**Supplementary Material for:**

**Genetics and community-based restoration can guide conservation of forest fragments for endangered primates**

María José Ruiz-Lópeza,l, Arleigh Jane Hitchcockb, Noah D. Simonsc, Jenneca McCarterb, Colin A. Chapmane,f,g,h, Dipto Sarkari, Patrick Omejaj, Tony L Goldbergk, and Nelson Tingb

aEstación Biológica de Doñana, CSIC, Dept. Wetland Ecology, Américo Vespucio s/n 41092 - Sevilla Spain,

bDepartment of Anthropology and Institute of Ecology and Evolution University of Oregon, Eugene, Oregon, USA

cDivision of Neuroscience and Division of Genetics, Oregon Health and Science University, Portland, Oregon, USA

eWilson Center,1300 Pennsylvania Avenue NW, Washington, D.C., USA

fCenter for the Advanced Study of Human Paleobiology, George Washington University, Washington DC, USA

gSchool of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa

hShaanxi Key Laboratory for Animal Conservation, Northwest University, Xi’an, China

iDepartment of Geography and Environmental Studies, Carleton University, Ottawa Canada,

jMakerere University Biological Field Station, P.O. Box 967, Fort Portal.

kDepartment of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin, Madison, Wisconsin

l CIBER Epidemiología y Salud Pública, Spain

Corresponding Author email: mjruiz@ebd.csic.es, nting@uoregon.edu

Orcid: María José Ruiz-López - 0000-0002-6849-644X, Arleigh Jane Hitchcock - Noah D. Simons - 0000-0002-9703-5802, Jenneca McCarter - , Colin A. Chapman - 0000-0002-8827-8140, Dipto Sarkar - 0000-0003-2254-049X, Tony L Goldberg - 0000-0003-3962-4913, and Nelson Ting - 0000-0002-0765-8177

**Details on sampling area and land use.**

Kibale National Park is dominated by mid-altitude (920 - 1590 m), moist-evergreen forest that receives an annual rainfall of 1660 mm (1970 – 2019). The process of forest fragmentation and habitat loss outside of Kibale was well advanced by the 1940s, and aerial photographs of the region taken in 1959 reveal only small fragments embedded in a matrix of agricultural fields. Improvement of the roads to the region in the early1990s and a rapidly growing human population density led to rapid forest clearing (Chapman et al., 2013). Between 2000 and 2020 the human population density within 1 km of the park’s boundary almost doubled, from 123 to 229 people / km2 (MacKenzie et al., 2017; WorldPop, 2020). Almost all (95%) of the local people (predominately Batoro and Bakiga tribes) are smallhold farmers, cultivating less than 5 ha (Mackenzie and Hartter, 2013). Farmers plant more than 20 species of subsistence crops, and the main staple foods are cooking bananas, sweet potatoes, Irish potatoes, beans, groundnuts, maize, and cassava. Cash crops other than tea and food crops (e.g., potatoes, maize, cooking bananas) are limited in the area, but since 2005 limited cultivation of tobacco and *Artemisia annua* has begun. Wood is commonly used for cooking, as well as for the production of charcoal, alcohol, bricks, and lumber (Naughton-Treves and Chapman, 2002; Naughton-Treves et al., 2007), such that woodlots have become a profitable source of income for many families.

***Sampling***

Between 2010 and 2013 we collected 299 red colobus samples from eight sites in or near Kibale National Park. We sampled three sites within Kibale (Kanyawara, Sebitoli, and Mainaro) and five fragments (Isunga, Kamakune, Byara, Lake Nkuruba, and Lake Kasenda). Of the collected samples, 54 were blood samples from Kanyawara and 245 were fecal samples across the other localities. Fecal samples were collected along transects except for Kanyawara where samples were collected during behavioral surveys from habituated, known individuals. Blood samples were collected in WhatmanTM FTATM cards (Whatman, Sigma-Aldrich, St Louis, Missouri, USA) as part of a broader study of health and conservation (Goldberg et al. 2012). Total genomic DNA was extracted from FTATM cards and fecal samples using a QIAmp® DNA micro Kit and a QIAmp®Stool Kits (Qiagen, Valencia, California, USA) respectively.

**Determination of the amount of host DNA in fecal samples using a qPCR**

To assess the quality of the fecal samples and estimate the number of repetitions necessary for genotyping, we quantified the amount of endogenous DNA using qPCR. Fecal samples were preserved in RNAlater and kept frozen at -20 C until DNA extraction. Extractions made from fecal samples can contain host, plant and bacterial DNA so a quantitative PCR (qPCR) was run for each sample to determine how much host DNA was present in each sample. An 81-bp section of the c-myc proto-oncogene was targeted from mouse (Genebank Accession no. X01023) and human (Genebank Accession no. J00120) using the 5’ nuclease assay and the protocol and Taqman probe following Morin et al. (2001). Probes were created using a 5’ 6-FAM reporter dye and a 3’-TAMRA quencher dye (Applied Biosystems, Foster City, California, USA). The oligonucleotide sequences used were: forward primer (CMYC\_E3\_F1U1) GCCAGAGGAGGAACGAGCT, reverse primer (CMYC\_E3\_R1U1) GGGCCTTTTCATTGTTTTCCA and probe (CMYC\_E3\_TMV) 6-FAM-TGCCTGCGTGACCAGATCC-TAMRA (Morin et al., 2001). The qPCR was run in a StepOnePlus Real-time PCR System (Applied Biosystems, Foster City, California, USA) in 20 µL containing 1x ‘TaqMan Universal PCR Master Mix II with UNG’ (Applied Biosystems, Foster City, California, USA), 200 nM probe, 8µg BSA (Biolabs), 300 nM of each primer and 2µL DNA extract. The cycle conditions were: initial incubation at 50˚C for 2 minutes and 95˚C for 10 minutes followed by 40 cycles of at 95˚ C for 15 seconds then 60˚ C for 1 minute. Each qPCR contained triplicate sets of standards of known quantities of DNA derived from a human DNA sample (Life Tech). The DNA quantities for each standard in 2 µl of DNA were: 20 µg, 5 µg, 1.25 µg, 0.312 µg, 0.078 µg, 0.019 µg, and 0.008 µg. Three negative controls were included in each qPCR and samples were run in duplicate. Stepone® Software v2.2.2 was used for analysis and each quantity checked on a standard curve. The standards were used to generate a standard curve, to which samples of unknown DNA quantities were compared to estimate DNA concentrations. To ensure accuracy of the genotypes and minimize the effects of allelic dropout. we only genotyped fecal samples that had an estimated concentration of host DNA higher than 0.05 ng/µl. All fecal samples containing between 0.05 ng/µl - 0.50 ng/µl were genotyped six times, 0.51 ng/µl - 0.100 ng/µl were genotypes 4 times and > 0.100 ng ng/µl 3 times (Morin et al., 2011, Wikberg et al., 2012).

**Multiplex microsatellite PCR**

All samples were genotyped at 15 microsatellite loci selected from the human genome and that have been used in previous studies of colobus monkeys (Table 1. Supplementary material; Allen et al, 2012; Wikberg et al., 2012). The 15 microsatellite loci were labeled with fluorescent dyes (NED, 6-fam or HEX) and co-amplified in seven multiplex Polymerase Chain Reactions (PCRs) following the Multiplex PCR Kit (Qiagen, Valencia, California) protocol for 35 cycles and the appropriate annealing temperature (see supplementary material). Each PCR was performed in 12.5 µl reactions with 2µl of DNA (containing at least 0.01ng of DNA), 1x QIAGEN Multiplex PCR Master Mix (contains 3mM $Mg^{2+}$), 0.2µM of each primer, 0.8mg/ml of BSA and RNase-free water. Fragment length analyses were performed on an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) at the Center for Genome Research and Biocomputing (Oregon State University).

**Genotyping and individual identification.**

Alleles were scored using GeneMapper® Software 5 (Applied Biosystems, Foster City, CA, USA). Heterozyogous genotypes were confirmed in at least two separate reactions and homozygous genotypes were confirmed in at least three separate reactions. Blood samples were genotyped twice, and fecal samples were genotyped from 3 to 6 times (see above). We assigned a consensus genotype per sample across runs and then calculated the error rate per locus, including both false allele and allelic drop out, using the program GIMLET (Valière, 2002). The mean allelic dropout (ADO) across all markers was 0.0424 with a range of 0.0115-0.085 and the mean false allele (FA) rate was 0.0406 with a range of 0.0085-0.079 (see supplementary material). To identify unique individuals, consensus genotypes were compared in the program GenAlEx 6.5 (Peakall and Smouse, 2006). We also tested the power of these markers to distinguish individuals by using a subset of 15 individuals and computing the Probability of identity (P(ID) random), the power to differentiate between randomly chosen individuals, as well as the probability of sibling identity (P(ID)sib) which is the power to differentiate between siblings. Because red colobus monkeys live in social groups composed of related individuals we used the more conservative P(ID)sib to estimate how many markers are needed to accurately identify individuals in this population (Waits et al., 2001). The probability of identity analyses indicated that at least 11 loci are needed to accurately identify individuals in all sites (P(ID) <1.0\*10-6 and P(ID)sib is <0.002 in all sites. When all sites were analyzed together, no locus deviated significantly from expected genotype frequencies under HWE after Bonferroni corrections or showed evidence of linkage disequilibrium. When analyses were carried out by site, one marker (locus D14s306) deviated significantly from HWE after Bonferroni corrections for the Isunga fragment (p-value = 0.0014). However, this marker did not show a significant heterozygote deficit or excess. None of the sites nor the overall population deviated significantly from HWE and no loci showed evidence of null alleles.

**Analyses in R of isolation by distance and genetic relatedness differences.**

We used the package adegenet 2.1.0 (Jombart and Ahmed, 2011) in R (R-Core-Team, 2020 to test for isolation by distance and analyze if mean relatedness and proportion of first-degree relationships decreased significantly with distance. supplementary for details). To do so we ran Mantel tests between the matrix of Euclidean geographic distances between sites and the matrix of genetic distances (either the Fst between sites or the pairwise genetic relatedness between sites). We used the function *mantel.randtest* and run 999 replicates.

**References**

Allen, J.M., Miyamoto, M.M., Akwarius, J., Margarini, K., Unvari-Martin, J., Carter, T., Chapman, C.A., 2012. Population genetics of habitat sensitive red colobus suggest long-term stability of Kibale National Park. Ecology and evolution 2, 2829-2842.

Chapman C.A., Ghai R.R., Jacob A.L. *et al.* (2013b). Going, going, gone: A 15-year history of the decline of primates in forest fragments near Kibale National Park, Uganda. in L.K. Marsh, C.A. Chapman editors. *Primates in fragments: complexity and resilience.* Springer, New York.

Goldberg T. L., Paige S.B., Chapman C.A. (2012). The Kibale EcoHealth Project: exploring connections among human health, animal health, and landscape dynamics in western Uganda. in A. A. Aguirre, P. Daszak, R. S. Ostfeld editors.New Directions in Conservation Medicine: Applied Cases of Ecological Health. New York, Oxford University Press**:** 452-465.

Jombart T., Ahmed I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* *27*, 3070-3071. doi: https://doi.org/ [10.1093/bioinformatics/btr521](https://doi.org/10.1093/bioinformatics/btr521)

Mackenzie C.A., Hartter J. (2013). Demand and proximity: drivers of illegal forest resource extraction. *Oryx* *47*, 288-297. doi: https://doi.org/10.1017/S0030605312000026

MacKenzie C.A., Salerno J., Chapman C.A., Reyna-Hurtado R., Tumusiime D.M., Drake M. (2017). Changing perceptions of protected area benefits and troubles around Kibale National Park, Uganda. *J Environ Manage* *200*, 217-228. doi: 10.1016/j.jenvman.2017.05.078

Morin P.A., Chambers K.E., Boesch C., Vigilant L. (2001). Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (Pan troglodytes verus). *Mol Ecol* *10*, 1835-1844.

Naughton-Treves L., Chapman C.A. (2002). Fuelwood resources and forest regeneration on fallow land in Uganda. *Journal of Sustainable Forestry* *14*, 19-32.

Naughton-Treves L., Kammen D.M., Chapman C. (2007). Burning biodiversity: woody biomass use by commercial and subsistence groups in western Uganda’s forests. *Biol Conserv* *134*, 232-241.

Peakall R., Smouse P.E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* *6*, 288-295.

R-Core-Team. (2020). R: A language and environment for statistical computing.R Foundation for Statistical Computing, Vienna, Austria.

Valière N (2002). Gimlet: a computer program for analysing genetic individual identification data. *Mol Ecol Notes* 2, 377-379

Waits L., Taberlet P, Luikart G. (2001). Estimating the probability of identity among genotypesin natural populations: cautions and guidelines. *Mol Ecol* 10, 249–256.

Wikberg E.C., Sicotte P., Campos F.A., Ting N. (2012). Between-group variation in female dispersal, kin composition of groups, and proximity patterns in a black-and-white colobus monkey (*Colobus vellerosus*). *PLOS one* *7*, e48740.

WorldPop. (2020). Open spatial demographic data and research.

**Table 1. Marker list information**.



**Table 2.** Pairwise mean r values on the lower half and proportion of first-degree relationships (Parent-offspring or Full-siblings) in the upper half.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Kanyawara  | Sebitoli  | Mainaro  | Isunga  | Kamakune  | Byara  | Nkuruba  | Kasenda  |
| Kanyawara  | - | 0.027 | 0.019 | 0.017 | 0.023 | 0.049 | 0.015 | 0.003 |
| Sebitoli  | 0.072 | - | 0.011 | 0.010 | 0.018 | 0.041 | 0.043 | 0.007 |
| Mainaro  | 0.071 | 0.056 | - | 0.014 | 0.018 | 0.053 | 0.046 | 0.007 |
| Isunga  | 0.063 | 0.067 | 0.082 | - | 0.020 | 0.030 | 0.006 | 0.011 |
| Kamakune  | 0.070 | 0.066 | 0.062 | 0.055 | - | 0.049 | 0.006 | 0.014 |
| Byara  | 0.088 | 0.077 | 0.101 | 0.072 | 0.081 | - | 0.035 | 0.010 |
| Nkuruba  | 0.070 | 0.078 | 0.052 | 0.062 | 0.061 | 0.080 | - | 0.016 |
| Kasenda  | 0.058 | 0.064 | 0.070 | 0.060 | 0.053 | 0.074 | 0.073 | - |