithecus*/*Lophocebus* vs. *Rungwecebus*/*Papio* | 4.97 (4.25-5.92) |
| *Theropithecus* vs. *Lophocebus* | 4.13 (2.94-5.30) |
| MRCA\* *Theropithecus* | 0.10 (0.01-0.23) |
| MRCA *Lophocebus* | 1.19 (0.65-1.79) |
| Southern Udzungwa (*R. kipunji* + *P. cynocephalus*) vs. remaining *Papio* + Mount Rungwe *R. kipunji* | 3.03 (2.06-4.09) |
| MRCA *Papio* (w/o Southern Udzungwa *P. cynocephalus*) | 1.92 (1.30-2.57) |
| Southern Udzungwa: *R. kipunji* vs. *P. cynocephalus* | 0.24 (0.07-0.44) |
| *P. cynocephalus* south vs. Mount Rungwe *R. kipunji* | 0.05 (0.00-0.13) |

\*MRCA: most recent common ancestor



Suppl. Figure 1. Chronogram showing phylogenetic relationships and divergence times among various baboon and kipunji mtDNA lineages. The mtDNA lineage of the Mount Rungwe kipunji clusters with southern yellow baboons, while yellow baboons from southern Udzungwa cluster with the Udzungwa kipunji. In contrast, yellow baboons from northern Udzungwa cluster with their conspecifics of the northern yellow baboon mtDNA clade. Blue-grey bars indicate 95% HPDs of estimated divergence times (see also suppl. Table 2). Numbers at nodes refer to ML BS and Bayesian PP values; \* denotes ML BS ≥ 95% and Bayesian PP ≥ 0.99.

*References*

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C, Xie D, Suchard M, Rambaut A, Drummond A. 2014 BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comp. Biol.* **10**, e1003537. (doi: 10.1371/journal.pcbi.1003537)

Brown WM, Prager EM, Wang A, Wilson AC. 1982 Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* **18**, 225-239. (doi: 10.1007/BF01734101)

Chernomor O, Klaere S, von Haeseler A, Minh BQ. 2016 Split diversity: measuring and optimizing biodiversity using phylogenetic split networks. In *Biodiversity conservation and phylogenetic systematics* (eds R Pellens, P Grandcolas), pp. 173-195. Cham, Switzerland: Springer.

Drummond AJ, Ho SYN, Phillips MJ, Rambaut A. 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88. (doi: 10.1371/journal.pbio.0040088)

Gelman A, Rubin DB. 1992 Inference from iterative simulation using multiple sequences. Statistical *Science* **7**, 457-472. (doi: 10.1214/ss/1177011136)

Gouy M, Guindon S, Gascuel O. 2010 SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221-224. (doi: 10.1093/molbev/msp259)

Jablonski NG, Frost SR. 2010 Cercopithecoidea. In *Cenozoic mammals of Africa* (eds L Werdelin, WJ Sanders WJ), pp. 393-428. Berkeley, CA: University of California Press.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587-589. (doi: 10.1038/nmeth.4285)

Knauf S, Barnett U, Maciej P, Klapproth M, Ndao I, Frischmann S, Fischer J, Zinner D, Liu H. 2015 High prevalence of antibodies against the bacterium *Treponema pallidum* in Senegalese Guinea baboons (*Papio papio*). *PLoS One* **10**, e0143100. (doi: 10.1371/journal.pone.0143100).

Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, Keyyu JD, Fyumagwa R, Lüert S, Godornes C, Liu H, Schwarz C, Šmajs D, Grange P, Zinner D, Roos C, Lukehart SA. 2016 Isolation of *Treponema* DNA from necrophagous flies in a natural ecosystem. *EBioMedicine* **11**, 85-90. (doi: 10.1016/j.ebiom.2016.07.033).

Minh BQ, Nguyen MAT, von Haeseler A. 2013 Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188-1195. (doi: 10.1093/molbev/mst024)

Newman TK, Jolly CJ, Rogers J. 2004 Mitochondrial phylogeny and systematics of baboons (*Papio*). *Am. J. Phys. Anthropol.* **124**, 17-27. (doi: 10.1002/ajpa.10340)

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268-274. (doi: 10.1093/molbev/msu300)

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012 MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539-542. (doi: 10.1093/sysbio/sys029)

Wildman DE, Bergman TJ, al-Aghbari A, Sterner KN, Newman TK, Phillips-Conroy JE, Jolly CJ, Disotell TR. 2004 Mitochondrial evidence for the origin of hamadryas baboons. *Mol. Phylogenet. Evol.* **32**, 287-296. (doi: 10.1016/j.ympev.2003.12.014)