

RESEARCH ARTICLE

# New Insights into Samango Monkey Speciation in South Africa

Desiré L. Dalton<sup>1,2\*</sup>, Birthe Linden<sup>3,4</sup>, Kirsten Wimberger<sup>5</sup>, Lisa Jane Nupen<sup>1,6</sup>, Adrian S. W. Tordiffe<sup>1,7</sup>, Peter John Taylor<sup>4</sup>, M. Thabang Madisha<sup>1</sup>, Antoinette Kotze<sup>1,2</sup>

**1** National Zoological Gardens of South Africa, Pretoria, South Africa, **2** Genetics Department, University of the Free State, Bloemfontein, South Africa, **3** Department of Zoology, School of Mathematical & Natural Sciences, University of Venda, Thohoyandou, South Africa, **4** SARChI Chair on Biodiversity Value & Change in the Vhembe Biosphere Reserve & Core Member of Centre for Invasion Biology, School of Mathematical & Natural Sciences, University of Venda, Thohoyandou, South Africa, **5** Department of Biological Sciences, University of Cape Town, Rondebosch, South Africa, **6** Percy FitzPatrick Institute, Department of Biological Sciences, University of Cape Town, Rondebosch, South Africa, **7** Department of Companion Animal Clinical Studies, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa

\* [desire@nzg.ac.za](mailto:desire@nzg.ac.za)



OPEN ACCESS

**Citation:** Dalton DL, Linden B, Wimberger K, Nupen LJ, Tordiffe ASW, Taylor PJ, et al. (2015) New Insights into Samango Monkey Speciation in South Africa. PLoS ONE 10(3): e0117003. doi:10.1371/journal.pone.0117003

**Academic Editor:** William M. Switzer, Centers for Disease Control and Prevention, UNITED STATES

**Received:** July 21, 2014

**Accepted:** December 17, 2014

**Published:** March 23, 2015

**Copyright:** © 2015 Dalton et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** PJT acknowledges funding from the National Research Foundation (NRF) and Department of Science Technology (DST) through the SARChI Chair on Biodiversity Value Change in the Vhembe Biosphere. DLD, AK and ASWT acknowledge NRF and the National Zoological Gardens of South Africa (NZG) for organizational core grant funding of support of research. BL thanks the University of Venda for research funding for field work. ASWT and KW would like to thank Ezemvelo KZN Wildlife for providing funding for the sample

## Abstract

The samango monkey is South Africa's only exclusively forest dwelling primate and represents the southernmost extent of the range of arboreal guenons in Africa. The main threats to South Africa's forests and thus to the samango are linked to increasing land-use pressure and increasing demands for forest resources, resulting in deforestation, degradation and further fragmentation of irreplaceable habitats. The species belongs to the highly polytypic *Cercopithecus nictitans* group which is sometimes divided into two species *C. mitis* and *C. albogularis*. The number of subspecies of *C. albogularis* is also under debate and is based only on differences in pelage colouration and thus far no genetic research has been undertaken on South African samango monkey populations. In this study we aim to further clarify the number of samango monkey subspecies, as well as their respective distributions in South Africa by combining molecular, morphometric and pelage data. Overall, our study provides the most comprehensive view to date into the taxonomic description of samango monkeys in South Africa. Our data supports the identification of three distinct genetic entities namely; *C. a. labiatus*, *C. a. erythrarchus* and *C. a. schwarzi* and argues for separate conservation management of the distinct genetic entities defined by this study.

## Introduction

The geographical distribution of the arboreal guenon *Cercopithecus albogularis* ranges from central and eastern to southern Africa where it occurs in different evergreen forest types including rainforest, Afromontane and riparian forests, as well as swamp and coastal forests [1]. The species belongs to the highly polytypic *Cercopithecus nictitans* group [2] which is sometimes divided into two species *C. mitis* and *C. albogularis*. Groves [3,4] recognises both species and uses the classification of *C. albogularis* for individuals distributed from Ethiopia to South

collection at those sites. KW would like to thank Novartis South African Veterinary Foundation Wildlife Research Fund (Grant 13/11/2010), and Primate Conservation Inc, for project funding, and Mazda Wildlife Fund for sponsorship of a research vehicle (Number: BD06MZGP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

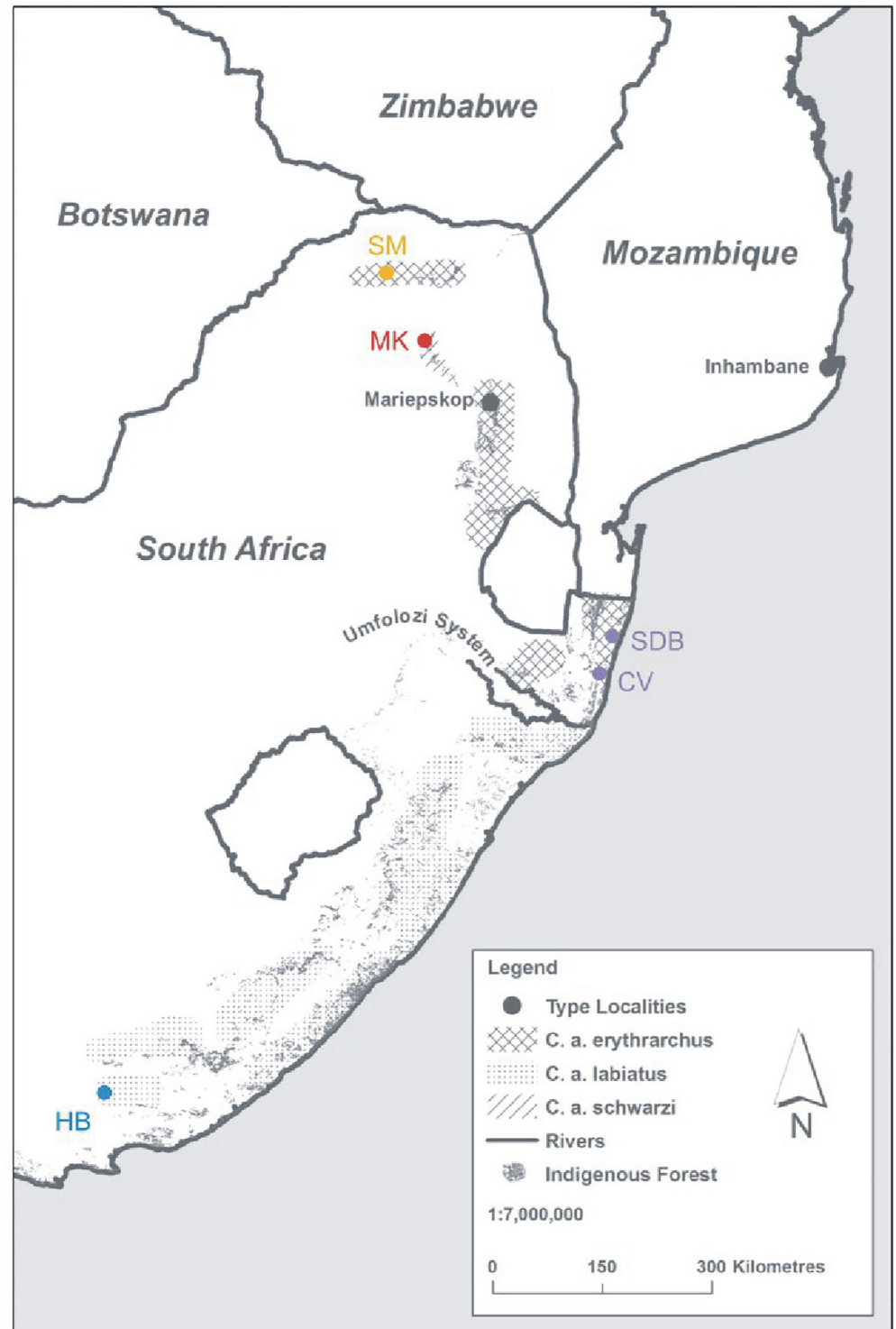
Africa and occurring in South and East Democratic Republic of Congo and North West Angola. Groves classifies *C. mitis* as those individuals found from the Congo-Oubangui River System to East African Rift Valley, as well as in Northern Angola and North Western Zambia. The splitting of the group into two separate species by Groves is based on differences in pelage colourations. Napier [5] followed by Kingdon [6] and Grubb [2] do not recognise *C. albogularis* as a separate species. Furthermore the International Union for Conservation of Nature (IUCN) has not yet recorded *C. albogularis* as a separate taxon and *C. mitis* is listed as least concern [7]. No genetic analysis has been done to date to support one or the other classification and in this paper we follow Groves [3,4].

The number of subspecies of *C. albogularis* recognised in South Africa (locally referred to as "samango monkeys" from the Zulu name iNsimango) is also disputed. Meester [8] followed by Grubb [2] recognise two subspecies namely *C. a. labiatus* [9] (type locality: South Africa), and *C. a. erythrarchus* [9] (type locality: Inhambane, Mozambique). Contrary to this, Roberts [10] followed by Dandelot [11] and Groves [3,4] recognise an additional third samango monkey subspecies in South Africa, namely *C. a. schwarzi* [10] (type locality: Mariepskop, South Africa). The present distribution of *C. a. labiatus* and *C. a. erythrarchus* is closely correlated with the distribution of Afromontane, Scarp and Indian Ocean coastal belt forests in southern Africa and the two subspecies do not overlap in their distribution [9]. *Cercopithecus a. labiatus* is distributed from the Pirie Forest in the Eastern Cape Province north-eastwards to the midlands of the KwaZulu-Natal Province [12]. *Cercopithecus a. erythrarchus* occurs from northern KwaZulu Natal through eastern Mpumalanga and central and eastern Limpopo in South Africa, through Zimbabwe and Mozambique up to Malawi (Fig. 1) [12]. The border between both subspecies appears to be at the St. Lucia and Umfolozi systems [9]. The distribution of *C. a. schwarzi* is typically from the Pilgrims Rest District, but also from Woodbush and intervening territory in the Eastern Transvaal (Mpumalanga) [3,10] (Fig. 1). Only the sub-species' *C. a. labiatus* and *C. a. erythrarchus* are recognised by both the latest national [12] and international conservation assessments (IUCN) [7]. In South Africa the subspecies *C. a. labiatus* is listed as Endangered as it is considered endemic to South Africa whereas the subspecies *C. a. erythrarchus* is listed as Vulnerable [12]. This national divergence is based on an assumed "rescue effect from neighbouring populations (Mozambique)" for the subspecies *C. a. erythrarchus* [12]. The species itself is listed as Vulnerable due to its patchy and highly restricted distribution which in turn can be explained by the very small size of the forest biome and its highly fragmented distribution in the country [12]. Globally, the IUCN lists the subspecies *C. m. erythrarchus* as least concern whereas *C. m. labiatus* is listed as Vulnerable [7]. Given the role that taxonomy has in determining the conservation status of the species /subspecies we believe it is vital to clarify this taxonomic issue. The current subspecies classification of samango monkeys in South Africa is based only on differences in pelage colouration and thus far no genetic research has been undertaken on South African samango monkey populations. In this study we aim to further clarify the number of samango monkey subspecies, as well as their respective distributions in South Africa by combining molecular, morphometric and pelage data.

## Materials and Methods

### Ethical and permit considerations

The project was approved by the National Zoological Gardens of South Africa's Research and Ethics Committee (Project no. P10/27). National Threatened or Protected Species (TOPS) permits (Permit no. 0 27520 and 0 27333) as well as provincial permits (Permit no. OP2407–2013-Kwazulu Natal and OP 0 8052 A-Eastern Cape) were obtained for the capture and anaesthesia



**Fig 1. Sampling localities and distribution of proposed samango species.** Geographic distribution of the proposed samango species (*Cercopithecus labiatus*, Hogsback; *Cercopithecus erythrarchus*, Cape Vidal and Sodwana Bay; *C. e. schwarzi*, Magoebaskloof and Soutpansberg Mountain) from this study. Type localities are included for *C. a. erythrarchus* (Inhambane) and *C. a. schwarzi* (Mariepskop). No exact type locality is available for the sub-species *C. a. labiatus*. Map based on the Endangered Wildlife Trust (EWT) 2004 distribution map and modified including additional distribution data from this study. Sampling localities (SM = Soutpansberg Mountains, MK = Magoebaskloof, SDB = Sodwana Bay, CV = Cape Vidal,

HB = Hogsback) are shown in relation to indigenous forests (DWAF 2004) and the Umfolozi System (Black and White Umfolozi Rivers).

doi:10.1371/journal.pone.0117003.g001

of the animals as well as the sample collection, transport and storage. Animals were not removed from the wild. They were captured and released at the same location.

### Sample collection

A total of 72 samango monkey blood and tissue samples were collected from five populations in Southern Africa (Table 1). All individuals included in the morphological analyses were also included in the microsatellite analyses for all five populations, with the exception of Hogsback, where data was unavailable for a subset of captured individuals. Samples were collected from the southernmost *C. a. labiatus* population in Hogsback (HB; -32.59526, 26.95675), and *C. a. erythrarchus* from two coastal populations in Cape Vidal (CV; -28.12329, 32.55638) and Sodwana Bay (SDB; -27.54692, 32.66939), a population along the Escarpment in Magoebaskloof (MK; -23.88859, 29.99582) and from two troops from the northernmost population in the Soutpansberg mountains (SM; -23.037881, 29.441949). See 1 for the geographic distribution of sampled populations and their acronyms. Trap-door metal cage traps (1.25m x 0.6m x 0.6m) were placed on the ground and were baited with either oranges or apples. Once secured in a trap, the samango monkeys were anaesthetised with an intramuscular injection of a tiletamine/zolazepam (Zoletil, Virbac, South Africa) combination at an estimated dose of 3.5–5 mg/kg. Three milliliters of blood was collected with a syringe and needle from the caudal saphenous vein and divided into serum and EDTA blood tubes (Vacutainer, BD, United Kingdom). The blood samples were placed on ice for up to 6 hours, after which they were frozen at -20°C until analysis. Several full length hairs were plucked out from between the shoulder blades of each monkey and placed in plastic Ziploc bags. Small (3mm in diameter) skin and cartilage samples were obtained from the pinna of each monkey using a normal leather punch. The punch was cleaned and disinfected (F10sc disinfectant, Health and Hygiene (Pty) LTD, South Africa) between samples. The skin and cartilage samples were placed in 1.5ml eppendorf tubes containing 95% ethanol. DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (GmbH, Germany) following the extraction protocol as outlined by the manufacturer. Each animal was

**Table 1. Environmental and demographic characteristics of sampled populations.**

Locality	Forest type	Altitude (masl)	Annual rainfall (mm)	Wet season (min-max°C)	Dry season (min-max °C)	Number samples
Soutpansberg (SM)	Limpopo Mistbelt Forest	1267	737	October-March (13–27)	April-September (4–25)	14
Magoebaskloof (MK)	Limpopo Mistbelt Forest	1518	1112	October-March (10–22)	April-September (3–21)	5
Sodwana Bay (SDB)	KwaZulu Natal Dune Forest	16	934	October-April (17–31)	May-September (12–27)	11
Cape Vidal (CV)	KwaZulu Natal Dune Forest	24	1057	October-April (17–30)	May-September (12–25)	5
Hogsback (HB)	Amatola Mistbelt Forest	1067	898	October-March (8–26)	April-September (4–21)	37

Sample locations are arranged from northernmost (top) to southernmost (bottom). Forest type descriptions follow Von Maltitz *et al.* (2003) and climatic data for each site were acquired using the Wordclim dataset by Hijmans *et al.* (2005). masl = meters above sea level, mm = millimetres and°C = degree Celsius.

doi:10.1371/journal.pone.0117003.t001

assigned to a gender and placed in one of three age classes, namely juvenile, sub-adult or adult based on tooth eruption where sub-adults have at least one permanent canine and adults have all four permanent third molar. Standard morphological measurements were taken (see below) before the individuals were placed in a darkened cage until they had recovered from the anaesthetic to be released back into the wild. A qualified veterinarian registered with the South African Veterinary Council (registration number D97/4000) and experienced in primate capture and anaesthesia was present at all the capture sites, administered the anaesthetic agents and ensured that animals were fully recovered prior to release.

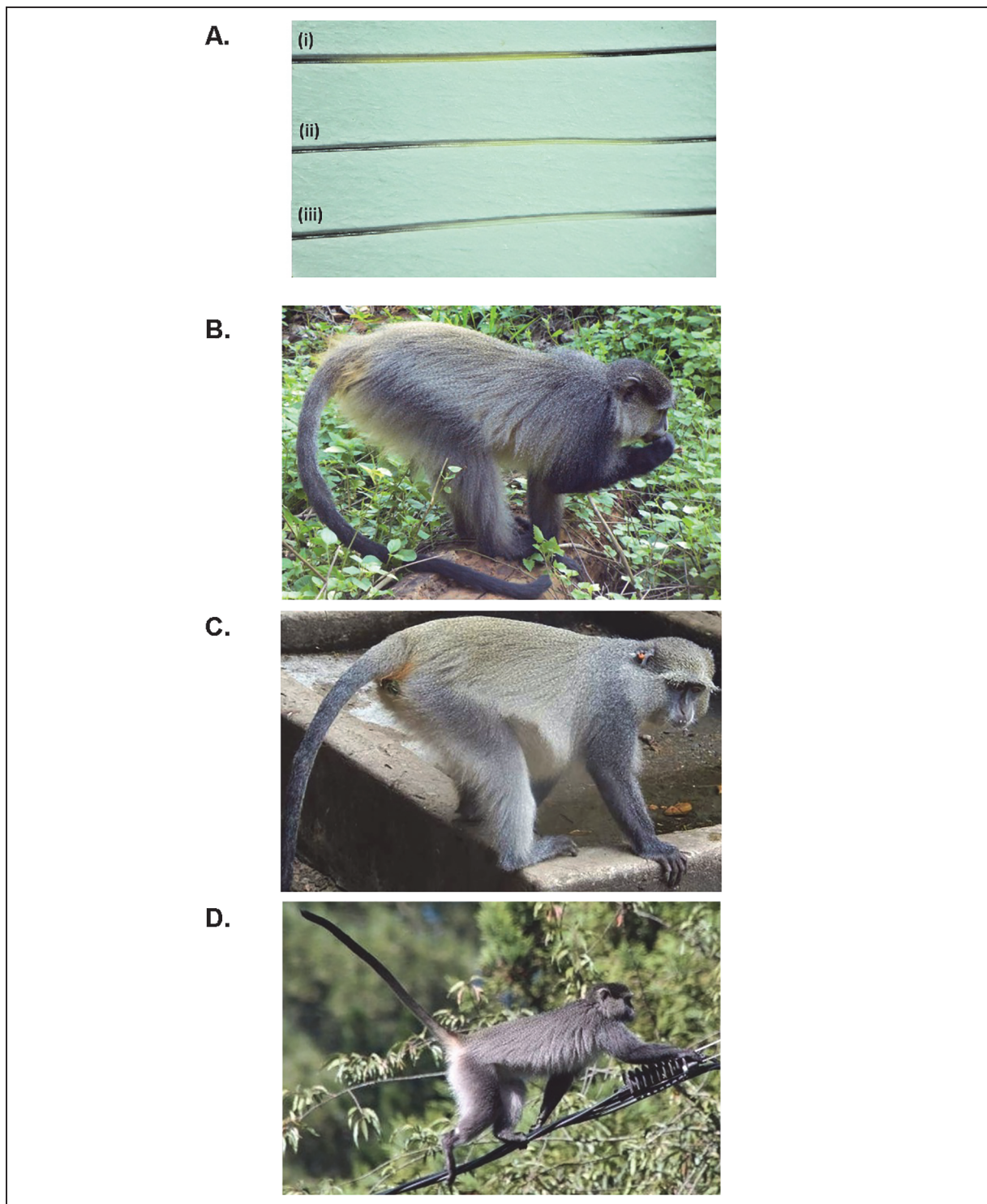
## Morphological measurements

For 101 individuals, body mass (BM, kg) was obtained using a digital hanging scale (25kg Sportsman's digital scale, Rapala, USA), whilst the following standard measurements were obtained (in cm) using both a tape measure and calliper: neck circumference (NC), head and body length from nose-tip to base of tail (HB), tail length (TL), hind foot length, without claw (HF), ear length (EL), ear width at base (EW), canine length (CW) and nipple length (NW). Full body pictures of individuals from different sites were also taken. Measurements made using a tape measure were converted using simple linear regression equations. Apart from nipple length which was only measured in females individuals, variables were log-transformed and analysed statistically for sexual dimorphism and age variation using 2-way Analysis of Variance (ANOVA) in the largest population from Hogsback ( $n = 67$ ). Owing to significant variation between the three age classes, data from adults were compared between all populations using principal component analysis (PCA). All statistical analyses were carried out using the programme PAST version 1.91 [13].

## Hair analysis

In order to assess variation in hair characteristics among the different samango populations sampled we analysed individual guard hairs of adult individuals only. As the Hogsback population was sampled at different months of the year (S1 Table) we chose five individuals per each study month to also allow for analysis of seasonal variation of hair characteristics. The number of adult individuals included for the hair analysis varied as it was depended on how many adults were captured. For each individual monkey we analysed five hairs. We analysed 34 adult individuals (20 females, 14 males) totalling 170 hairs. Hair samples were analysed with the naked eye under a 20 Watt 240 Volt halogen lamp. Hair length was measured using an ordinary ruler with millimetre divisions. Five hairs originating from the same individual were fixed on the base onto a white piece of paper using transparent tape. Analysis under a dissecting microscope proved to be unsatisfactory as the hairs did not fit in the field of view in their entire length. Furthermore, the bright LED light source of the microscope made it much more difficult to detect the subtle colour differences. Hair characteristics analysed included: colour of hair base and tip, number of bands, colour of bands, number of bands per colour, width of bands and total hair length. Bands were counted starting with the first visible light band from the base of the hair. Band length was scored using three categories: bands equal, light/dark band longer and light/dark band double the length.

Hair colour was scored using three different categories: white, light yellow and dark yellow. These categories were established after comparing hairs from all different geographic locations (Fig. 2) in order to identify the range of colours (shades and extremes). Furthermore hairs of all three colour categories were fixed next to each other with transparent tape onto the white analysing sheet, constantly visible during hair analysis. In order to statistically test specifically for species/subspecies differences in hair characteristics, three geographical groups were formed:



**Fig 2. Three different colour categories identified for samango hair samples from different geographic locations.** (A). (i) Cape Vidal (dark yellow category), (ii) Soutpansberg Mountain (light yellow category), (iii) Hogsback (white category). The image also illustrates the typical, alternating dark and light banding of hairs. B to D: Comparisons of pelage colouration between samango monkeys from three geographic locations. B: Adult female, Soutpansberg (Inland) (photo: Birthe Linden), C: Adult female, Cape Vidal (Coast) (photo: Joan Chalmers), D: Adult female, Hogsback (Hogsback) (photo: Kirsten Wimberger).

doi:10.1371/journal.pone.0117003.g002

Inland (Soutpansberg Mountain and Magoebaskloof), Coast (Cape Vidal and Sodwana Bay) and Hogsback (remained on its own). All statistical analysis was carried out using the software R version 3.0.1. Full body *in situ* photographs of individuals from the three populations were compared to examine differences in pelage colouration.

### Microsatellite genotyping

All samples collected during the present study were genotyped for polymorphism at 21 microsatellite loci based on previous studies: D15S108, D14S306, D13S765, D18S536, D12S67, D9S922, D8S1106, D4S243, D5S1457, D10S611, D6S311, D5S1466, D1S518, D1S207, D3S1768, D10S1432, D7S503, D2S1326, D11S956, D11S925 and D17S1290 [14–27]. Amplification was carried out using a 15  $\mu$ l reaction volume and polymerase chain reaction was conducted with PromegaGoTaqFlexi DNA polymerase, (Promega Corporation) which has a 1 x buffer containing 10 milli molar (mM) Tris-HCl (pH 9.0), 50 mM potassium chloride (KCl) and 0.1% Triton X-100. The final reaction conditions were as follows: 1 X PCR buffer, 1.5–2.5 mM MgCl<sub>2</sub>, 200 micro molar ( $\mu$ M) of each 2'-deoxynucleotide triphosphate (dNTP), 10 picomol (pmol) of each of the forward and reverse primer, 1 unit (U) TaqDNA polymerase and 10–20 nano gram (ng) genomic DNA template. The conditions for PCR amplification were as follows; 5 minutes (min) at 95°C initial denaturation, 30 cycles for 30 seconds (sec) at 95°C, 30 sec at 50–65°C and 30 sec at 72°C, followed by extension at 72°C for 20 min. The PCR reaction was carried out in the BOECO TC-PRO Thermal Cycler. PCR products were pooled together and run against Genescan500 LIZ (Applied Biosystems, Inc.) internal size standard on an ABI 3130 Genetic Analyzer. Samples were genotyped using GeneMapper v. 4.0.

### mtDNA sequence analysis

The purified DNA was used as a template to amplify mitochondrial gene fragments using the Cyt B primers L14724 cgaagcttgatgaaaaaccatcggtg and H15149 aaactgcagcccctcagaatgattttgtcctca [28] and 16S primers 16SA cgctgtttaacaaaaacat and 16SB ctccggttgaactcagatca [29, 30]. Polymerase chain reaction (PCR) was carried out using a 25  $\mu$ l reaction volume and was conducted with Thermo Scientific's DreamTaq Green PCR master mix which has a 1 x buffer containing 10 mM TrisHCl (pH 9.0), 50 mM potassium chloride (KCl) and 0.1% Triton X100. The final reaction conditions were as follows: 1 X PCR buffer, 2.5 mM MgCl<sub>2</sub>, 200 M of each dNTP, 5 pmol of each of the forward and reverse primer, 1 unit (U) *Taq* DNA polymerase and 20 ng genomic DNA template. PCR cycling conditions were as follows: Initial denaturation at 95°C for 5 min, annealing at 52°C for 50 sec and extension at 72°C for 1 min, 30 cycles of denaturation at 94°C for 30 sec, annealing at 50°C for 50 sec and extension at 72°C for 1 min. A final extension step of 72°C for 20 min concluded the cycling. After the initial PCR amplification, products were purified according to the Exo/Sap amplicon purification method described by Werle [31]. Cycle sequencing of the PCR products obtained was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the manufacturer's protocol. Sequencing products were purified with the ZR DNA Sequencing Clean-Up Kit (Zymo Research) and sequenced on an ABI 3130 genetic analyser in forward and reverse. The raw sequence data were analyzed using the ABI Prism DNA Sequencer software v3.4.1.

### Analysis of nuclear genetic diversity

MICRO-CHECKER [32] was used to detect possible genotyping errors, allele dropout and non-amplified alleles (null alleles) for each microsatellite locus. Population genetic analyses were carried out at two scales: firstly, with populations defined as recognised subspecies (*C. a. labiatus* and *C. a. erythrarchus*), and secondly as four populations, largely reflecting collection

localities, that were identified from spatially independent multivariate analyses of the two currently recognised subspecies. The latter populations are Hogsback (HB), Sodwana Bay and Cape Vidal combined (SDBCV), Magoebaskloof (MK) and Soutpansberg Mountains (SM). To estimate the levels of genetic diversity within populations, the mean number of alleles per locus ( $N_A$ ), observed heterozygosities ( $H_o$ ), expected heterozygosities ( $H_E$  and  $uH_E$ ) and deviations from Hardy-Weinberg (HW) proportions were calculated using MS TOOLKIT [33] and GENALEX version 6.5 [34,35]. Linkage disequilibrium between pairs of microsatellite loci within each population and locus was evaluated using GENEPOP 4.1.4 [36,37]. Associated probability values were corrected for multiple comparisons using Bonferroni adjustment for a significance level of 0.05. To estimate levels of divergence among populations, spatially explicit and spatially independent analyses were carried out. The former included the estimation of various fixation and differentiation indices by permutation, and  $F_{ST}$ - and  $R_{ST}$ -based analyses of molecular variance (AMOVA) [38], that were carried out in GENALEX [34,35]. Spatially independent analyses included assignment tests; principal coordinates analysis (PCoA), factorial correspondence analysis (FCA) and Bayesian cluster analysis in STRUCTURE.

## Population and regional genetic structure

The level of genetic differentiation was determined among and within the currently recognized subspecies, and the four distinctive populations. The genetic relationship between populations and individual assignments of samples was inferred via a Bayesian clustering analysis using the statistical programme STRUCTURE version 2.3.3 [39]. The programme was run without prior population information (option USEPOPINFO = 0, no LOCPRIOR), to ensure that the pre-defined populations were in agreement with the genetic data. STRUCTURE was initially run for all 72 samples with 20 replicates from  $K = 1-9$ , with a run-length of 1 million repetitions of Markov chain Monte Carlo (MCMC), following the burn-in period of 100,000 iterations. A second STRUCTURE run that included only *C. a. erythrarchus* ( $n = 35$ ) samples was run with the same settings for  $K = 1-5$ . The  $K$  with the greatest increase in posterior probability ( $\Delta K$ ) [40] was identified as the optimum number of sub-populations using STRUCTURE HARVEST [41]. CLUMPP [42] generated the average of these 20 runs for the appropriate  $K$ , and the results were visualised in DISTRUCT [43].

## Analysis of mtDNA data

Alignment and vetting of sequence data were carried out using Bioedit v7.0.9.0 [44] and FinchTV v 1.4 (Geospiza Inc.), and published sequences for closely related species were included in alignments to confirm amplification of the correct target region. Standard population genetic diversity and differentiation indices (number of haplotypes, haplotype diversity,  $h$ , nucleotide diversity,  $\pi$ , the average number of nucleotide differences,  $k$  and  $\phi_{ST}$ ) were calculated using DNAsp v 5.10.01 [45] based on combined mitochondrial sequence data. Maximum Parsimony (MP), Maximum Likelihood (ML) and Bayesian MCMC phylogenetic analyses were conducted for each gene region separately, and for combined datasets. Maximum Parsimony (MP) haplotype networks [46] were generated in Network v 4.6.1.0 ([www.fluxus-engineering.com](http://www.fluxus-engineering.com)) and Splittree v 4.10 [47] (10 000 bootstrap replicates) using the default settings. Statistical model selection was carried out for each dataset in MEGA v 5.0 [48] and jModeltest [49]. ML phylogenetic trees were generated in MEGA v 5.1 [48], with branch support evaluated using bootstrap-resampling (10 000 replicates). Bayesian phylogenetic trees were estimated using MCMC in MrBayes Version 3.1.2 [50] (with one million generations, trees sampled every 100 generations) and BEAUti/BEAST v2.1.1 [51] (twenty million MCMC generations, priors as in Hart et al. (2012) and the whole tree constrained with a lognormal



distribution around 6 My, mean  $M = 1.79$ , standard deviation  $S = 0.1$ ), with a ten to 25% burn-in. Divergence date estimates were inferred using a Bayesian approach implemented in BEAST [52]. These analyses were run separately for the two gene regions and for a combined (concatenated) dataset. The results of four independent BEAST runs were checked for adequate mixing and convergence using Tracer 1.5. Once convergence was achieved, BEAST tree files were combined using LogCombiner, summarized using TreeAnnotator 1.5.3 and visualized in FigTree 1.3.1. CorelDraw was used to edit the haplotype networks and phylogenetic trees.

## Results

### Morphometric analyses

As expected, age explained a significant proportion of morphometric variance (8–54%SS) in all variables except for ear length (Table 2). Significant sexual dimorphism was present in all variables except for ear length and width, neck circumference and head and body length (Table 2). Significant interaction (age-sex) terms for all variables except ear length (Table 2) signalled divergent growth projectories in males and females associated with secondary sexual dimorphism and progressively greater increases in most male morphometric variables following puberty. This is exemplified by plots of mean body mass for both sexes in the three age classes; sexual dimorphism is negligible in juvenile and subadults but very marked in adults (S1 Fig). Based on these significant differences between age and sex classes, for subsequent analyses of population variation we excluded juveniles and subadults and analysed males and females separately. From PCA of seven log-transformed morphometric variables, both males and females showed Soutpansberg animals to be distinctly larger-sized (higher PC1 scores) than other populations, with Magoebaskloof (in females) being somewhat intermediate (Fig. 3A and 3B; S2 Table). Soutpansberg animals tended to have disproportionately longer tails than in other populations, as indicated by positive scores for PC2 on the scatterplots and positive loadings for tail length on PC2 compared with negative loadings for other variables (Fig. 3A and B).

**Table 2. Results of 2-way ANOVA of sexual dimorphism and age variation in 67 samangos from Hogsback.**

Variable	Age % SS	F(2,66)	Sex % SS	F(2,66)	Interaction % SS	F(2,66)
Weight (kg)	53.8	143.2 ***	2.1	11.4**	23.8	63.5***
Neck circumference	42.4	41.3 ***	1.9	3.7 NS	17.0	16.6***
Head-body	58.7	85.8 ***	0.0002 NS	0.0008 NS	20.3	29.7 ***
Tail	24.1	21.5 ***	8.6	15.4 **	21.1	18.8 ***
HF	7.9	17.5 ***	33.8	37.5 ***	18.0	19.9 ***
Ear length	1.1	0.8 NS	4.6	1.6 NS	8.3	77.5 NS
Ear width	12.9	11.1 **	2.5	1.1 NS	7.6	3.2 *
Canine length	8.1	26.9 ***	34.5	57.4 ***	25.4	42.3 ***

Significance of F-values indicated as non-significant (NS)

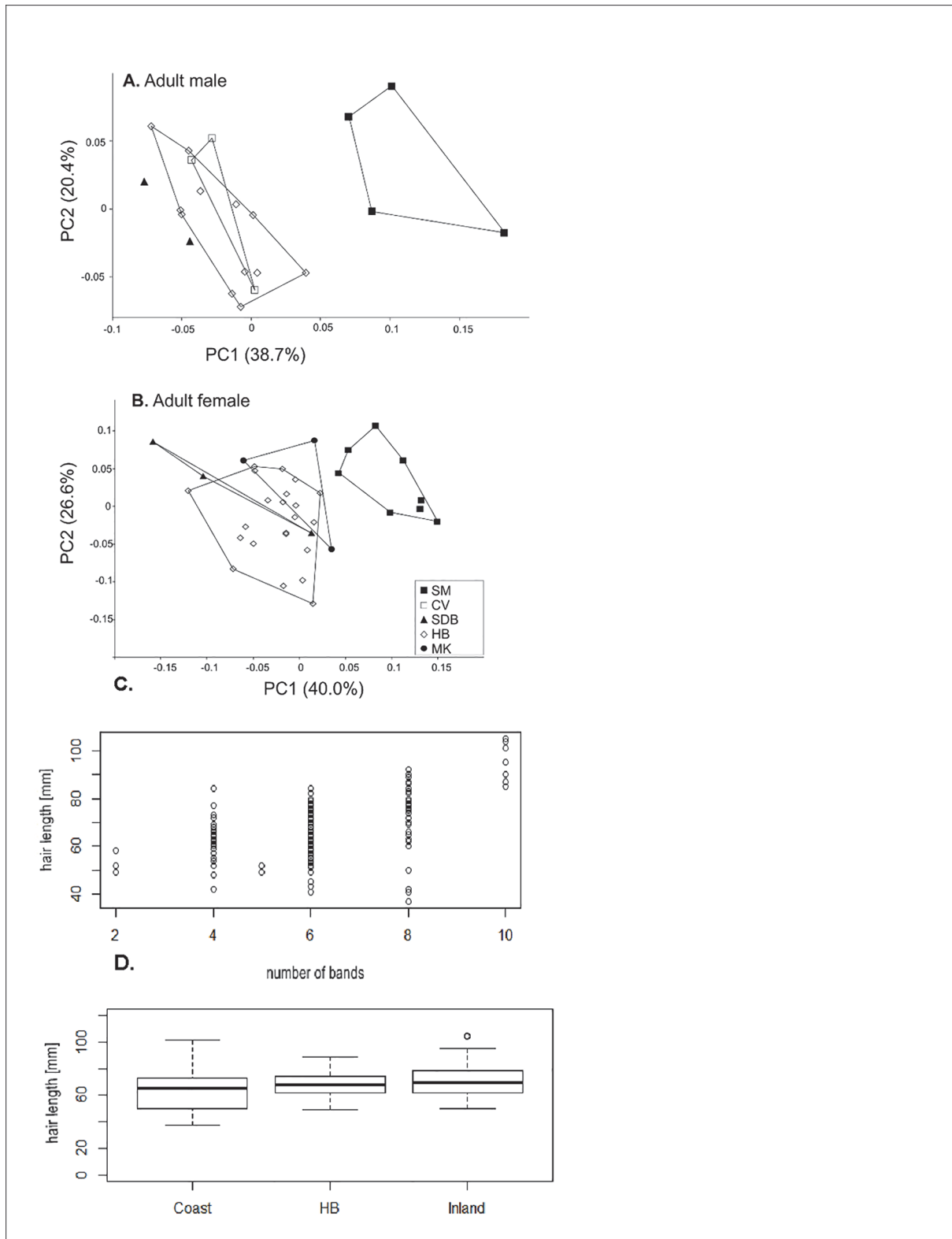
$P < 0.01$  (\*)

$P < 0.001$  (\*\*)

$P << 0.0001$  (\*\*\*)

SS = Sum of Squares. Value in parentheses following F-values represents degrees of freedom). Except for body mass (kg), all measurements were expressed in centimetres.

doi:10.1371/journal.pone.0117003.t002



**Fig 3. PCA of log-transformed morphometric variables.** (A) adult male and (B) female samangos from five populations (closed squares = Soutpansberg Mountain (SM); open squares = Cape Vidal (CV); closed triangle = Sodwana Bay (SDB); open diamonds = Hogsback (HB); closed circles—Magoebaskloof (MK). (C) Pearson's Correlation between hair length and number of bands per hair from all study sites. (D) Boxplots illustrating variation of hair length between three different geographic locations (Coast = Cape Vidal, Sodwana Bay, HB = Hogsback and Inland = Soutpansberg Mountain, Magoebaskloof).

doi:10.1371/journal.pone.0117003.g003

**Table 3. Total number of hairs found per colour category across the three geographic locations: Inland (Soutpansberg Mountain, Magoebaskloof), Coast (Cape Vidal, Sodwana) and HB (Hogsback).**

	dark yellow	light yellow	white	black double	black longer	black equal
<b>inland</b>	1	55	/	/	55	1
<b>coast</b>	30	9	/	/	21	18
<b>HB</b>	/	11	64	72	3	/

doi:10.1371/journal.pone.0117003.t003

### Hair analysis

When analysing variability in hair colour, significant differences between the Soutpansberg Mountains and Magoebaskloof (Inland), coastal populations (Sodwana Bay and Cape Vidal) and Hogsback populations were found ( $\chi^2 = 232.52$ ,  $df = 4$ ,  $p\text{-value} < 0.0001$ ), with Inland showing light yellow bands, Coast showing mostly dark yellow bands and Hogsback showing white bands (Table 3). The relative length of black and white hair bands also differed significantly in frequency of three different categories between the three populations ( $\chi^2 = 208.52$ ,  $df = 6$ ,  $p\text{-value} < 0.0001$ ). Black bands were just longer than the light bands in 55 out of 56 Inland individuals, double the length of light bands in 72 out of 75 Hogsback individuals and somewhat intermediate in Coast individuals with roughly equal case where black bands were longer (21) or equal (18) (Table 3). A Pearson's Correlation analysis of hair characteristics revealed a significant correlation between hair length and number of bands ( $t = 8.2603$ ,  $df = 168$ ,  $p\text{-value} = < 0.0001$ ) and we found that, except in two hairs, the total number of bands was always an even number (2, 4, 6, 8, 10) and that the number of black bands and light bands was always equal in numbers (Fig. 3C). The base and tip of the hairs were found, without exception, to have the same colour in all individuals across study sites, with the base being white and the tip being black. An analysis of variance (ANOVA) of hair length showed significant differences between the three study sites ( $df = 2$ ,  $f = 4.52$ ,  $p\text{-value} = 0.0122$ ). A post hoc Tukey HSD test showed that the Inland and Coast populations differed significantly in hair length ( $p\text{-value} = < 0.01$ ); hair lengths between Hogsback and Coast as well as Hogsback and Inland did not differ significantly (Fig. 3D). Using the Hogsback population only for analysis of seasonal variation in hair length (ANOVA), we did not find significant differences of hair length between the different sampling months ( $df = 2$ ,  $f = 0.41$ ,  $p\text{-value} = 0.665$ ). Furthermore, a t-test showed no significant differences between Hogsback males and females regarding hair length ( $t = -0.56$ ,  $df = 70.36$ ,  $p\text{-value} = 0.5801$ ). When comparing full body in situ photographs of individuals from the three populations, pelage colouration differences are clearly visible (Fig. 2B to 2D). It can be seen that the Coast individual has an overall lighter appearance when compared to the Hogsback and Inland individuals. A marked difference regarding the darkness of individuals are the black arms of the Hogsback and Inland monkeys compared to the grey arms of the Coast monkey. The yellow wash or shine on the back is most visible and most extensive in the Coast individual and near to absent in the Hogsback individual. The ischial regions also show clear colouration differences being most prominent and orange in the Coast individual, yellow in the Inland individual and white in the Hogsback individual. Further colour differences worth mentioning are the very conspicuous white ear tufts and white underside of the tail (about the first quarter) in the Hogsback monkey compared to less obvious white ear tufts and dark tail undersides in the Inland and Coast monkeys.

Table 4. Genetic variation estimates.

Subspecies	Microsatellites						mtDNA				
	N	N <sub>a</sub>	N <sub>eA</sub>	H <sub>O</sub>	H <sub>E</sub>	uH <sub>E</sub>	N	h	Hd	π	k
<b><i>C. a. labiatus</i></b>	<b>36.9</b>	<b>3.33</b> <b>±0.35</b>	<b>1.87±0.15</b>	<b>0.42</b> <b>±0.06</b>	<b>0.38</b> <b>±0.05</b>	<b>0.39±0.05</b>	<b>7</b>	<b>2</b>	<b>0.57</b>	<b>0.0019</b>	<b>1.71</b>
Hogsback ( <i>C. a. labiatus</i> )	36.9	3.33 ±0.35	1.87±0.15	0.42 ±0.06	0.38 ±0.05	0.39±0.05	7	2	0.57	0.0019	1.71
<b><i>C. a. erythrarchus</i></b>	<b>31.81</b>	<b>5.00</b> <b>±0.41</b>	<b>2.990</b> <b>±0.257</b>	<b>0.48</b> <b>±0.04</b>	<b>0.61</b> <b>±0.04</b>	<b>0.617</b> <b>±0.039</b>	<b>21</b>	<b>6</b>	<b>0.81</b>	<b>0.004</b>	<b>3.56</b>
Soutpansberg ( <i>C. a. erythrarchus</i> )	13.4	2.71 ±0.31	1.92±0.16	0.41 ±0.06	0.4±0.05	0.41±0.05	4	2	0.5	0.0017	1.5
Magoebaskloof (MK; <i>C. a. erythrarchus</i> )	4.1	2.86 ±0.22	2.27±0.18	0.50 ±0.08	0.49 ±0.05	0.56±0.06	5	2	0.4	0.0004	0.4
Sodwana Bay and Cape Vidal ( <i>C. a. erythrarchus</i> )	14.3	3.81 ±0.31	2.71±0.25	0.53 ±0.06	0.57 ±0.04	0.59±0.04	12	2	0.53	0.0012	1.06
<b>Total (mean±SE)</b>	<b>17.2</b>	<b>3.18</b> <b>±0.16</b>	<b>2.19±0.1</b>	<b>0.46</b> <b>±0.03</b>	<b>0.46</b> <b>±0.03</b>	<b>0.49±0.03</b>	<b>28</b>	<b>8</b>	<b>0.87</b>	<b>0.005</b>	<b>4.37</b>

The sample size (N, mean across microsatellite loci and for combined mtDNA genes), number of alleles (N<sub>a</sub>±SE), number of effective alleles (N<sub>eA</sub>), observed (H<sub>O</sub>) and expected (H<sub>E</sub>) heterozygosities, and unbiased expected heterozygosity (uH<sub>E</sub>) over all loci; the number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), and average number of nucleotide differences (k) are given for the combined 16S and Cyt b dataset (911 bp). Populations are defined as the currently recognised subspecies (in bold) as well as among four populations of samango monkeys.

doi:10.1371/journal.pone.0117003.t004

### Genetic analysis: Populations defined as currently recognized subspecies

**Analysis of nuclear and mitochondrial genetic diversity.** In the *C. a. labiatus* (Hogsback, HB) samples, only two loci significantly deviated from HWE (loci D12S67 and D2S1326), and two loci were monomorphic (loci D11S956 and D6S311). Among *C. a. erythrarchus* individuals, the pattern was strikingly different, with 11 loci deviating from HWE. Each pair of loci in each population was tested for linkage disequilibrium (420 tests), and 81 of these were significant. However, only eight were significant within *C. a. labiatus*, and only three locus pairs showed significant evidence of linkage in both populations (loci D5S1466 and D10S1432, D10S1432 and D2S1326, D5S1457 and D2S1326). Genetic diversity was lower among *C. a. labiatus* than among *C. a. erythrarchus* in terms of, among others, the number of alleles, heterozygosity and mtDNA haplotype diversity (Table 4). The mean number of alleles per locus over the two populations ranged between 1.5 and 8, with an overall mean of 4.17 over loci and populations. All 37 *C. m. labiatus* (HB) samples exhibited at least one of 15 private alleles found at 11 loci in that population. Additionally, at least one of the 50 private alleles found among *C. m. erythrarchus* samples, i.e. alleles that were not found among *C. m. labiatus* (HB) samples, was exhibited by all individuals (n = 35) from this population. This pattern indicates strong divergence and restricted gene flow between the currently recognized subspecies.

**Spatially explicit analyses.** Various fixation indices (F<sub>ST</sub>, G<sub>ST</sub>), corrected (standardized) fixation indices (G<sup>\*</sup><sub>ST</sub>) and a pure estimate of differentiation (D<sub>EST</sub>) were calculated to estimate genetic divergence between the two putative subspecies. Pairwise estimates of population differentiation were all high and highly significant (Table 5, P = 0.0001). The maximum possible value for G<sub>ST</sub>, given the dataset was G<sub>STmax</sub> = 0.33±0.04. Locus D6S311 consistently showed the strongest signal of differentiation, and locus s was the only locus that did not show significant (P<0.05) differentiation among the two subspecies (locus s P = 0.479) based on G<sub>ST</sub>.

**Table 5. Microsatellite-based fixation indices.**

	Microsatellites						mtDNA	
	F <sub>ST</sub>	G <sub>ST</sub>	G' <sub>STNEI</sub>	G' <sub>STHED</sub>	G'' <sub>ST</sub>	D <sub>EST</sub>	φ <sub>ST</sub>	G <sub>ST</sub>
<b>Between subspecies</b>	0.159±0.029	0.151±0.029	0.262±0.044	0.454±0.085	0.526±0.085	0.358±0.082	0.55	0.14
<b>Between populations</b>	0.251±0.028	0.212±0.029	0.264±0.034	0.484±0.066	0.518±0.065	0.345±0.065	0.79	0.404

Standardized fixation indices and Jost's (D<sub>EST</sub>) estimate of differentiation based on the two recognized subspecies and between populations defined as the four clusters identified in the spatially independent analyses (all P<0.001). G<sub>ST</sub> (Nei 1973) and φ<sub>ST</sub> (Hudson et al. 1992) based on the two subspecies and four populations (both P<0.001).

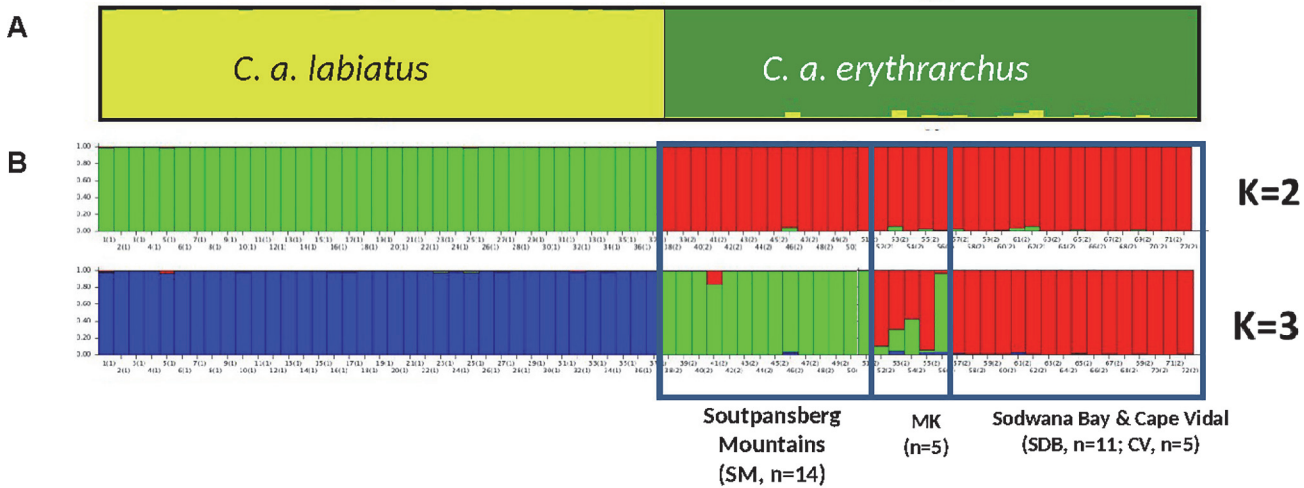
doi:10.1371/journal.pone.0117003.t005

Corrected F<sub>ST</sub> (F'<sub>ST</sub>) and R<sub>ST</sub> were estimated during AMOVA [53]. The F<sub>ST</sub>-based AMOVA showed that 25% of the variance in allele frequencies was explained among populations, and 14% among individuals. F<sub>STmax</sub> was 0.467 and F<sub>ST</sub> was estimated at 0.25 (P = 0.0001), giving an F'<sub>ST</sub> of 0.526. F<sub>IS</sub> (0.188) and F<sub>IT</sub> (0.391) were also high and highly significant (P<0.001), indicating strong population structure. The R<sub>ST</sub>-based AMOVA generated a highly significant (P<0.001) R<sub>ST</sub>-value of 0.157, with 16% of variance distributed among the two subspecies. Estimates based on the combined mtDNA sequence data for the two subspecies also indicated significant divergence (Table 5). Interestingly, these estimates all increased when populations were defined at a finer scale i.e. collection localities.

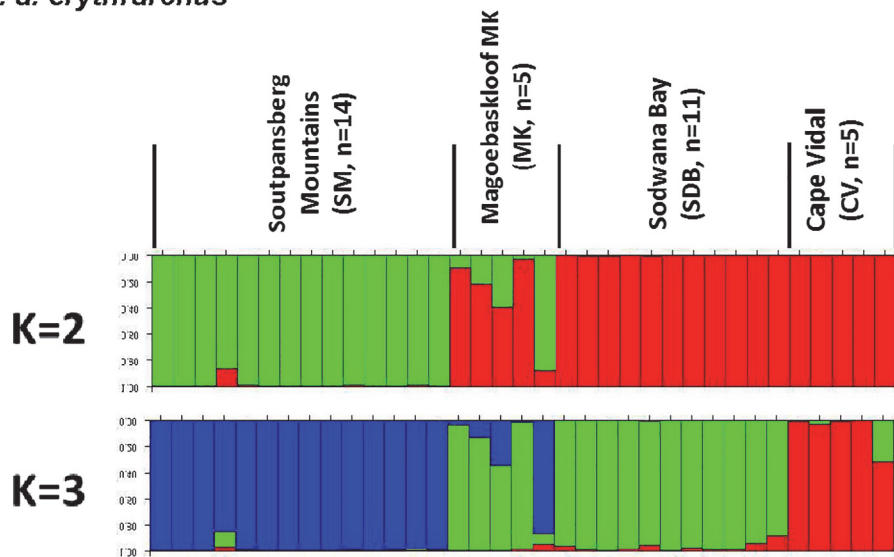
**Spatially independent analyses.** Bayesian cluster analysis in STRUCTURE clearly identified the two subspecies, with members of each subspecies being highly likely to belong to that subspecies based on their multi-locus genotypes (Fig. 4A and 4B). Posterior probabilities (Ln) using Bayesian admixture analysis were calculated for K = 1–9, with K = 2 being identified as the most likely true K value based on STRUCTURE HARVESTER results (S2 Fig.; K = 2 displayed the greatest posterior probability). Results indicated two distinct clusters for *C. a. labiatus* and *C. a. erythrarchus* (Fig. 4A and 4B). All 72 individuals were correctly assigned to their subspecies based on their multi-locus genotype (S3 Fig.), indicating that the two taxa are highly distinctive.

**Multivariate Principal Coordinates and Correspondence analyses.** *C. a. labiatus* individuals from Hogsback formed a tight, distinctive cluster in the principal coordinate's analysis and are clearly divergent from all *C. a. erythrarchus* individuals (S4 Fig.), and the forest and vilage troops appear to be genetically homogenous based on their multi-locus genotypes. Samples of the latter subspecies, however, appear to show higher levels of structure, and fall into three groups that largely reflect their sampling locality (S4 Fig.). Factorial correspondence analysis conducted in GENETIX showed a very similar pattern to the principle coordinates analysis in that *C.a. labiatus* individuals from Hogsback formed a tight, distinct group, and *C.a. erythrarchus* from other sampling localities were also distinct from each other, although they were clearly more similar to each other than to HB samples (S5 Fig.). However, one Soutpansberg Mountain sample (SMBT6) appears to be more closely related to individuals from Magoebaskloof (S4 Fig.).

**Phylogenetic trees.** Two distance measures based on microsatellite data were used to construct UPGMA phylogenetic trees of individuals with bootstrapping over loci (1000 replicates): DSW and Cavalli-Sforza and Edwards Dc (chord distance, S6 Fig.). Both methods produced similar results that reflected the multivariate analyses above. All *C. a. labiatus* individuals from Hogsback form their own distinctive clade (Clade A, S6 Fig.). The Soutpansberg Mountain and Magoebaskloof *C. a. erythrarchus* individuals appear to be very closely related (Clades C and D



C. Only *C. a. erythrarchus*



**Fig 4. Bayesian cluster analysis in STRUCTURE.** (A) The average cluster membership (over 20 runs) for  $K = 2$  of microsatellite genotypes of *C. a. labiatus* and *C. a. erythrarchus*. (B) Example bar plots of runs where  $k = 2$  and  $k = 3$ . (C) Example histograms from STRUCTURE analysis for  $K = 2$  and  $K = 3$  where only *C. a. erythrarchus* individuals are included. Each individual is represented by a single horizontal line, with lengths proportional to the estimated membership in each cluster.

doi:10.1371/journal.pone.0117003.g004

respectively, [S6 Fig.](#)), but are quite distinct from the coastal Sodwana Bay and Cape Vidal individuals (Clade B, [S6 Fig.](#)).

Fine-scale analyses: Populations defined as collection localities

Results based on the full microsatellite dataset of 72 individuals suggest that further structure exists among the *C. a. erythrarchus* individuals sampled. A second STRUCTURE analysis was therefore run that omitted *C. a. labiatus*, which is clearly very distinct from all *C. a. erythrarchus* populations. Settings for this STRUCTURE analysis were 100 000 burnin, 1 million MCMC, admixture model, no LOCPRIOR, and 20 replicates for  $K = 1$  to  $K = 5$ . STRUCTURE-HARVESTER identified the most likely number of clusters as  $K = 2$  ([Fig. 4C](#)). Structure results for the 20 runs (1 million) were averaged using CLUMPP, and visualized using DISTRUCT.

The two *C. a. erythrarchus* clusters strongly reflect their collection localities in the Soutpansberg Mountains, and in the coastal Sodwana Bay and Cape Vidal populations (Fig. 4C). However, those samples collected in Magoebaskloof are not as distinctive. The observed population structure appears to reflect coastal and inland forms of *C. a. erythrarchus*, with restricted gene flow between them. However, samples from Magoebaskloof (MK) resemble hybrids in that most are equally likely to have originated in either of the aforementioned coastal and inland clusters.

**HWE, LD, genetic diversity, private alleles.** Population genetic analyses were repeated with populations defined as the three *C. a. erythrarchus* clusters identified by STRUCTURE, and the Hogsback *C. a. labiatus* population. Where deviations from HWE remained the same for *C. a. labiatus*, three loci were monomorphic in the Soutpansberg Mountain (SM, loci D6S311, D10S1432, D11S956) population, and two deviated from HWE (D4S243 and D1S207). Loci D10S1432, D1S207 and D5S1466 deviated from HWE in the MK population, and D11S956 was monomorphic. Loci D2S1326, D10S1432, D5S1466, D10S611, D8S1106 and D12S67 deviated from HWE in the SDCV population. 37 of the 840 pairwise linkage tests were significant, but no locus pair showed significant linkage in all four populations. Over all four populations, GENEPOP identified six pairs of loci as significantly linked (S3 Table). Although 171 pairwise comparisons were affected by missing data and monomorphic loci, it does not appear that linkage disequilibrium will influence these results. Across all four populations and all loci, the mean number alleles ( $N_a$ ) was 3.4 and mean observed heterozygosity ( $H_o$ ) was 0.46 (Table 4). Observed heterozygosity and number of alleles was highest in the SDB and CV population. The highest haplotype and nucleotide diversity was found among *C. a. labiatus* individuals from Hogsback (Table 4) and the lowest among *C. a. erythrarchus* from Magoebaskloof (Table 4). All *C. a. labiatus* individuals still exhibited at least one of the 12 private alleles in that population, but the each of the *C. a. erythrarchus* clusters also exhibited private alleles: SM exhibited one private allele (exhibited by one individual, SMN8), MK had five private alleles (at least one of which is exhibited by all samples). The combined SDB and CV populations exhibited 20 private alleles, and all individuals except one (SDB 3) exhibited at least one of these private alleles.

**Spatially explicit analyses.** Most indices of overall population structure increased compared to the previous analysis comparing only the two subspecies (Table 5).

**AMOVA-based estimates of population differentiation.** Overall,  $F_{ST}$  based on AMOVA was 0.29 ( $P < 0.001$ ),  $F_{STmax}$  was 0.50, and corrected  $F_{ST}$ , therefore, equals 0.58. Also, 29% of variance is explained among the four populations, and 9% among individuals. Pairwise  $F_{ST}$  values estimated during AMOVA (Table 6) show that the HB population is consistently significantly and highly differentiated from all other populations, and that within *C. a. erythrarchus*, Soutpansberg Mountain population is most different from the Sodwana Bay and Cape Vidal population. The analogous analysis of the two subspecies produced an  $F_{ST}$  of 0.24 ( $P = 0.0001$ ),

**Table 6. AMOVA-based estimates of population differentiation based on microsatellites.**

Population	HB	SM	MK	SDBC
Hogsback (HB, <i>C. a. labiatus</i> )	-	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Soutpansberg Mountains (SM; <i>C. a. erythrarchus</i> )	0.389 (0.156)	-	0.000 (0.001)	0.000 (0.000)
Magoebaskloof (MK; <i>C. a. erythrarchus</i> )	0.371 (0.752)	0.221 (0.320)	-	0.000 (0.000)
Sodwana Bay and Cape Vidal (SDB, CV; <i>C. a. erythrarchus</i> )	0.304 (0.375)	0.251 (0.146)	0.163 (0.265)	-

Pairwise  $F_{ST}$  ( $R_{ST}$  values indicated in brackets) values estimated during AMOVA (below diagonal) and the associated probabilities (above diagonal).

doi:10.1371/journal.pone.0117003.t006

**Table 7. AMOVA-based estimates of population differentiation based on mtDNA.**

Population	HB	SM	MK	CV	SDB
Hogsback (HB; <i>C. a. labiatus</i> )	-	0.65	0.86	0.85	0.85
Soutpansberg Mountains (SM; <i>C. a. erythrarchus</i> )	0.29	-	0.85	0.66	0.82
Magoebaskloof (MK; <i>C. a. erythrarchus</i> )	0.33	0.38	-	0.97	0.97
Cape Vidal (CV; <i>C. a. erythrarchus</i> )	0.51	0.63	0.66	-	1
Sodwana Bay (SDB; <i>C. a. erythrarchus</i> )	0.54	0.64	0.69	1	-

Pairwise population  $\phi_{ST}$  (above the diagonal) and  $G_{ST}$  (below the diagonal).

doi:10.1371/journal.pone.0117003.t007

which is lower than the estimate between these two *C. a. erythrarchus* populations. Importantly, all populations are significantly different from one another. Overall  $R_{ST}$  was 0.23 ( $P < 0.001$ ), and pairwise  $R_{ST}$ -values also indicate high, and highly significant, differentiation among the four predefined populations. The largest differences are again between *C. a. labiatus* and each of the three *C. a. erythrarchus* subpopulations (all pairwise  $R_{ST} > 0.15$ , Table 6), although the latter are also significantly different from each other. Analogous results based on the combined mtDNA dataset corroborated the finding of significant population differentiation (Table 7, all  $P < 0.001$ ), although the pattern of divergence was different.

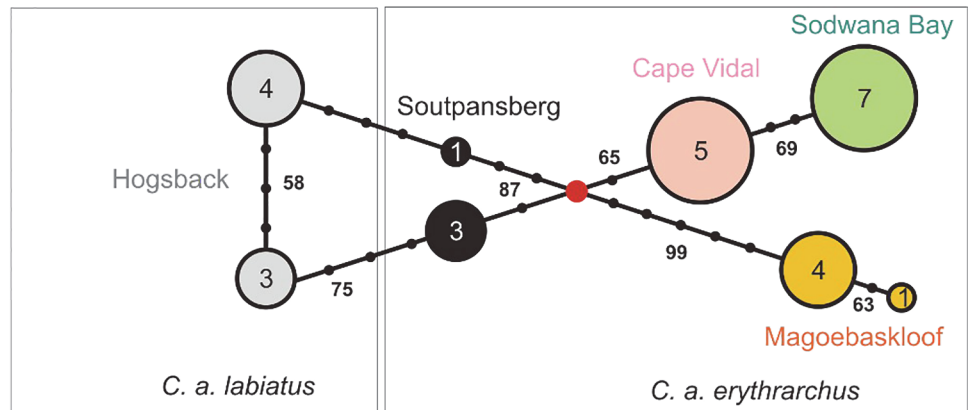
### Spatially independent analyses

**Population assignment test.** Overall, 99% of individuals were correctly assigned to their population. The only exception was one individual from Magoebaskloof (MK5), which was assigned to the Soutpansberg Mountain population.

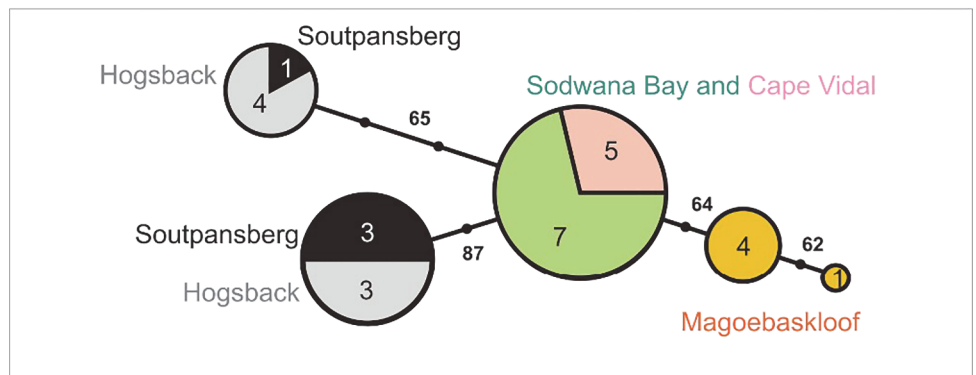
**Phylogenetic analysis of mtDNA data.** A total of 911 bp of mtDNA sequence data (416 bp of cyt b and 495 bp of 16S; Genbank accession numbers KP120559 to KP120615) were generated in 28 samples (HB = 7, SDB = 7, CV = 5 and SM = 4 and MK = 5). Each gene region contained five haplotypes, and when analysed together, there were a total of eight haplotypes (haplotype diversity = 0.87). MP-based haplotype networks largely reflect the pattern of divergence estimates from analyses of population differentiation (Fig. 5). Haplotypes are largely restricted to single populations, and are therefore not likely to represent numts (nuclear copies of mitochondrial genes). Interestingly, some individuals from Hogsback (*C. a. labiatus*) share 16S haplotypes with individuals from the SM population (*C. a. erythrarchus*). The pattern of reticulate evolution reflected in the haplotype networks cannot be represented as a bifurcating phylogenetic tree, but is evident as reduced nodal support in the ML and Bayesian trees (S7–12 Figs.). The combined ML analysis placed *C. a. labiatus* as the basal lineage separate from the *C. a. erythrarchus*. In addition, ML analysis did not favour a monophyletic *C. a. erythrarchus* group but rather split the samples into sister groups as shown in S9 Fig. This pattern was identical to that of the cyt b Bayesian tree (S11 Fig.). However, when 16S is incorporated into analyses, the shared haplotypes between HB and SM changes the tree topology (S12 Fig.). The Bayesian tree constructed in BEAST (S13 Fig.) reflects the complexity of the relationships within *C. a. erythrarchus* and the reticulate nature of these relationships in that it places the SM population as sister to CV. It also clearly shows the distinctiveness of the HB, MK and SDB populations.



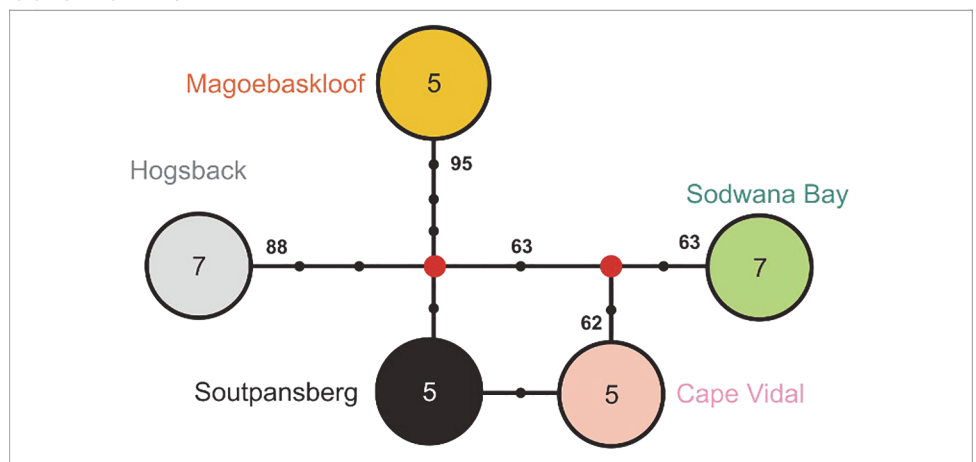
(a) *cytb* and 16S (n=28)



(b) 16S (n=28)



(c) *cytb* (n=29)



**Fig 5. Maximum Parsimony reduced-median haplotype networks.** (a) Combined 16S and *cyt b* data, (b) 16S data and (c) cytochrome b data. The sizes of circles, and numbers in circles, represent the number of individuals sharing a haplotype, red circles represent missing haplotypes, and tick marks along connecting lines are nucleotide changes. Numbers in bold along the connecting lines are bootstrap support.

doi:10.1371/journal.pone.0117003.g005

## Discussion

### Morphological variation

Morphometric variables in isolation, could not distinguish the two genetic lineages; *C. a. labiatus* and *C. a. erythrarchus* described by Meester [8] followed by Grubb [2]. Instead, the Soutpansberg (SM) and northern Escarpment (Magoebaskloof, MK) populations currently classified as *C. a. erythrarchus* are distinctly heavier and larger in most body metrics than both the Hogsback (Eastern Cape, HB) population of *C. a. labiatus* as well as the coastal forest populations of *C. a. erythrarchus* from Sodwana Bay (SDB) and Cape Vidal (CV) in KwaZulu-Natal. Body size in many animals is known to be strongly influenced by environmental factors resulting in well established “rules” of geographic variation, e.g. towards increasing body size at higher latitudes and elevations (Bergman’s Rule) [54]. Whilst Bergmann’s Rule might explain the larger body size of SM animals compared with those from coastal populations (SDB, CV), the data from HB are anomalous since both the high elevation, high latitude and similar climatic conditions to the Soutpansberg of this site would predict the largest or at least equal body size. It may be that nutritional constraints or other environmental factors limit body size in the Hogsback population. We found clear pelage colour polymorphism in South African samango monkeys and were able to identify three distinct geographic colour morphs: Hogsback, Inland (Soutpansberg, Magoebaskloof) and Coast (Cape Vidal, Sodwana Bay). This finding supports the presence of three sub-species as first proposed by Roberts [10] and as currently accepted by Groves [3]. Within the three geographic colour morphs we found that the HB individuals (currently classified as *C. a. labiatus*) show hardly any yellow colouration, neither on the back nor in the ischial region (observations also made by both Roberts [10] and Groves [3,4] for *C. a. labiatus*) and clearly setting them apart from Inland (SM, MK) and Coast (SDB, CV) individuals. The most pronounced difference between Coast individuals and all the other populations found was their generally lighter pelage colour, the prominent and comparatively extensive yellow wash on the back and the orange ischial region. These results are very similar to those of Roberts [10] and Groves [3] who found that the coastal *C. a. erythrarchus* individuals had the most conspicuous reddish colouration in the ischial region. The overall lighter appearance of the pelage in Coast individuals found in our study could be explained by the fact that in 46% of Coast hairs analysed the light bands were equal to the length of the dark bands as opposed to Hogsback and Inland individuals where black bands were either double the length (Hogsback) or longer than light bands (Inland), resulting in an overall darker appearance. The conspicuously black arms found in both Hogsback and Inland individuals further increase the effect of this dark appearance. Roberts [10] and Groves [3] found that the main colour difference between *C. a. schwarzi* and *C. a. erythrarchus* is that the ischial region is buffy yellow (Roberts) or reddish (Groves) in *schwarzi* but not as reddish as it is in coastal *C. a. erythrarchus* (Roberts). This description compares well to the pelage colouration we observed for the Inland individuals which had a yellow rather than orange (Coast) ischial region. The Inland population was made up of individuals from Magoebaskloof and the Soutpansberg as they showed the same hair characteristics. Morphological results from our study could not be compared to Meester [8] and Grubb [2] as they do not give any other reasoning rather than distribution to delineate the two subspecies, *erythrarchus* and *labiatus*.

### Nuclear genetic diversity and population structure

Across the whole dataset of 72 individuals, the number of alleles detected per microsatellite locus ranged from two to 11, with a mean of four, and a mean expected heterozygosity ( $H_E$ ) of 0.5. The total  $H_E$  was 0.59. When the dataset was subdivided to reflect the two currently

recognised subspecies, *C. a. labiatus* and *C. a. erythrarchus*, the former showed lower levels of genetic diversity in terms of number of alleles and heterozygosity ( $H_E = 0.38$ ) and fewer deviations from HWE relative to the latter ( $H_E = 0.62$ ). The  $H_E$  value for *C. a. labiatus* falls into the lower range of what has previously been reported for Neotropical primate species [55], for example, among others, Capuchin monkeys (*Cebus apella*,  $H_E = 0.378$ ) and Squirrel monkeys ( $H_E = 0.239$ ). However, among *C. a. erythrarchus*, genetic diversity estimates are among the highest described for primate species, for example red howler monkeys (*Alouatta seniculus*,  $H_E = 0.638$ ) and brown woolly monkeys (*L. lagotricha*,  $H_E = 0.54$ ). The strong deviations from HWE observed among *C. a. erythrarchus* samples seem to be as a result of population subdivision i.e. the Wahlund effect, as inbreeding and strong genetic drift would have resulted in lower heterozygosity and numbers of alleles [55]. Also, spatially independent analyses detected some structure among samples belonging to *C. a. erythrarchus*. In terms of genetic structure and divergence between the currently recognised subspecies, we found strong evidence for isolation based on mutually exclusive private alleles in each subspecies i.e. all the individuals belonging to each taxon exhibited alleles that were not found in the other taxon. This pattern was reflected in the highly significant pairwise estimates of population differentiation (e.g.  $G_{ST} = 0.146$ ,  $P < 0.001$ ;  $\phi_{ST} = 0.55$ ), which were as high, or higher, than those reported for currently recognised primate species [55]: for example, among *Cebus* species ( $G_{ST} = 0.08$  to  $0.11$ ), *Ateles* species ( $G_{ST} = 0.105$  to  $0.123$ ) and *Alouatta* species ( $G_{ST} = 0.189$  to  $0.259$ ). The  $R_{ST}$ -value estimated during AMOVA, however, was lower between samango subspecies ( $R_{ST} = 0.15$ ,  $P < 0.001$ ) than those reported for other primate species e.g. *Cebus* species pairwise  $R_{ST}$ -values ranged from  $0.26$  to  $0.33$ .

Spatially independent analyses of the nDNA dataset consistently identified strong subdivision between the two currently recognized subspecies, and also within the *C. a. erythrarchus* samples. The finer scale structure within this taxon co-reflect the collection locality of individuals in the Soutpansberg Mountains and Magoebaskloof (inland), and coastal populations in Sodwana Bay and Cape Vidal. Bayesian clustering analysis indicates that the coastal and inland forms are distinctive based on their multi-locus genotypes, but that the Magoebaskloof population that falls geographically between the coastal and Soutpansberg populations, resembles both inland and coastal forms, indicating a possible “hybrid zone” between these two population units.

Expected heterozygosity within these *C. a. erythrarchus* sub-populations ranged from  $0.34$  in the Soutpansberg Mountains to  $0.58$  in the coastal population (Sodwana Bay and Cape Vidal). Importantly, population differentiation estimates based on the four predefined populations increased compared to analyses comparing only the two currently recognised subspecies. This, together with the results of the finer scale spatially independent analyses e.g. the assignment test (99% correct assignment) and Bayesian clustering, indicate strong structure among coastal and inland forms of *C. a. erythrarchus*, and exceptionally strong differentiation between these populations and that of *C. a. labiatus*.

## Mitochondrial population structure

Mitochondrial DNA analysis generated discordant patterns from nuclear DNA in terms of reconstructed relationships among groups. In contrast to both phenotypic descriptions and nuclear DNA, the HB and SM individuals were found to be closely related based on 16S and the SM and coastal populations (SDB and CV) were more similar based on cyt b. The coastal and MK populations were differentiated into two separate groups, and MK was consistently distinctive from all other populations. Discordance between mtDNA and patterns observed for nuclear genomes has been widely reported and has been found to be the highest in mammals [56].

Discordance may be due to inadequate data, homoplasy, nucleotide composition or hybridisation [57]. Discordant gene trees have been previously reported in Colobine monkeys where the authors attributed their findings to ancestral hybridisation [57]. In this study, discordance is most likely due to geographic isolation of HB (*C. a. labiatus*) and SM (*C. a. erythrarchus*) populations followed by secondary contact which could occur either via hybridisation, due to human mediated interventions or due to sex-biased dispersal. There is evidence that females are philopatric while males leave their troops before sexual maturity at 6–8 years old [58–60]. However, further sampling would have to be conducted in the intervening areas to properly ascertain whether clinal variation occurs between these two populations. In this case, a hybrid zone between these populations may be evident.

## Phylogeography

Paleoclimatic data of the late Quaternary (last 150,000 years) was used by Lawes [9] to explain the current distribution of samango monkeys and forests in South Africa. *Cercopithecus a. labiatus* is suggested to represent the first dispersal event from East Africa along Afromontane forests down into South Africa before the last glacial maximum (LGM) whereas *C. a. erythrarchus* is thought to have entered South Africa in a second dispersal event, after the LGM within the last 12,500 years, primarily along the expanding coastal belt forest on the Mozambique seaboard. *Cercopithecus a. labiatus* was until then in all likelihood the only representative of the species present in southern Africa prior to and during the last glacial maximum [9]. Our study shows that *labiatus* and *erythrarchus* diverged about 1.7 Mya (1.6–2.6 Mya) during the mid-Pleistocene. During this period the subtropical African climate periodically oscillated between markedly wetter and drier conditions, with step-like increases of variability and aridity near 2.8 Ma, 1.7 Ma, and 1.0 Ma and also suggesting more varied and open habitats after 1.8 Ma [61]. These wet-dry oscillations will have had a marked influence on the extent and continuity of forest habitat on the continent and DeMenocal [62] suggests that they offered discrete opportunities for ecologic fragmentation and genetic isolation. The subspecies *labiatus* was likely isolated by retreating forests during one of the drier Pleistocene periods, indicating that some forest refugia remained south of the Umfolozi system, not only throughout the LGM but also throughout the mid Pleistocene. There is also no genetic indication of secondary hybridisation with later radiations of the species which suggests that a geographic barrier of some sort, such as the Umfolozi system, as suggested by Lawes [9], will have prevented subsequent radiations dispersing further south. Our data supports the theory of separate radiation events of the species into southern Africa. However, the suggestion of only two recent radiation events suggested by Lawes [9] is possibly, considering our results, a bit too simplistic. The first radiation of what we currently refer to as *labiatus* must have, according to our results, occurred prior to the mid-Pleistocene and it thus seems more likely that several older Pleistocene radiation events occurred before the LGM. There is also no reason to assume that gallery forests of river systems (e.g. Limpopo system) connecting forests did not play a role in the radiation of the species. Our results indicate that there are genetic as well as morphological differences between the coastal and inland *erythrarchus* populations in South Africa. Future sampling of inland and coastal populations from Zimbabwe (inland) and Mozambique (coast and inland) further north will show if this might be a general pattern within the species range and also give additional insights about possible radiation routes and the number of radiation events.

## Taxonomic conclusions

Our study gives the most comprehensive insights to date into the taxonomic description of samango monkeys in South Africa. Spatially independent analyses of the nDNA dataset and

data based on pelage colour consistently identified strong subdivision between the two currently recognized subspecies, and also within the *C. a. erythrarchus* samples. The reticulate evolutionary pathway evident in the mtDNA haplotype networks, discordant from nDNA, likely represents an ancient pattern of hybridisation or introgression among samango monkey populations, and highlights the complex evolutionary history of this group. Differences in morphology and nDNA markers, however, suggest that these populations are on distinct evolutionary trajectories. Our data thus support the two recognised subspecies namely; *C. a. labiatus* and *C. a. erythrarchus*. Within *C. a. erythrarchus* two lineages should be proposed. The Soutpansberg is currently classified as *C. a. erythrarchus* by Friedmann and Daly [12] and is geographically not explicitly mentioned in any samango subspecies descriptions by either Roberts [10] or Groves [3]. However, as Magoebaskloof geographically falls under the distribution range described for *C. a. schwarzi* by Roberts and Groves and as the Soutpansberg and Magoebaskloof show the same pelage characteristics, the Soutpansberg population should more accurately be classified as *C. a. schwarzi*. Whereas, the coastal populations (CV and SDB) may be assigned to *C. a. erythrarchus* (type locality Inhambane on the coast of Mozambique). However, the taxonomy of the Magoebaskloof population remains unresolved. This population is highly divergent based on mtDNA sequence data, but is relatively uniform genetically (nDNA) and phenotypically (morphometrics and pelage colour) to the SM population (*C. a. schwarzi*). However, it is unclear whether this population can be attributed to *C. a. schwarzi* or if it is a hybrid zone or a distinct lineage. Therefore, the above hypothesis should be tested by further sampling in intermediate localities and adjacent countries to properly ascertain whether clinal variation occurs throughout the range of these taxa, or if they are isolated, as it currently unclear whether these lineages are distinct species or sub-species. Additional markers (such as X- and Y-chromosomes and mobile elements) are also required to fully resolve the taxonomy of this species and test the validity of *C. mitis* versus *C. albogularis*. The final taxonomic outcome is dependent on choice of species concept. According to the Phylogenetic Species Concept (PSC), the three lineages are diagnosable (on three independent characters) and therefore represent species. The PSC has gained widespread acceptance although its application to large mammals has been recently questioned by some conservationists and defended by systematists [63–70]. Regardless of the taxonomic outcome, our data argue for separate conservation management of the three distinct genetic entities defined by this study, as Evolutionarily Significant Units (*C. a. erythrarchus* and *C. a. labiatus*) and Management Units (coastal and inland populations of *C. a. erythrarchus*) following the definitions of Moritz [71]. Distinct genetic entities need to be conserved to protect the loss of genetic diversity, as this diversity is an essential part of biodiversity conservation.

## Supporting Information

**S1 Fig. Growth in mean body mass in male and female samango monkeys from Hogsback.** (TIFF)

**S2 Fig. Output from STRUCTURE HARVESTER showing (A) Probability (-LnPr) of K = 1 to 9 averaged over 20 runs. (B) Delta K values for real population structures of K = 1–9. Output from STRUCTURE HARVESTER showing (C) Probability (-LnPr) of K = 3 and (D) Delta K values for real population structures of K = 3.** (TIFF)

**S3 Fig. Log-likelihood plot for all 72 individuals indicating their log-likelihood of being assigned to their own subspecies, or the other (log-likelihoods on axes are negative values and**

the highest (least negative) value indicates the most likely population).  
(TIFF)

**S4 Fig. Principal coordinates analysis showing (a) the subspecific identity of samples and (b) their collection localities.**  
(TIFF)

**S5 Fig. Spatially independent correspondence analysis incorporating all 72 samples of samango monkey.** (a) and (b) are different projections of the same results. Individuals are colour-coded according to their collection locality (HB, Hogsback; SBCV, Sodwana Bay and Cape Vidal; MK, Magoebaskloof; and SM, Soutpansberg Mountains).  
(TIFF)

**S6 Fig. Individual unrooted UPGMA phylogenetic trees based on two measures of genetic distance: (a) Cavalli-Sforza and Edwards  $D_c$  (chord distance) and (b)  $D_{sw}$ .**  
(TIFF)

**S7 Fig. Maximum likelihood phylogenetic tree based on the 16S sequence data for 28 individuals.** Numbers at the nodes are bootstrap support.  
(TIFF)

**S8 Fig. Maximum likelihood tree based on cytochrome b sequence data for 29 individuals.** Numbers at the nodes are bootstrap support.  
(TIFF)

**S9 Fig. Maximum-likelihood phylogenetic tree based on 911 bp of mtDNA sequence data (16S and cyt b).** Numbers at the nodes are bootstrap support (<50% not shown).  
(TIFF)

**S10 Fig. Bayesian Phylogenetic tree based on 16S sequence data for 28 individuals.** Numbers at the nodes are Bayesian posterior probabilities.  
(TIFF)

**S11 Fig. The Bayesian 50 percent majority rule consensus tree based on cytochrome b sequence data (416 bp and 29 individuals).** Numbers at the nodes are posterior probabilities.  
(TIFF)

**S12 Fig. Bayesian phylogenetic tree based on the combined, concatenated (partitioned) mtDNA sequence data (16S and cyt b).** Numbers at the nodes are posterior probabilities.  
(TIFF)

**S13 Fig. Bayesian analysis in BEAST, showing the estimated dates of divergence of *Cercopithecus* subspecies (Genbank accession numbers are given for each sequence).**  
(TIFF)

**S1 Table. Hairs analysed from the five sampling sites detailing the sampling month, the number of males and females, the total number of individuals and the total number of hairs analysed per site with a total of five hairs analysed per individual.**  
(DOCX)

**S2 Table. Variable loadings from PCA of seven variables in 21 adult male and 36 adult female samango monkeys from four populations.**  
(DOCX)

**S3 Table. The six significant pairs of loci identified by LD analysis in GENEPOP.**  
(DOCX)

## Acknowledgments

We would like to thank Hogsback residents and Magoesbaskloof Hotel for allowing us to capture on their properties. KW would like to thank University of Kwa-zulu Natal for the capture cages. We would like to thank Kerryn Bullock, Jabu Linden, Griet Njies and Nadine Viljoen for assistance with the captures and sampling.

## Author Contributions

Conceived and designed the experiments: DLD BL KW LJNI ASWT PJT MTM AK. Performed the experiments: DLD BL KW ASWT MTM. Analyzed the data: DLD BL KW LJNI ASWT PJT MTM AK. Contributed reagents/materials/analysis tools: DLD BL ASWT PJT AK. Wrote the paper: DLD BL KW LJNI ASWT PJT MTM AK. Sample collection: BL KW ASWT.

## References

1. Skinner JD, Chimimba CT (2006) The Mammals of the Southern African Sub-region. Cambridge University Press.
2. Grubb P, Butynski TM, Oates JF, Bearder SK, Disotell TR, et al. (2003) Assessment of the Diversity of African Primates. *Int J Primatol* 24:1301–1357.
3. Groves C (2001) Primate Taxonomy. Smithsonian Institution Press, Washington DC.
4. Groves CP (2005) Order Primates. In: Mammal Species of the World. A Taxonomic and Geographic Reference, Wilson DE and Reeder DM (eds.), pp.111–184. Third Edition. Johns Hopkins University Press, Baltimore.
5. Napier PH (1981) Catalogue of Primates in the British Museum (Natural History) and Elsewhere in the British Isles. Part III: Family Cercopithecidae, Subfamily Cercopithecinae. British Museum (Natural History), London.
6. Kingdon JS (1997) The Kingdon Field Guide to African Mammals, Academic Press, London.
7. Kingdon J (2013) The Kingdon field guide to African mammals. A&C Black.
8. Meester J, Rautenbach IL, Dippenaar NL, Baker CM (1986) Classification of southern African mammals. *Transvaal Museum Monograph*, 5: 1–359. *Models. Bioinformatics* 19:1572–1574.
9. Lawes MJ (1990) The Distribution of the Samango Monkey (*Cercopithecus mitis erythrarchus* Peters, 1852 and *Cercopithecus mitis labiatus* I. Geoffroy, 1843) and Forest History in Southern Africa. *J Biogeogr* 17: 669–680.
10. Roberts A (1951) The Mammals of South Africa. Trustees of the 'Mammals of South Africa' Book Fund, Pretoria, 700 pp.
11. Dandelot P (1974) Order Primates. In: The Mammals of Africa. An Identification Manual, part 3, Meester J, Setzer H eds., Smithsonian Institution Press, Washington DC, pp. 1–43.
12. Friedmann Y, Daly B (2004) Red Data Book of the mammals of South Africa: A Conservation Assessment. CBSG Southern Africa, Conservation Breeding Specialist Group (SSC/IUCN) Endangered Wildlife Trust, South Africa.
13. Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 9: 1–9.
14. Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44: 388–396. PMID: [2916582](#)
15. Constable JJ, Packer C, Collins DA, Pusey AE (1995). Nuclear DNA from primate dung. *Nature* 373: 393. PMID: [7830789](#)
16. Kayser M, Ritter H, Bercovitch F, Mrug M, Roewer L, et al. (1996) Identification of highly polymorphic microsatellites in the rehesus macaque *Macaca mulatta* by cross-species amplification. *Mol Ecol* 5:157–159. PMID: [9147691](#)
17. Coote T, Bruford MW (1996) Human microsatellites applicable for analysis of genetic variation in apes and Old World monkeys. *J Hered* 87:406–410. PMID: [8904839](#)
18. Perelygin AA, Kammerer CM, Stowell NC, Rogers J (1996) Conservation of human chromosome 18 in baboons (*Papio hamadryas*): a linkage map of eight microsatellites. *Cytogenet Cell Genet* 75: 207–209. PMID: [9067425](#)

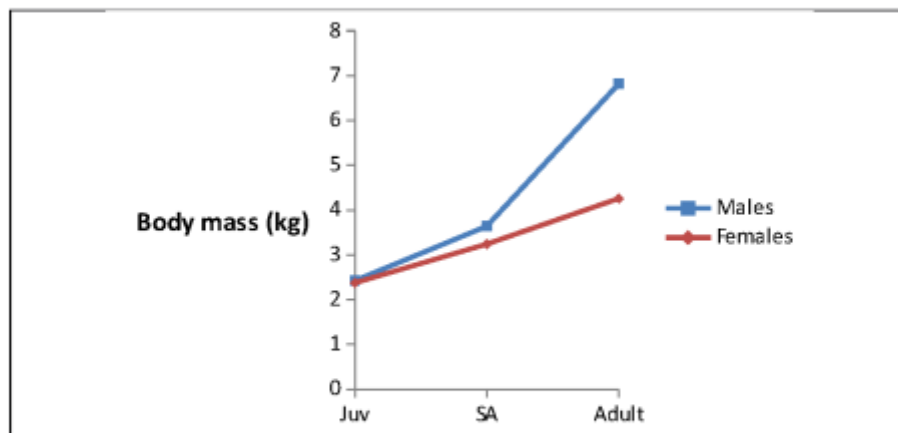
19. Morin PA, Kanthaswamy S, Smith DG (1997) Simple sequence repeat (SSR) polymorphisms for colony management and population genetics in rhesus macaques (*Macaca mulatta*). *Am J Primatol* 42: 199–213. PMID: [9209585](#)
20. Morin PA, Mahboubi P, Wedel S, Rogers J (1998) Rapid Screening and Comparison of Human Microsatellite Markers in Baboons: Allele Size Is Conserved, but Allele Number Is Not. *Genomics* 53: 12–20. PMID: [9787073](#)
21. Ely JJ, Gonzalez DL, Reeves-Daniel A, Stone WH (1998) Individual identification and paternity determination in chimpanzees (*Pan troglodytes*) using human short tandem repeat (STR) markers. *Int J Primatol* 19: 255.
22. Clifford SL, Jeffrey K, Bruford MW, Wickings EJ (1999) Identification of polymorphic microsatellite loci in the gorilla (*Gorilla gorilla gorilla*) using human primers: application to noninvasively collected hair samples. *Mol Ecol* 8: 1551. PMID: [10564463](#)
23. Bayes MK, Smith KL, Alberts SC, Altmann J, Bruford MW (2000) Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. *Conserv Genet* 1: 173–176.
24. Rogers J, Mahaney MC, Witte SM, Nair S, Newman D et al. (2000) A genetic linkage map of the baboon (*Papio hamadryas*) genome based on human microsatellite polymorphisms *Genomics*, 67: 237–247. PMID: [10936045](#)
25. Zhang YW, Morin PA, Ryder OA, Zhang YP (2001) A set of human tri- and tetra-nucleotide microsatellite loci useful for population analyses in gorillas (*Gorilla gorilla gorilla*) and orangutans (*Pongo pygmaeus*). *Conserv Genet*, 2: 391–395.
26. Grobler JP, Jacquier M, DeNys H, Blair M, Whitten PL et al. (2006) Primate sanctuaries, taxonomy and survival: a case study from South Africa. *Ecological and Environmental Anthropology* 2: 12–16.
27. Yan G, Zhang G, Fang X, Zhang Y, Li C et al. (2011) Genome sequencing and comparison of two non-human primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat Biotechnol* 29:1019–23. doi: [10.1038/nbt.1992](#) PMID: [22002653](#)
28. Hsieh HM, Chiang HL, Tsai LC, Lai SY, Huang NE et al. (2001) Cytochrome b gene for species identification of the conservation animals. *Forensic Sci Int* 122: 7–18 PMID: [11587860](#)
29. Palumbi SR, Martin AP, Romano S, McMillan WO, Stice L et al. (1991) The simple fool's guide to PCR. Spec Publ, Department of Zoology, University of Hawaii, Honolulu, Hawaii.
30. Xiong B, Kocher TD (1991) Comparison of mitochondrial DNA sequences of seven morpho-species of black flies (Diptera: Simuliidae). *Genome* 34: 306–311. PMID: [2055453](#)
31. Werle E, Schneider C, Renner M, Völker M, Fiehn W (1994) Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res* 22: 4354. PMID: [7937169](#)
32. Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.
33. Park SDE (2001) The Excel microsatellite toolkit. Website <http://animalgenomics.ucd.ie/sdepark/ms-toolkit>, accessed August 2007.
34. Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6: 288–295.
35. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*: 28: 2537–2539. PMID: [22820204](#)
36. Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Her* 86: 248–249.
37. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol ecol resour* 8: 103–106. doi: [10.1111/j.1471-8286.2007.01931.x](#) PMID: [21585727](#)
38. Excoffier L (1992) winamova ver 1.55-analysis of molecular variance-graphical windows 3 x program for the analysis of population structure from molecular or conventional genetic data. LGB/software/winamova.
39. Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959. PMID: [10835412](#)
40. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol ecol* 14: 2611–2620. PMID: [15969739](#)
41. Earl DA (2012) BM vonHoldt. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359–361.
42. Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806. PMID: [17485429](#)



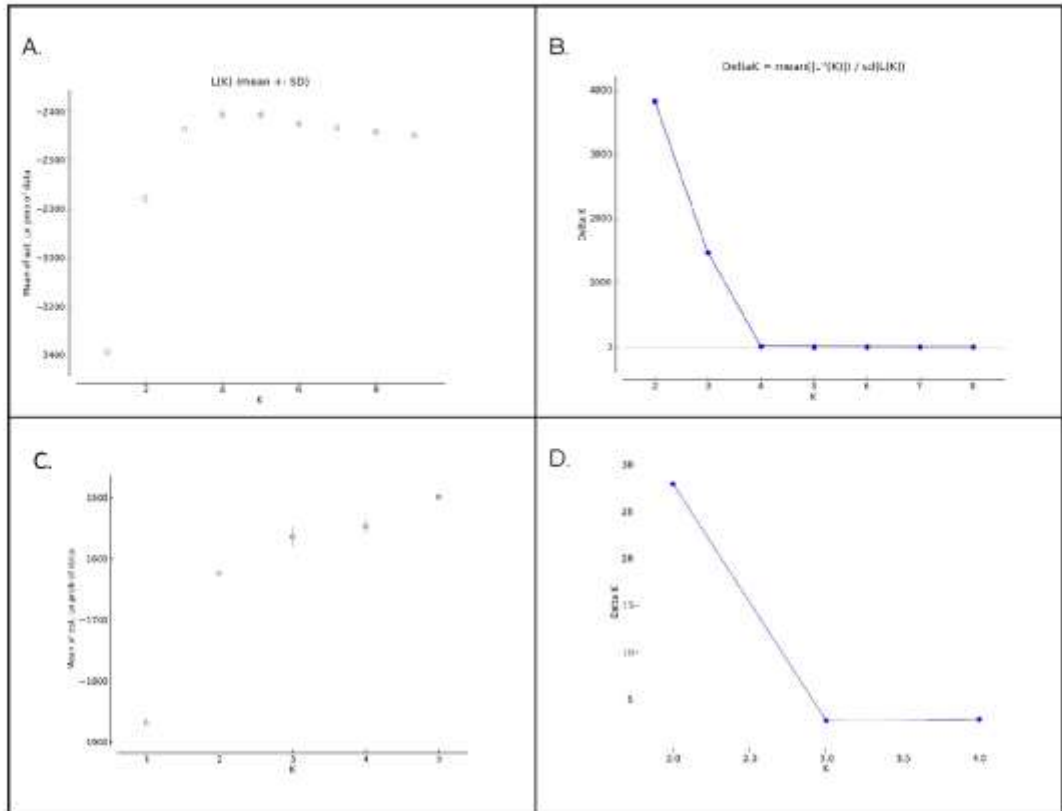
43. Rosenberg NA (2003) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4: 137–138.
44. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:95–98.
45. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–2. doi: [10.1093/bioinformatics/btp187](https://doi.org/10.1093/bioinformatics/btp187) PMID: [19346325](https://pubmed.ncbi.nlm.nih.gov/19346325/)
46. Polzin T, Daneshmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. *Oper Res Lett* 31:12–20.
47. Hudson RR, Slatkint M, Maddison WP (1992) Estimation of levels of gene flow from DNA Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol biol evol* 23: 254–267.
48. Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol biol evol* 28: 2731–2739. doi: [10.1093/molbev/msr121](https://doi.org/10.1093/molbev/msr121) PMID: [21546353](https://pubmed.ncbi.nlm.nih.gov/21546353/)
49. Posada D (2008) jModelTest: phylogenetic model averaging. *Mol biol evol* 25: 1253–1256. doi: [10.1093/molbev/msn083](https://doi.org/10.1093/molbev/msn083) PMID: [18397919](https://pubmed.ncbi.nlm.nih.gov/18397919/)
50. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–4. PMID: [12912839](https://pubmed.ncbi.nlm.nih.gov/12912839/)
51. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10: e1003537. doi: [10.1371/journal.pcbi.1003537](https://doi.org/10.1371/journal.pcbi.1003537) PMID: [24722319](https://pubmed.ncbi.nlm.nih.gov/24722319/)
52. Hart JA, Detwiler KM, Gilbert CC, Burrell AS, Fuller JL et al. (2012) Lesula: A new species of Cercopithecus monkey endemic to the Democratic Republic of Congo and implications for conservation of Congo's Central Basin. *PLoS ONE* 7: e44271. doi: [10.1371/journal.pone.0044271](https://doi.org/10.1371/journal.pone.0044271) PMID: [22984482](https://pubmed.ncbi.nlm.nih.gov/22984482/)
53. Avise JC (2004) *Molecular markers, natural history, and evolution*, 2nd ed. Sinauer, Sunderland, MA.
54. Bergmann C (1847) Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien* 3: 595–708. PMID: [25242907](https://pubmed.ncbi.nlm.nih.gov/25242907/)
55. Ruiz-Garcia M, Escobar-Armel P, Alvarez D, Mudry M, Ascunce M et al. (2007) Genetic variability in four Alouatta species measured by means of nine DNA microsatellite markers: Genetic structure and recent bottlenecks. *Folia Primatologica* 78: 73–87. PMID: [17303937](https://pubmed.ncbi.nlm.nih.gov/17303937/)
56. Toews DP, Brelford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Mol Ecol* 21: 3907–3930. doi: [10.1111/j.1365-294X.2012.05664.x](https://doi.org/10.1111/j.1365-294X.2012.05664.x) PMID: [22738314](https://pubmed.ncbi.nlm.nih.gov/22738314/)
57. Roos C, Walter L, Zinner D (2011) Natural hybridization in primates. In: Weinbauer GF, Vogel F (eds.), *Future trends in primate toxicology and biotechnology*. Waxmann, 9–33.
58. Henzi SP, Lawes MJ (1987) Breeding influxes and the behaviour of adult male samango monkeys (*Cercopithecus mitis albogularis*). *Folia Primatol* 48:125–136. PMID: [3443416](https://pubmed.ncbi.nlm.nih.gov/3443416/)
59. Henzi SP, Lawes MJ (1988). Strategic responses of male samango monkeys (*Cercopithecus mitis*) to a decline in the number of receptive females. *Int J Primatol* 9:479–495.
60. Ekernas LS, Cords M (2007) Social and environmental factors influencing natal dispersal in blue monkeys (*Cercopithecus mitis stuhlmanni*). *Anim Behav* 73: 1009–1020.
61. deMenocal PB (2004) African climate change and faunal evolution during the Plio-Pleistocene. *Earth and Planetary Science Letters* 220:3–24.
62. deMenocal PB (1995) Plio-Pleistocene African Climate. *Science* 270: 53–59. PMID: [7569951](https://pubmed.ncbi.nlm.nih.gov/7569951/)
63. Cotterill FPD, Taylor PJ, Gippoliti S, Bishop JM, Groves CP (2014) Why one century of phenetics is enough: response to 'Are there really twice as many bovid species as we thought?' *Syst Biol* 63: 819–832. doi: [10.1093/sysbio/syu003](https://doi.org/10.1093/sysbio/syu003) PMID: [24415680](https://pubmed.ncbi.nlm.nih.gov/24415680/)
64. Gippoliti S, Groves CP (2012) "Taxonomic inflation" in the historical context of mammalogy and conservation. *Hystrix, Italian Journal of Mammalogy* 23: 6–9.
65. Gippoliti S, Cotterill FPD, Groves C (2013) Mammal taxonomy without taxonomists: a reply to Zachos and Lovari. *Hystrix, Italian Journal of Mammalogy*, In Press.
66. Groves CP (2013) The nature of species: A rejoinder to Zachos et al. *Mamm biol* 78: 7–9.
67. Heller R, Frandsen P, Lorenzen ED, Siegmund HR (2013) Are there really twice as many bovid species as we thought? *Syst Biol* 62:490–493. doi: [10.1093/sysbio/syt004](https://doi.org/10.1093/sysbio/syt004) PMID: [23362112](https://pubmed.ncbi.nlm.nih.gov/23362112/)
68. Zachos FE, Lovari S (2013) Taxonomic inflation and the poverty of the Phylogenetic Species Concept—a reply to Gippoliti and Groves. *Hystrix, Italian Journal of Mammalogy*, In Press.

69. Zachos FE, Apollonio M, Bärmann EV, Festa-Bianchet M, Göhlich U, et al. (2013). Species inflation and taxonomic artefacts—A critical comment on recent trends in mammalian classification. *Mamm biol* 78: 1–6.
70. Zachos FE, Clutton-Brock TH, Festa-Bianchet M, Lovari S, Macdonald DW et al. (2013) Species splitting puts conservation at risk. *Nature* 494:35. doi: [10.1038/494035a](https://doi.org/10.1038/494035a) PMID: [23393651](https://pubmed.ncbi.nlm.nih.gov/23393651/)
71. Moritz C (1994) Defining “Evolutionarily Significant Units” for conservation. *Trends ecol evol* 9:373–375. doi: [10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4) PMID: [21236896](https://pubmed.ncbi.nlm.nih.gov/21236896/)

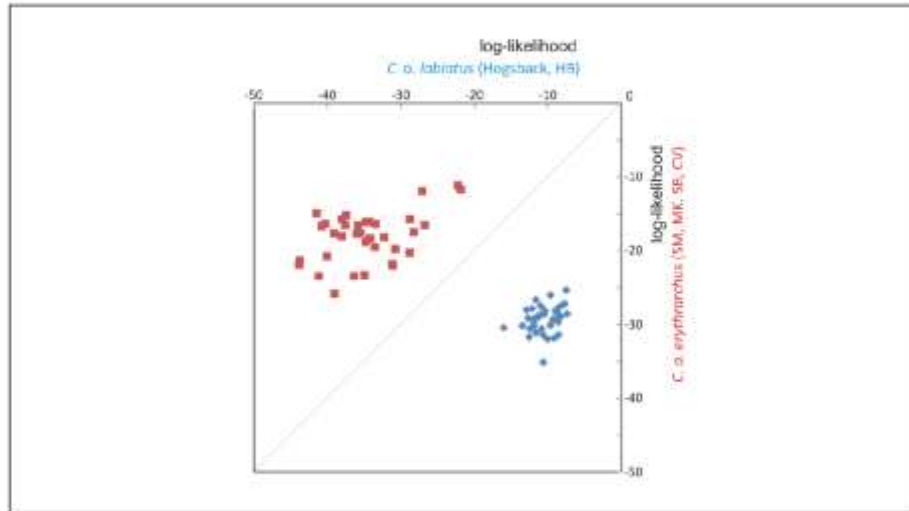
## Supporting Information



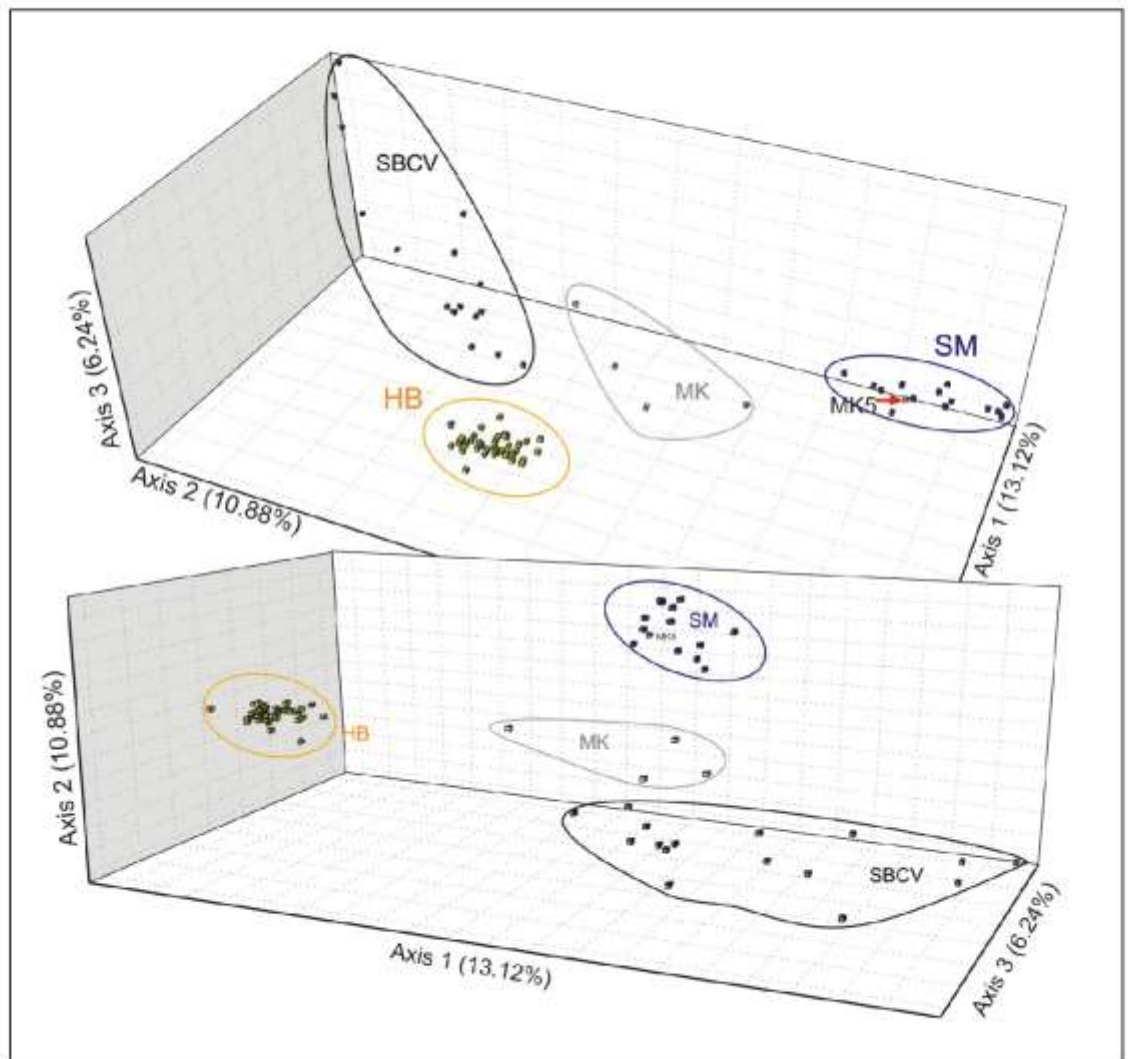
S1 Fig. Growth in mean body mass in male and female samango monkeys from Hogsback.



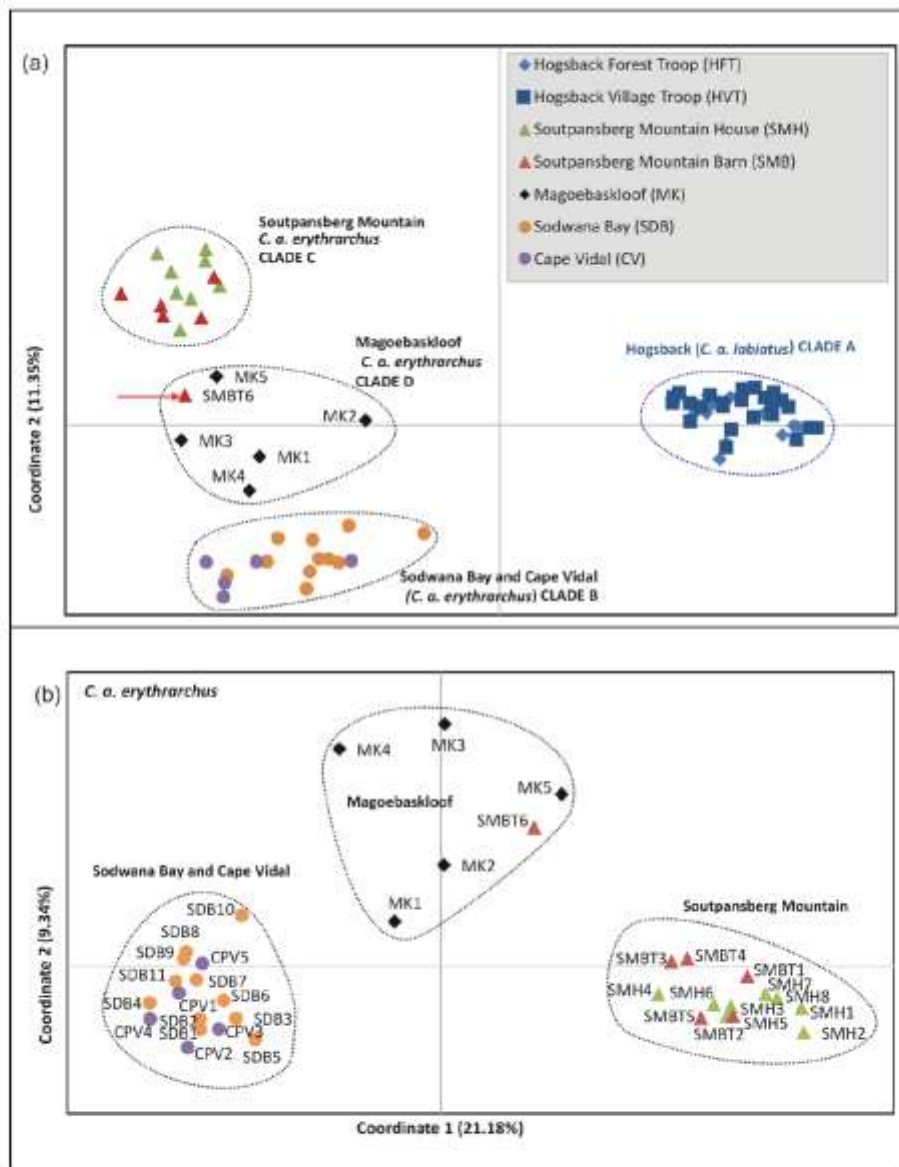
S2 Fig. Output from STRUCTURE HARVESTER showing (A) Probability (-LnPr) of K = 1 to 9 averaged over 20 runs. (B) Delta K values for real population structures of K = 1-9. Output from STRUCTURE HARVESTER showing (C) Probability (-LnPr) of K = 3 and (D) Delta K values for real population structures of K = 3.



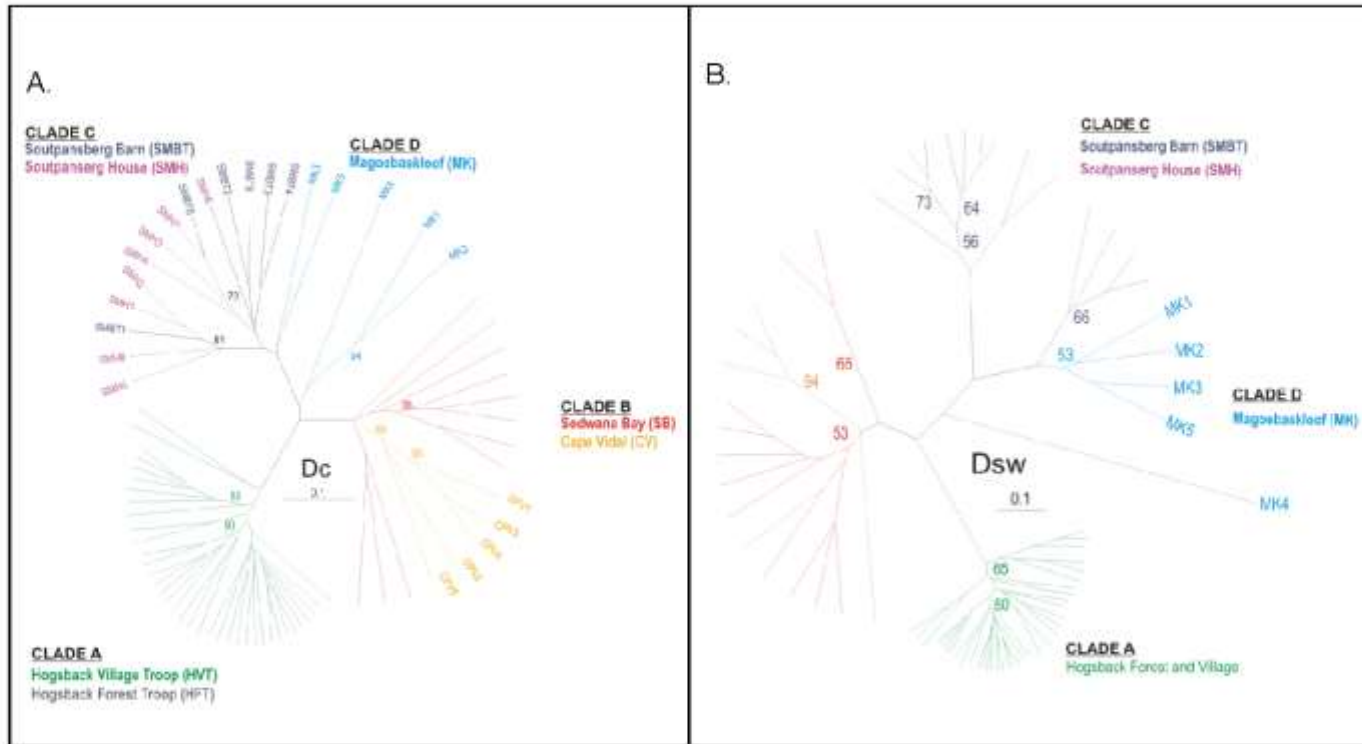
S3 Fig. Log-likelihood plot for all 72 individuals indicating their log-likelihood of being assigned to their own subspecies, or the other (log-likelihoods on axes are negative values and the highest (least negative) value indicates the most likely population).



S4 Fig. Principal coordinates analysis showing (a) the subspecific identity of samples and (b) their collection localities.

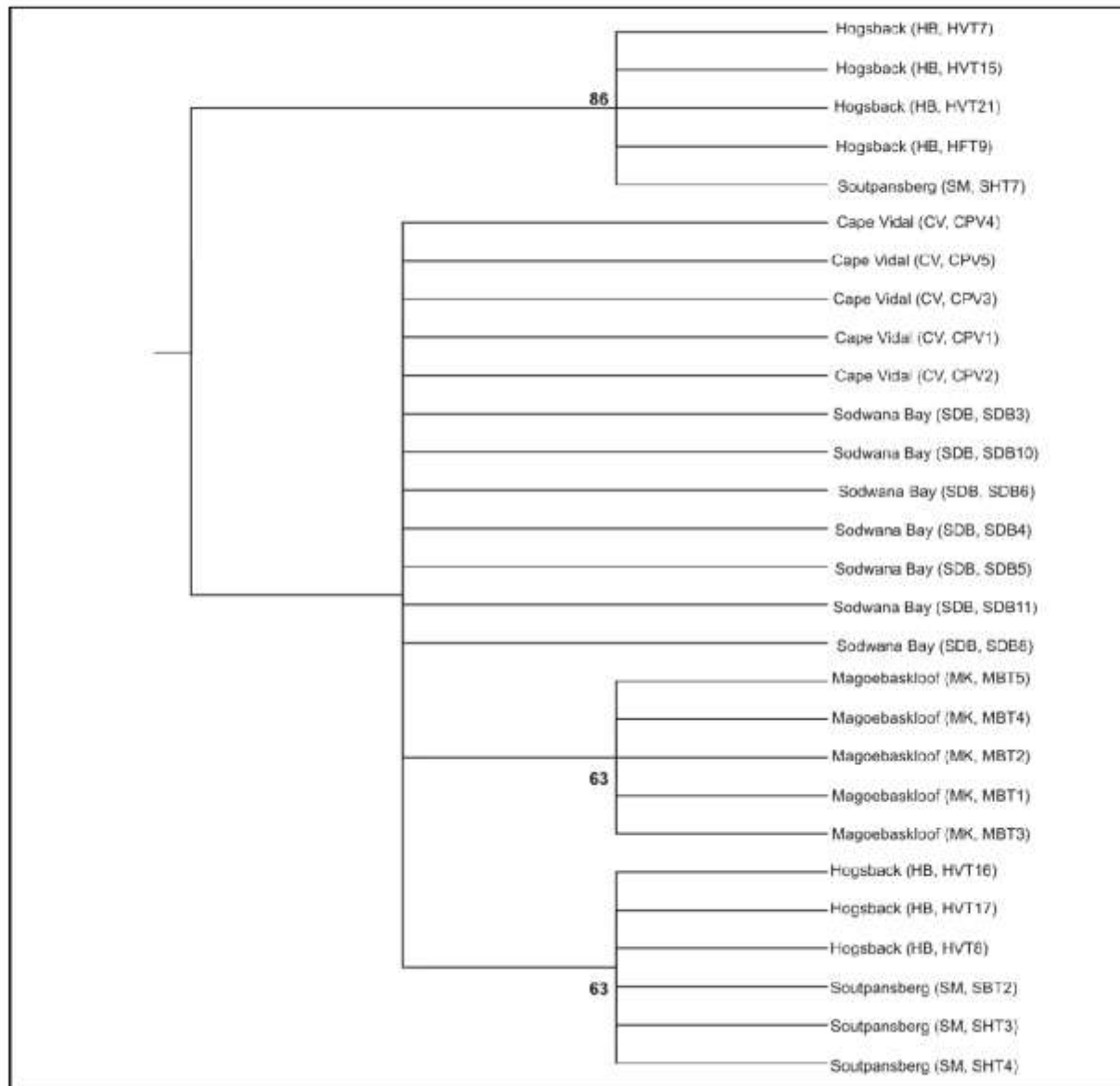


S5 Fig. Spatially independent correspondence analysis incorporating all 72 samples of samango monkey. (a) and (b) are different projections of the same results. Individuals are colour-coded according to their collection locality (HB, Hogsback; SBCV, Sodwana Bay and Cape Vidal; MK, Magoebaskloof; and SM, Soutpansberg Mountains).



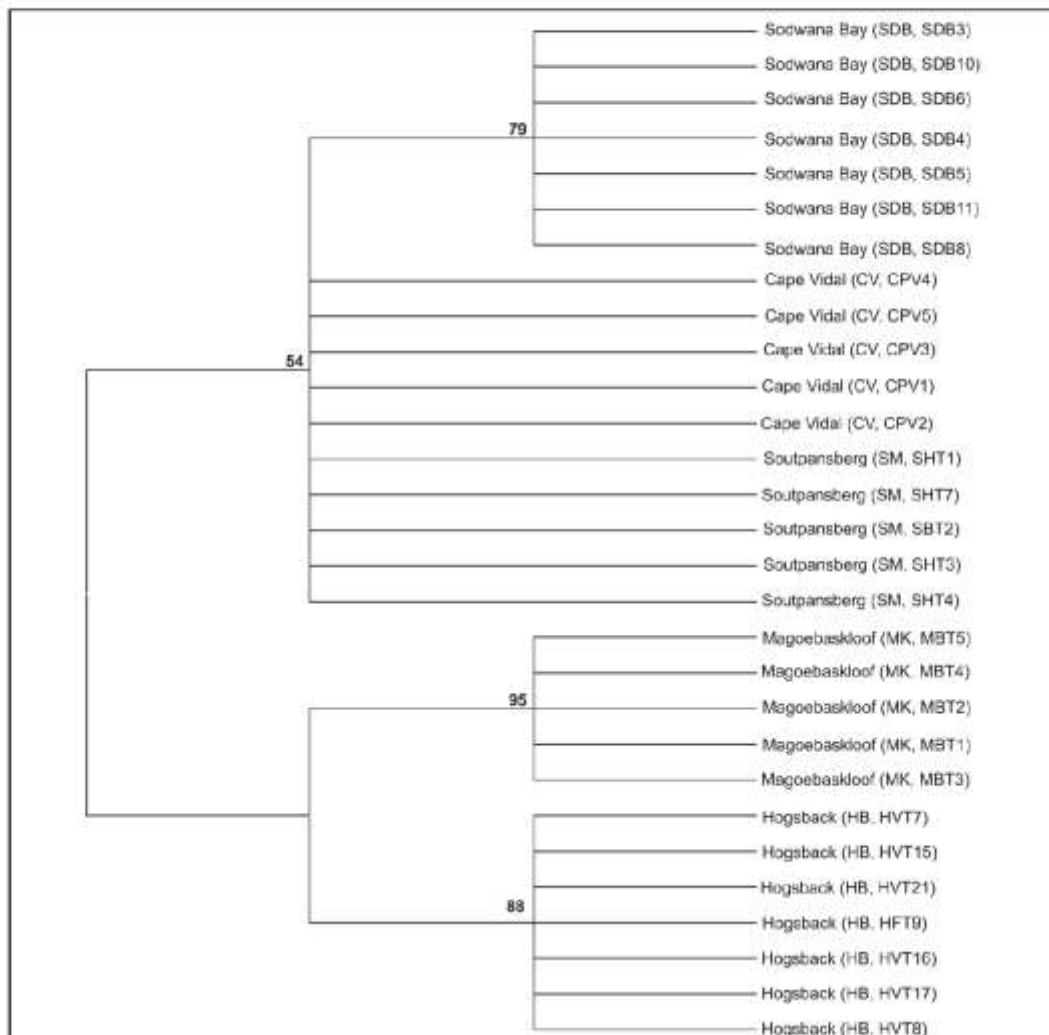
S6 Fig. Individual unrooted UPGMA phylogenetic trees based on two measures of genetic distance: (a) Cavalli-Sforza and Edwards Dc (chord distance) and (b) Dsw.



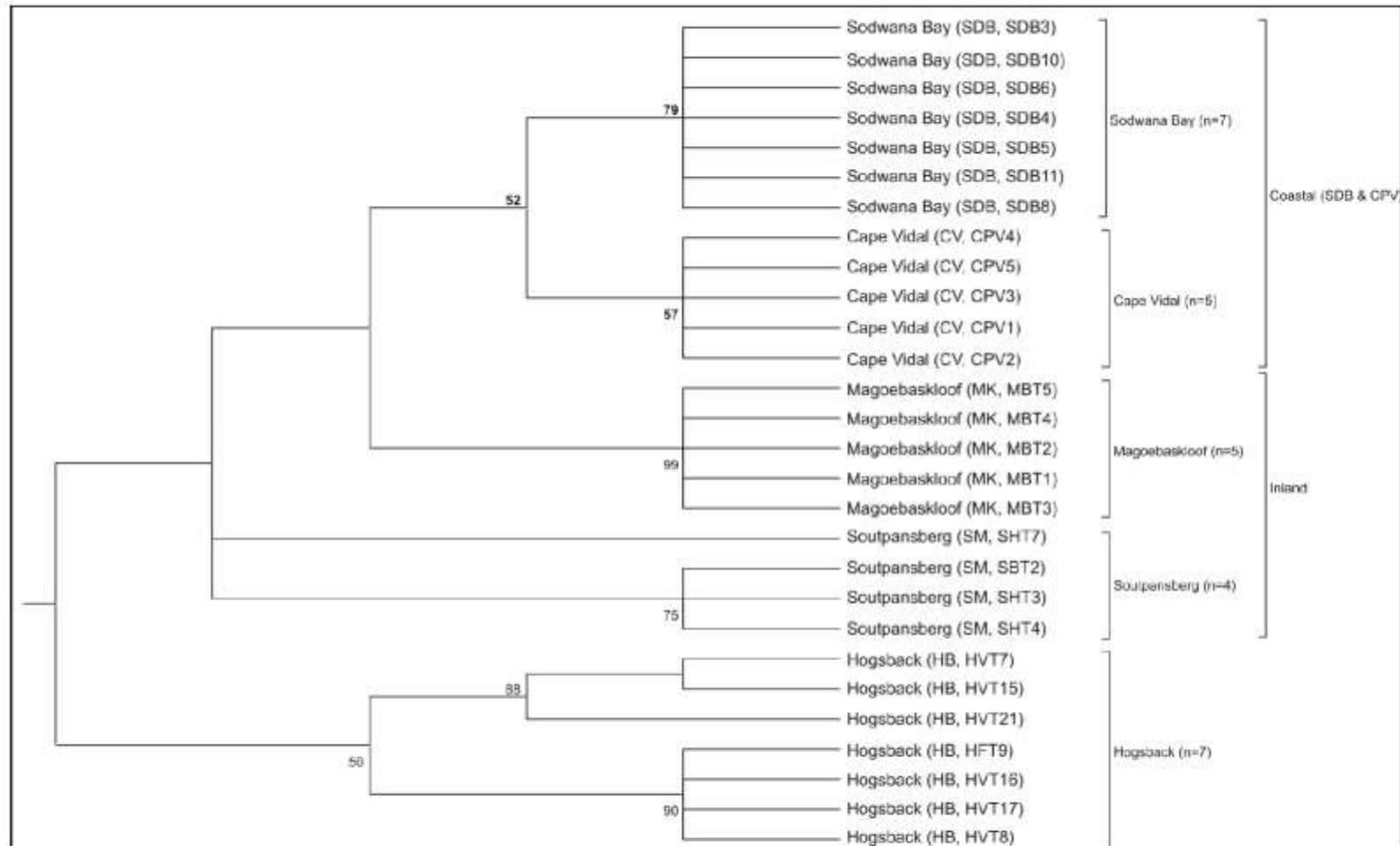


S7 Fig. Maximum likelihood phylogenetic tree based on the 16S sequence data for 28 individuals.

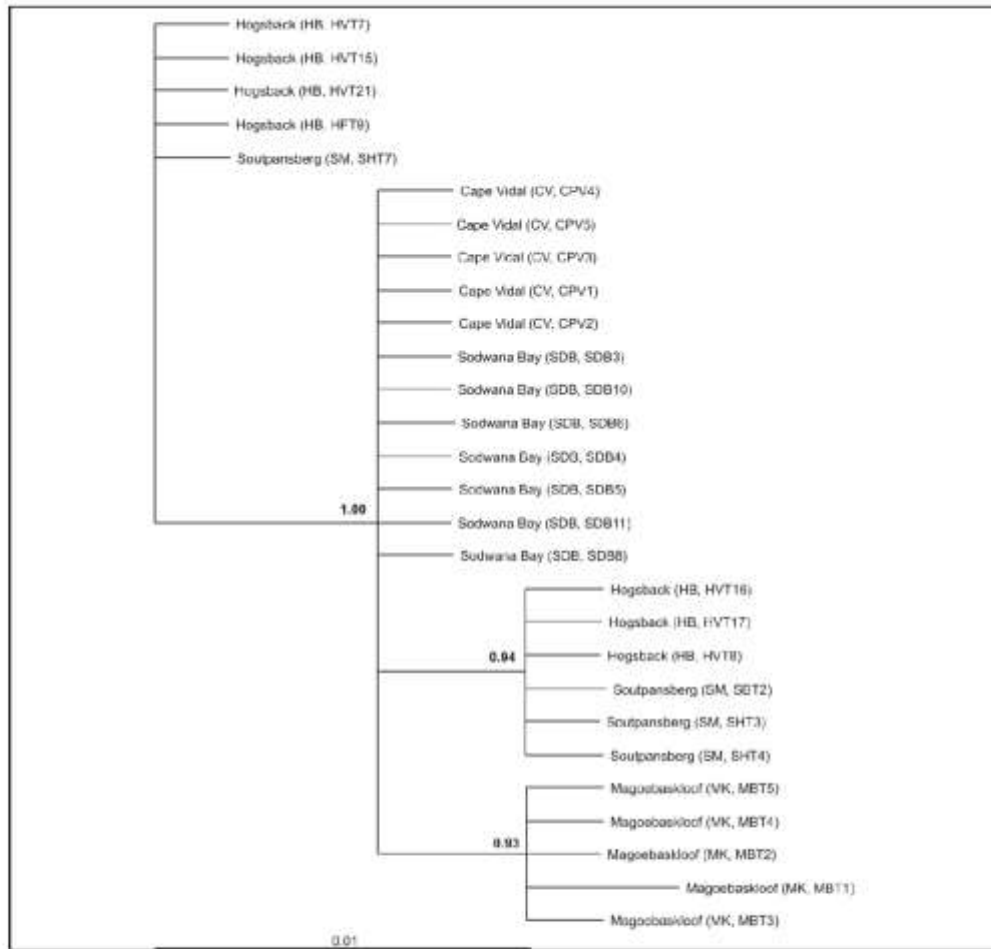
Numbers at the nodes are bootstrap support.



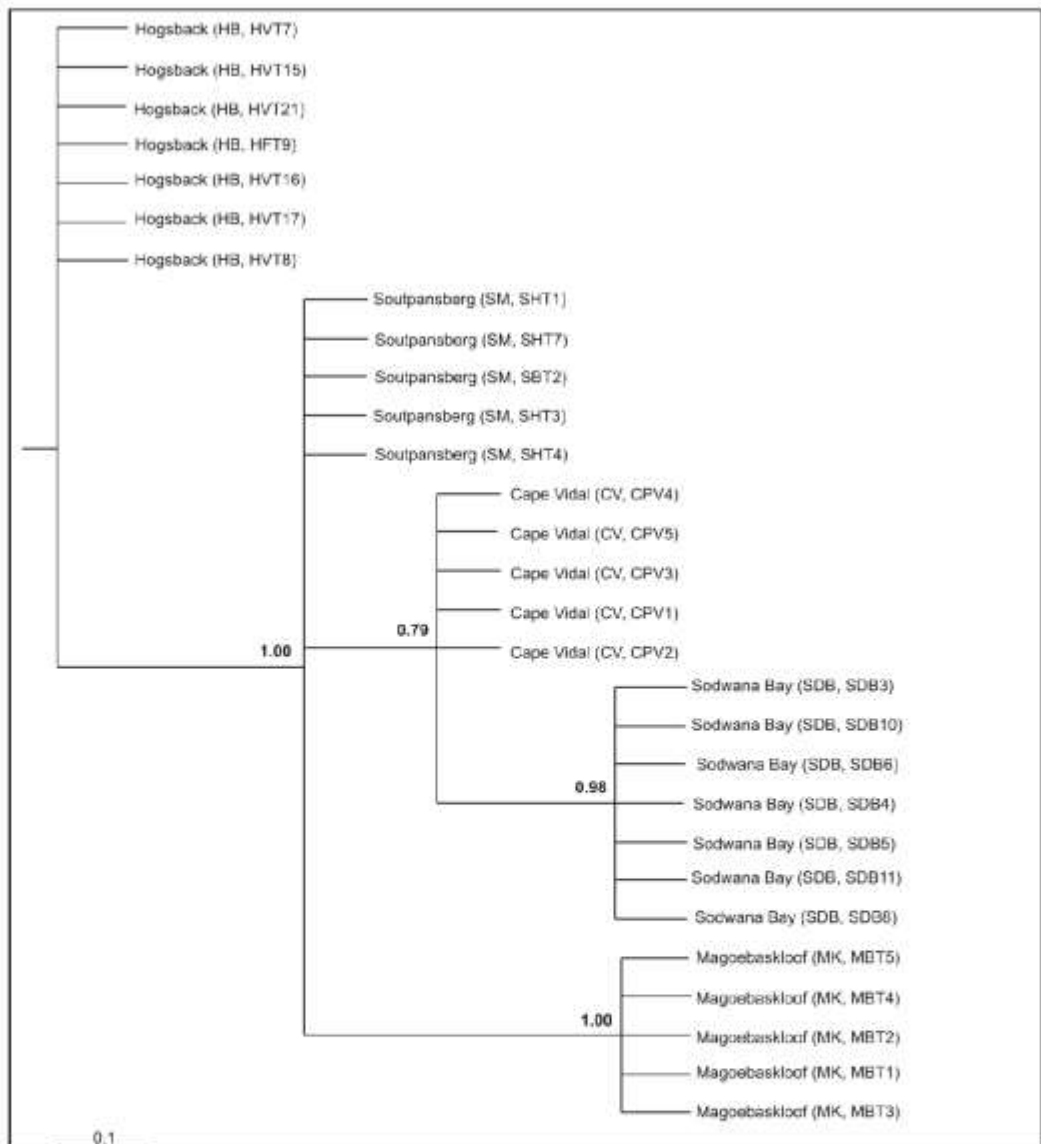
S8 Fig. Maximum likelihood tree based on cytochrome b sequence data for 29 individuals. Numbers at the nodes are bootstrap support.



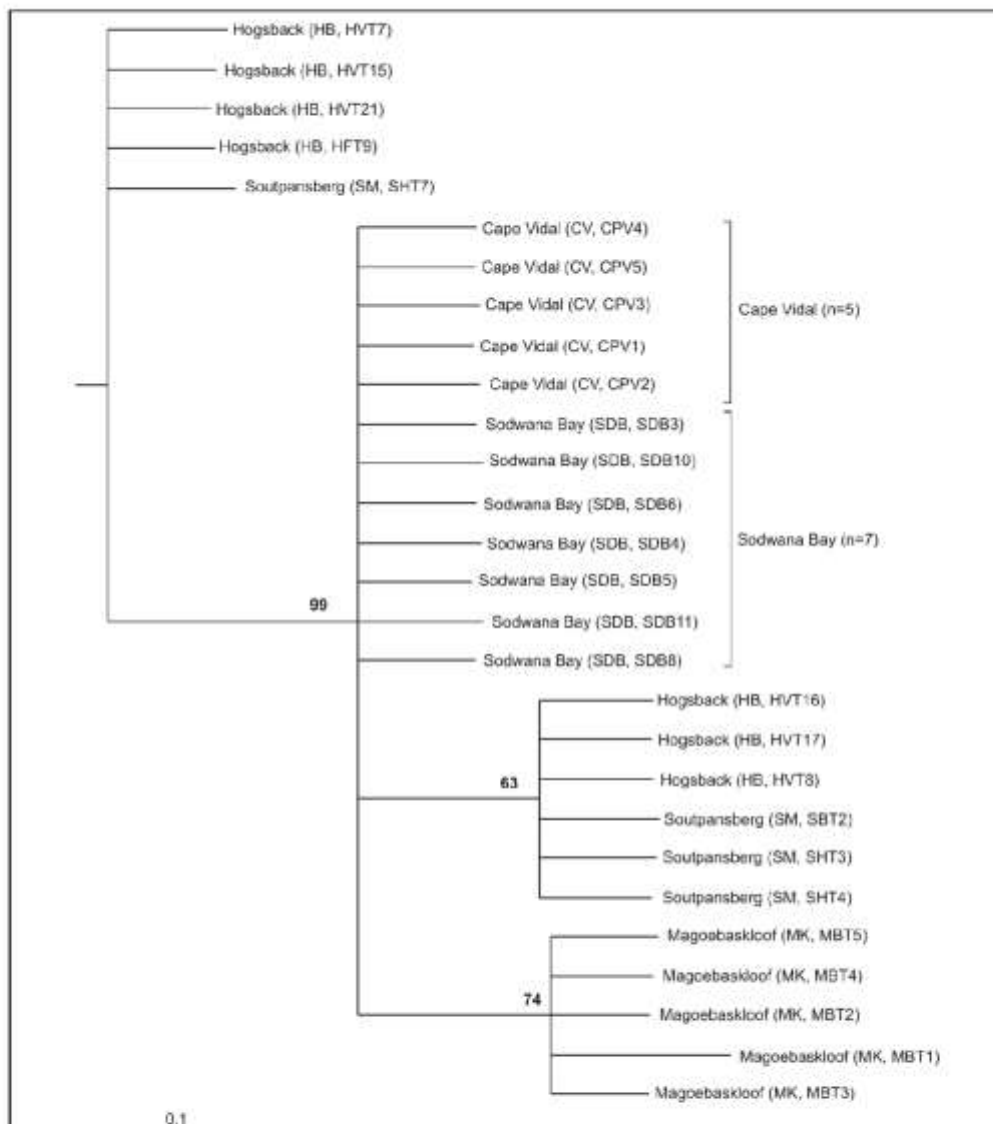
S9 Fig. Maximum-likelihood phylogenetic tree based on 911 bp of mtDNA sequence data (16S and cyt b). Numbers at the nodes are bootstrap support (<50% not shown).



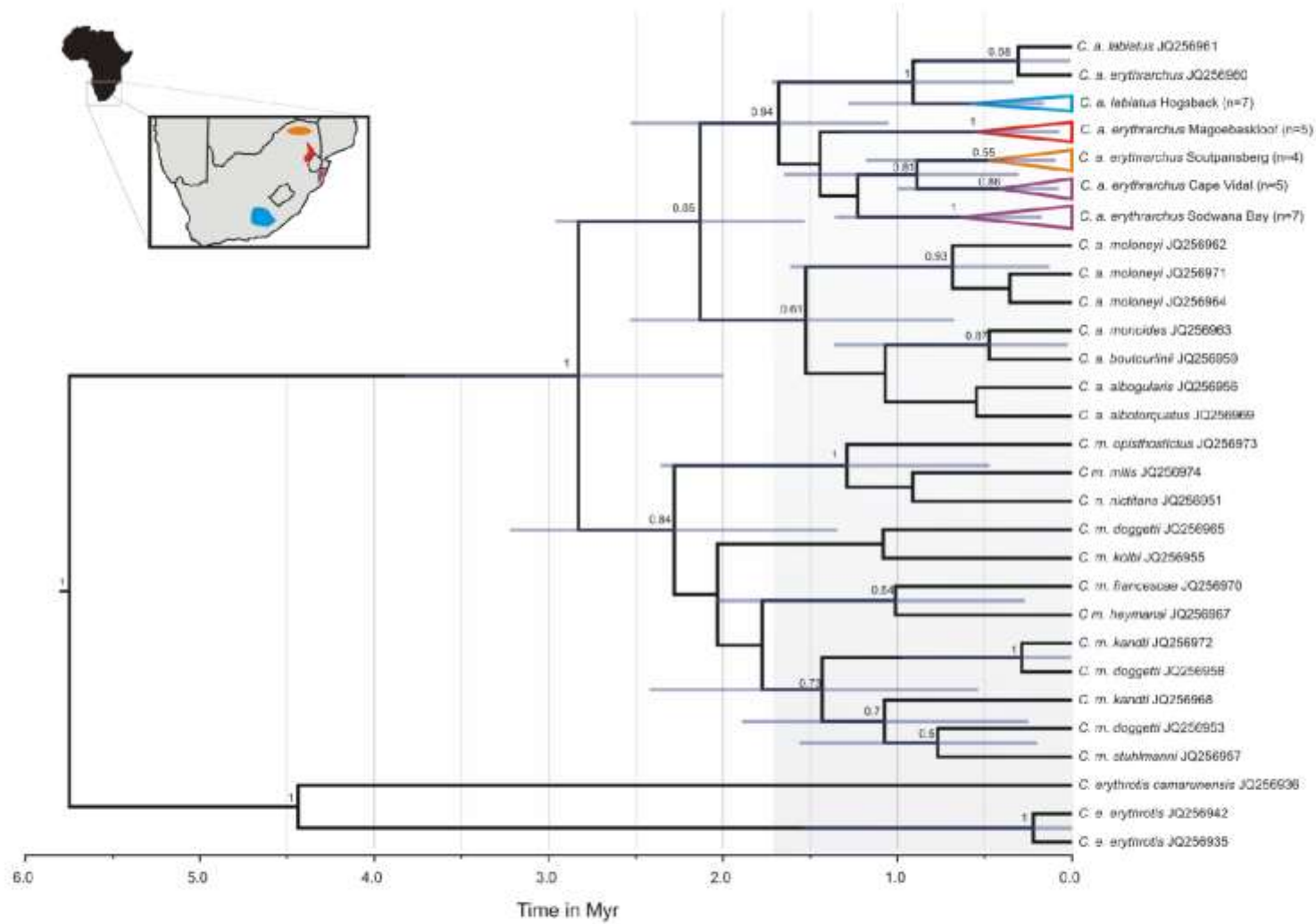
S10 Fig. Bayesian Phylogenetic tree based on 16S sequence data for 28 individuals. Numbers at the nodes are Bayesian posterior probabilities.



S11 Fig. The Bayesian 50 percent majority rule consensus tree based on cytochrome b sequence data (416 bp and 29 individuals). Numbers at the nodes are posterior probabilities.



S12 Fig. Bayesian phylogenetic tree based on the combined, concatenated (partitioned) mtDNA sequence data (16S and cyt b). Numbers at the nodes are posterior probabilities.



S13 Fig. Bayesian analysis in BEAST, showing the estimated dates of divergence of *Cercopithecus* subspecies (Genbank accession numbers are given for each sequence).

Locality	Month	Number females	Number males	Total individuals	Total hairs
Soutpansberg	August	6	2	8	40
Magoebaskloof	August	3	/	3	15
Sodwana Bay	May	3	2	5	25
Cape Vidal	May	/	3	3	15
Hogsback	March, August, September	8	7	15	75

S1 Table. Hairs analysed from the five sampling sites detailing the sampling month, the number of males and females, the total number of individuals and the total number of hairs analysed per site with a total of five hairs analysed per individual.

Male							
Variable	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6	Axis 7
<b>BM (kg)</b>	0.75	-0.05	-0.29	-0.52	0.29	0.00	0.02
<b>NC</b>	0.30	-0.65	0.22	0.46	0.18	-0.39	-0.17
<b>HB</b>	0.22	-0.20	0.49	-0.09	-0.28	0.68	-0.34
<b>TL</b>	0.41	0.51	-0.26	0.51	-0.29	-0.05	-0.39
<b>HF</b>	0.32	0.15	0.22	0.37	0.05	0.24	0.79
<b>EL</b>	0.12	-0.27	-0.12	-0.20	-0.85	-0.27	0.26
<b>EW</b>	-0.12	-0.42	-0.71	0.26	0.01	0.49	0.05
Female							
<b>BM (kg)</b>	0.64	-0.19	0.66	-0.06	0.14	-0.32	-0.02
<b>NC</b>	0.06	-0.89	-0.37	-0.08	0.24	0.01	0.06
<b>HB</b>	0.19	-0.09	0.22	0.24	-0.13	0.76	0.50
<b>TL</b>	0.70	0.27	-0.62	0.03	-0.13	-0.11	0.17
<b>HF</b>	0.23	-0.09	-0.03	0.25	-0.14	0.39	-0.84
<b>EL</b>	-0.10	-0.08	-0.01	0.93	0.01	-0.34	0.09
<b>EW</b>	-0.05	-0.28	0.06	-0.09	-0.93	-0.20	0.05

S2 Table. Variable loadings from PCA of seven variables in 21 adult male and 36 adult female samango monkeys from four populations.

Locus pair		P-Value
a	e	0.001
d	u	0.014
c	g	0.018
g	i	0.028
b	p	0.030
i	r	0.031

S3 Table. The six significant pairs of loci identified by LD analysis in GENEPOP.