

ORIGINAL ARTICLE

A novel landscape genetic approach demonstrates the effects of human disturbance on the Udzungwa red colobus monkey (*Procolobus gordonorum*)

MJ Ruiz-Lopez^{1,8}, C Barelli^{2,3,4,8}, F Rovero², K Hodges³, C Roos⁵, WE Peterman^{6,9} and N Ting^{1,7}

A comprehensive understanding of how human disturbance affects tropical forest ecosystems is critical for the mitigation of future losses in global biodiversity. Although many genetic studies of tropical forest fragmentation have been conducted to provide insight into this issue, relatively few have incorporated landscape data to explicitly test the effects of human disturbance on genetic differentiation among populations. In this study, we use a newly developed landscape genetic approach that relies on a genetic algorithm to simultaneously optimize resistance surfaces to investigate the effects of human disturbance in the Udzungwa Mountains of Tanzania, which is an important part of a universally recognized biodiversity hotspot. Our study species is the endangered Udzungwa red colobus monkey (*Procolobus gordonorum*), which is endemic to the Udzungwa Mountains and a known indicator species that thrives in large and well-protected blocks of old growth forest. Population genetic analyses identified significant population structure among Udzungwa red colobus inhabiting different forest blocks, and Bayesian cluster analyses identified hierarchical structure. Our new method for creating composite landscape resistance models found that the combination of fire density on the landscape and distance to the nearest village best explains the genetic structure observed. These results demonstrate the effects that human activities are having in an area of high global conservation priority and suggest that this ecosystem is in a precarious state. Our study also illustrates the ability of our novel landscape genetic method to detect the impacts of relatively recent landscape features on a long-lived species.

Heredity (2016) **116**, 167–176; doi:10.1038/hdy.2015.82; published online 16 September 2015

INTRODUCTION

Comprehensive understanding of the extent to which human activities are impacting today's ecosystems is essential for conservation planning and mitigation of future losses in biodiversity. This task is particularly important in tropical rainforest ecosystems, which harbor extremely high levels of species richness and endemism but are currently experiencing unprecedented degrees of anthropogenically driven habitat loss and fragmentation. Despite a well-developed theoretical framework regarding the genetic consequences of such habitat modification, including increases in genetic drift, inbreeding, population differentiation and extinction susceptibility, empirical studies of tropical rainforest taxa inhabiting fragmented landscapes have drawn mixed conclusions. Most have been unable to detect the effects of fragmentation or have shown ambiguous results (reviewed in Radespiel and Bruford, 2014). Furthermore, those that have detected the effects of fragmentation are often unable to discern whether these effects are due to fragmentation that has occurred naturally or fragmentation that has been anthropogenically driven (Radespiel and

Bruford, 2014). Landscape genetics, which uses spatial models to examine how landscape features affect the spatial distribution of genetic variation, can be used to clarify this issue (Holderegger and Wagner, 2008). Despite the potential for landscape genetics to aid in understanding the effects of forest loss and fragmentation in the tropics, such studies are still relatively scant and thus represent a current research priority for this rapidly developing field (Manel and Holderegger, 2013).

Although tremendous methodological and statistical advances have been seen since landscape genetics was formally coined in 2003 (Manel and Holderegger, 2013), parameterization and analysis of resistance surfaces remains one of the greatest challenges faced by landscape genetic practitioners. Over the years, numerous methods have been used to assign resistance values to landscape features. Recently, Peterman *et al.* (2014) used Ricker and monomolecular data transformations in combination with linear mixed effects models (maximum likelihood population effects parametrization; Van Strien *et al.*, 2012) and a nonlinear optimization algorithm to optimize

¹Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA; ²Sezione di Biodiversità Tropicale, MUSE—Museo delle Scienze, Trento, Italy; ³Reproductive Biology Unit, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany; ⁴Department of Biodiversity and Molecular Ecology, Fondazione Edmund Mach, San Michele all'Adige, Trento, Italy; ⁵Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany; ⁶Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, USA and ⁷Department of Anthropology, University of Oregon, Eugene, OR, USA
Correspondence: Dr W Peterman, School of Environment and Natural Resources, The Ohio State University, 2021 Coffey Road, 210 Kottman Hall, Columbus, OH 43210, USA.
E-mail: bill.peterman@gmail.com

or Dr N Ting, Department of Anthropology, University of Oregon, Eugene, OR 97403-1218, USA.

E-mail: nting@uoregon.edu

⁸These authors contributed equally to this work.

⁹Current address: School of Environment and Natural Resources, The Ohio State University, Columbus, OH, USA.

Received 23 January 2015; revised 30 July 2015; accepted 4 August 2015; published online 16 September 2015

resistance surfaces to describe variation in pairwise genetic distances. Importantly, this approach requires no *a priori* assumptions about the scale and direction of the resistance relationship. Peterman (2014) extended this approach to include optimization of categorical resistance surfaces as well as the simultaneous optimization of multiple resistance surfaces through the use of genetic algorithms. Using simulated data, Peterman (2014) has demonstrated that genetic algorithms are a powerful optimization framework for accurately parameterizing resistance surfaces in a landscape genetic analysis. The use of genetic algorithms for resistance surface optimization provides significant advances over existing approaches, which require *a priori* determination of the direction of the resistance relationship and assess only a limited range of potential parameter space (for example, Dudaniec *et al.*, 2013).

In this study, we leverage the framework and newly developed capacity of optimization procedures described by Peterman (2014) and Peterman *et al.* (2014) to investigate the effects of landscape resistance on gene flow in the Udzungwa red colobus monkey (*Procolobus gordonorum*), which is an endangered primate endemic to the forests of the Udzungwa Mountains in Tanzania. Specifically, we assessed (i) the effects of forest fragmentation on the genetic variation of the Udzungwa red colobus monkey, and (ii) which landscape features best explain genetic differentiation. To do so, we first characterized genetic diversity across five different forest blocks and identified population structure based only on genetic data. We then assessed how landscape features, both natural (topography, past and present forest cover) and anthropogenic in origin (fire density, distance to the nearest village, railroads), related to genetic differentiation. All of these variables have been shown to have an effect on red colobus populations (Dinesen *et al.*, 2001). To combine the resistance information in these variables, our new approach simultaneously optimizes multiple surface types (for example, both categorical and continuous) to create composite resistance models as implemented in the R package 'ResistanceGA'. Given that red colobus monkeys are long-lived arboreal forest-adapted species, we expected that historical forest cover would be the most important predictor explaining genetic differentiation across the landscape.

This research is of great importance because the Udzungwa red colobus is an indicator species and the Udzungwa Mountains are part of a major biodiversity hotspot facing increasing levels of forest loss and fragmentation. No wildlife study has yet quantified how this habitat alteration has affected the levels of genetic diversity and gene flow throughout this ecosystem or whether it is natural and/or human-related landscape features that affect genetic differentiation. Thus this study will help not only to validate a new method for elucidating the causes of genetic differentiation in a tropical-forest adapted species but will also provide insight into the overall ecosystem health of a biodiversity hotspot through the study of an important indicator species.

MATERIALS AND METHODS

Study system—site and species

The Udzungwas occupy an area of approximately 19 000 km² (7°40' S to 8°40' S and 35°10' E to 36°50' E; see Figure 1) and comprise the largest forests (±1300 km²) remaining in the Eastern Arc Mountain chain (Rovero *et al.*, 2014). They are thus of critical importance to the Eastern Afromontane Biodiversity Hotspot. Although the forests within the Udzungwas were most likely connected to one another at some point in the past, they are currently fragmented into numerous blocks owing to a combination of natural factors (for example, geology, climate, terrain morphology, aspect) and human disturbance (for example, subsistence and commercial logging, pole cutting,

agriculture, bushfires) (Dinesen *et al.*, 2001; Struhsaker *et al.*, 2004; Marshall *et al.*, 2010; Rovero *et al.*, 2012). These blocks vary widely in size, shape, altitudinal range, type of forest and protection level (Figure 1; Supplementary Table S1). The five forests sampled in this study (with forest area in parentheses from Marshall *et al.*, 2010) are Mwanihana (151 km²), Matundu (526 km²), Ndundulu (231 km²), Udzungwa Scarp (here after referred to as U. Scarp; 314 km²) and Magombera (12 km²). It is not known how long these blocks have been isolated from one another, but some have likely been separated for >100 years with anthropogenic activity maintaining the separation (Struhsaker *et al.*, 2004). Further information on altitudinal range and protection status is displayed in Supplementary Table S1.

Red colobus monkeys are arboreal, forest adapted, leaf-eating monkeys that are considered important indicator species in the African tropics (Struhsaker *et al.*, 2004). They are sensitive to habitat degradation as they have a specialized diet; they are also extremely vulnerable to hunting as they live in large social groups (up to >60–70 individuals) (Marshall *et al.*, 2010, Rovero *et al.*, 2012). They have a long generation time (6–8 years), making recoveries from population declines relatively slow. Additionally, as forest-adapted species, red colobus monkeys have a limited ability to disperse through non-forested or matrix habitats. Although the Udzungwa red colobus is locally common and distributed throughout the majority of the Udzungwa forest blocks with a total estimated abundance of 60–65 000 individuals (Araldi *et al.*, 2014), the species is IUCN-Endangered as they are endemic to the Udzungwas and fragmented into several sub-populations, some of which are decreasing in size (Struhsaker *et al.*, 2008). In fact, ecological and demographic studies have demonstrated that this monkey is rare or absent in forests where they are hunted, and it primarily thrives in well-protected blocks of old growth forest (Struhsaker *et al.*, 2004; Marshall *et al.*, 2010; Rovero *et al.*, 2012; Araldi *et al.*, 2014).

Sampling and genotyping

We non-invasively collected fecal samples of red colobus monkey from 2011 to 2012 following the two-step ethanol–silica storage method (Nsubuga *et al.*, 2004). We conducted reconnaissance walks along line transects that were 2 km in length, systematically distributed 1 km apart and covering most of the forest extent (see details in Araldi *et al.*, 2014). Using disposable gloves, we collected fecal samples directly after an individual had defecated whenever red colobus groups were detected along or near transects. To avoid potential re-sampling, we sampled each site only once and simultaneously flagged potential samples on the ground before collection. Sampling locations are displayed in Figure 1 (see also Figure 4 and Supplementary Figures S1–S3). We conducted the study without direct contact or interaction with the animals and under permission of the Tanzania Commission for Science and Technology (COSTECH; Permit Nos. 2011-85-NA-2011-33; 2011-84-NA-2011-33; 2011-351-NA-2011-68; 2011-346-NA-2011-183), Tanzania Wildlife Research Institute (TAWIRI) and Tanzania National Parks (TANAPA).

We extracted total genomic DNA from 170 fecal samples using the Gen-IAL First-DNA-All-Tissue-Kit (Gen-IAL, Troisdorf, Germany) following the manufacturer's protocol. To determine the sex of the individuals, we used a PCR-based gonosomal sexing assay (Roos, unpublished) and carried out micro-satellite genotyping using the following 11 polymorphic tetranucleotide markers: D1S548, D4S243, D6S493, D12S67, D13S321, D3S1766, c19a, D11S2002, D14 S306, D6S311, and D5S1457 (see Supplementary Material S1 and Supplementary Table S2 for details). Details regarding PCR reactions and cycling parameters are available in Supplementary Material S1. PCR products were run together with a size standard (Gene Scan TM 400 HD Rox, Applied Biosystems, Foster City, CA, USA) on an ABI 3130xL Genetic Analyzer (Applied Biosystems). Allele sizes were determined by using the Peak Scanner Software (Applied Biosystem, version 1.0). To assure accuracy, genotyping was repeated several times leading to a consensus genotype. Accordingly, individuals were called heterozygotes at a locus when two different alleles appeared together in at least two replicates, and homozygotes were called when only one allele appeared in at least five replicates. Based on the different runs, we calculated the allelic drop out rate, false allele rate and quality indices per locus across samples (see Supplementary Material S1 and Supplementary Table S2).

Final genotypes were compared in the program GENALEX 6 (Peakall and Smouse, 2006), and all samples that differed by four or fewer alleles were

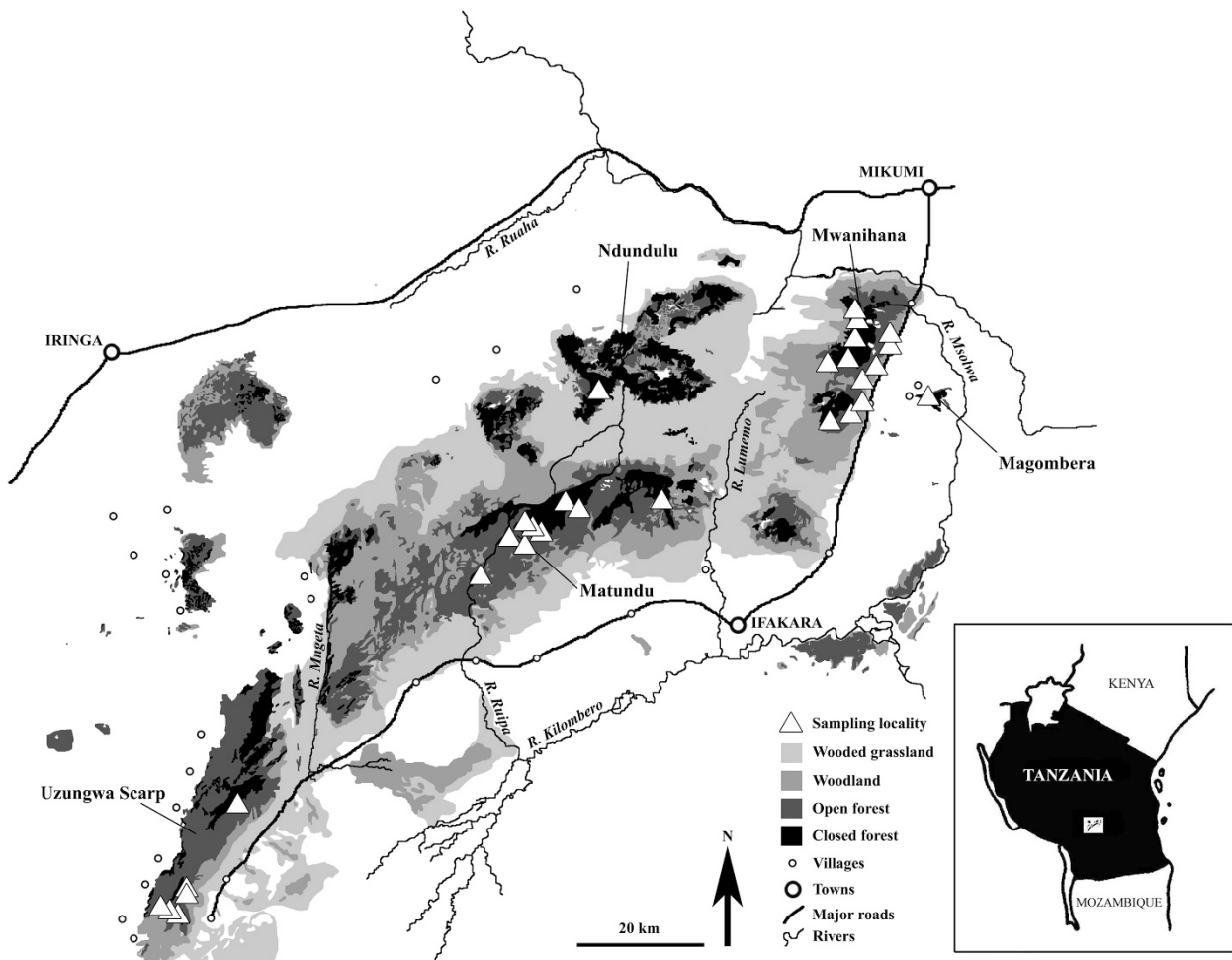


Figure 1 Map of the Udzungwa Mountains showing the five sampled forests (Uzungwa Scarp, Matundu, Ndundulu, Magombera, Mwanihana) and sampling localities. Modified from Marshall *et al.*, 2010.

identified and rechecked for potential problems with allelic dropout. Those samples that differed at ≤ 2 alleles but matched in sex and were collected at the same point were considered the same individual. We used this conservative approach to avoid scoring erroneous genotypes as different individuals. Because red colobus monkeys live in groups with related individuals, when working with non-invasive samples it is important to accurately identify different individuals. We thus tested the power of these markers to distinguish individuals in a subset of 15 individuals per forest block by computing $P(\text{ID})_{\text{random}}$, which is the power to differentiate between randomly chosen individuals, and $P(\text{ID})_{\text{sib}}$, which is the power to differentiate between siblings and is a more conservative measure (Waits *et al.*, 2001). We used the results of the more conservative test and the molecular sexing assay to ensure that we were not resampling the same individuals (Waits *et al.*, 2001).

Population genetic analyses

We tested for deviation from Hardy–Weinberg equilibrium and linkage disequilibrium in each forest block and across the five forest blocks using the Fisher’s exact test and Markov chain algorithms (1000 batches, 10 000 iterations) implemented in GENEPop version 4.0.10 (Rousset, 2008). Significance levels were adjusted for multiple comparisons using a Bonferroni correction (Rice, 1989). To test for evidence of null alleles, we used the program MICROCHECKER (Van Oosterhout *et al.*, 2004). To assess the levels of genetic diversity, we calculated the mean unbiased expected and observed heterozygosity using GENALEX 6 (Peakall and Smouse, 2006) for each forest block. Allelic diversity was analyzed by calculating mean allelic richness and the number of private alleles per forest block. These estimates were adjusted for

sample size using rarefaction in HP-RARE 1.0 (Kalinowski, 2004). Pairwise F_{ST} and population differentiation were evaluated at different levels. We calculated pairwise F_{ST} values across blocks using Fstat v. 2.9.3.2 (Goudet, 1995; 10 000 permutations) and tested for significant differences across them. We corrected for multiple testing using sequential Bonferroni correction (Rice, 1989). To infer population structure, we used both a Bayesian clustering method implemented in the program STRUCTURE v. 2.3.4 (Pritchard *et al.*, 2000) using ΔK to identify the number of clusters (Evanno *et al.*, 2005) and a Discriminant Analysis of Principal Components (DAPC; Jombart *et al.*, 2010). Both programs were run in all the individuals and in a subset of individuals that did not include full sibling and parent offspring associations identified through MLrelate (Kalinowski *et al.*, 2006) to ensure that highly related individuals did not influence the results. To infer recent migration (over the past few generations) across the five forest blocks, we used BayesAss 3.0 (Wilson and Rannala, 2003). Further information regarding the detailed population structure and migration analyses are in Supplementary Material S2.

Landscape genetic analyses

We estimated the genetic dissimilarity among the individuals in our sample by calculating the Rousset’s genetic distance a_i (Rousset, 2000) between pairs of individuals using the program SPAGEDI (Hardy and Vekemans, 2002). Because many sample locations contained multiple individuals belonging to a single social group, we used only a single observation per GPS (Global Positioning System) sample location to reduce redundancy. For these locations, we took the harmonic mean of the Rousset’s a ($n = 22$ locations, average of 2.03 ± 1.04 s.d. individuals per location). Before doing this, we also removed full sibling and

parent offspring associations identified through MLrelate (Kalinowski *et al.*, 2006) so that highly related individuals did not influence the means. The final number of unique spatial sample locations was 36.

To determine the effects of topography, land cover and human disturbance on gene flow in red colobus monkeys, we developed and assessed eight resistance surfaces for their relationship with pairwise genetic distance. Categorical surfaces included contemporary forest cover (open canopy, closed canopy, no forest; derived from Landsat ETM1; Global Land Cover Facility/ US Geological Survey; 25 October and 1 November 1999; Paths 167–8; Rows 65–6 following Marshall *et al.* (2010); Supplementary Figure S1), historical land cover (from 1981 to 1994; <http://glcf.umd.edu/data/landcover/> (Hansen *et al.*, 1998); Supplementary Figure S2) and railroad (Supplementary Figure S2). Continuous surfaces included topographic position, topographic ruggedness, distance to the nearest village (Supplementary Figure S3) and fire density (Supplementary Figure S3). Railroad and distance to the nearest village layers are available at <http://easternarc.or.tz/downloads/GIS-data/Udzungwa.zip>. Topographic position index (TPI) and topographic ruggedness were derived from a 90-m resolution digital elevation model (<http://hydrosheds.cr.usgs.gov/hydro.php>) using the Topography Tools toolkit (Dilts 2010). TPI is a relative measure that describes the position (for example, ridge, slope, valley) of a 90-m cell on the landscape relative to the surrounding landscape, while ruggedness describes the amount of elevational change within a focal area. Using the locations of fires on the landscape from 2000 to 2007 (Supplementary Figure S3A), we calculated the density of fires within a 5-km moving window to estimate the number of fires per square kilometer on the landscape, and using the locations of villages in our study area (Supplementary Figure S3B), we calculated distance to the nearest village as the Euclidean distance of every 90-m cell to the nearest village. All surfaces were then resampled to a resolution of 500 m for analysis. Geographic Information Systems data processing was done with ArcGIS 9.3 (ESRI, Redlands, CA, USA).

We measured the resistance distance between sample locations using CIRCUITSCAPE v.4.0.5 (McRae, 2006). This approach assesses all possible pathways between any two points and may better represent gene flow that occurs over multiple generations (McRae, 2006). For this analysis, we assessed connectivity based on average resistances using an eight neighbor connection scheme. To determine resistance values of each continuous surface, we used methods described by Peterman *et al.*, (2014), which uses monomolecular and Ricker functions to transform continuous resistance surfaces. This approach extensively explores parameter space to determine the optimal resistance values for each surface that best describe the pairwise genetic differentiation (a_r) while making no *a priori* assumptions concerning the direction or magnitude of the relationship. Extending the approach by Peterman *et al.*, (2014), we also optimized the resistance value of land cover types in each categorical resistance surface. To accomplish this, we held one cover type constant at a resistance value of 1 and then adjusted the resistance values of all other levels, exploring values ranging from 0 to 3500. Because samples of individuals were not evenly distributed across the landscape, we conducted bootstrap resampling of our data to control for potential bias in our results. Twenty-seven individuals (that is, 75%) were randomly selected without replacement and each surface was then fit to the sampled individuals. Following 10 000 iterations we determined, the average rank, the average model weight ($\bar{\omega}$), and the frequency a surface

was the top ranked model ($\hat{\pi}$). Burnham and Anderson (2002) describe $\hat{\pi}$ as being the bootstrap equivalent to ω , providing an estimate of the uncertainty in the top model.

New to this study, following optimization and bootstrapping of surfaces in isolation we simultaneously optimized all relevant combinations of the top resistance surfaces to generate unique composite surfaces. All surfaces that had greater selection frequency $\hat{\pi}$ than distance alone were included in the composite surface optimization. We again conducted bootstrap model selection using 10 000 random samples to determine the average rank, $\bar{\omega}$, and $\hat{\pi}$ of each composite model and component surface. Finally, we calculated the cumulative $\bar{\omega}$ of each surface to determine its relative importance across the models considered (Burnham and Anderson, 2002).

Our novel framework for optimization and selection of resistance surfaces used the R package ResistanceGA (Peterman, 2014; <https://github.com/wpeterman/ResistanceGA>), which relies on a genetic algorithm (GA; Scrucca, 2013) to adaptively search parameter space and find the resistance surface transformation that best explains pairwise genetic differentiation. We used AICc (Akaike's information criterion corrected for small/finite sample size; Akaike, 1974) as our objective criteria during optimization. Mixed effects models were fit by maximum likelihood in *lme4* (Bates *et al.*, 2014), using the maximum likelihood population effects parameterization (Van Strien *et al.*, 2012). Details of our new approach can be found in Peterman (2014), and optimization methods and genetic algorithm settings can be found in Supplementary Materials S3.

Using the best supported resistance surface, we visualized the current flow among the sampled forest fragments. In addition, using the same resistance surfaces and based on bootstrap model selection, we visualized current flow on the landscape following methods described by Koen *et al.*, (2014), buffering our focal landscape by 33–43% and assessing current flow based on 100 nodes evenly distributed around the perimeter of the buffered landscape. This method allows for an unbiased visual interpretation of current flow across the resistance landscape, with the density of current in each grid cell representing the probability of use by a dispersing animal (Koen *et al.*, 2014).

RESULTS

Population genetics

The probability of identity analyses indicated that at least nine loci are needed to accurately identify individuals in all forest blocks ($P(\text{ID}) < 1.0 \times 10^{-6}$ and $P(\text{ID})_{\text{sib}}$ is < 0.002 in all five forests). After removing all the individuals genotyped at < 9 loci, the final sample size was 121 individuals (Table 1). In addition, locus D14S306 was not considered for downstream analyses given the high proportion of missing values (43%), leaving a final panel of 10 microsatellite loci.

No population or locus deviated significantly from expected genotype frequencies under Hardy–Weinberg equilibrium after Bonferroni corrections. No loci showed evidence of linkage disequilibrium. Two loci, c19 for U. Scarp and D12S67 for Mwanihana, showed evidence of null alleles. Locus D12S67 had an estimated frequency of null alleles of 0.15 (Oosterhout measure) in Mwanihana

Table 1 Genetic diversity measures for the five forest blocks included in the analyses

Pop	N	Na	Ne	Ac	Pc	Ho	He	UHe	F_{IS}
U. Scarp	16	5.2	2.993	4.98	0.84	0.633	0.631	0.652	0.030
Matundu	21	5.0	3.230	4.43	0.12	0.683	0.657	0.673	–0.015
Ndundulu	23	4.9	3.164	4.41	0.15	0.709	0.658	0.673	–0.055
Mwanihana	27	4.8	3.320	4.24	0.11	0.684	0.675	0.688	0.006
Magombera	34	4.8	2.911	4.06	0.08	0.661	0.614	0.623	–0.061
Total mean	121	4.9	3.124	4.42	0.26	0.674	0.647	0.662	–0.024
σ		1.1	0.928	0.31	0.29	0.156	0.119	0.122	0.034

Abbreviations: Ac, allelic richness corrected by sample size; He, expected heterozygosity; Ho, observed heterozygosity; N, sample size per fragment; Na, mean number of alleles; Ne, mean number of effective alleles (number of alleles with frequency $> 5\%$); Pc, proportion of private alleles corrected by sample size; UHe, unbiased expected heterozygosity; σ , standard deviation.

and locus c19a had an estimated frequency of null alleles of 0.20 (Oosterhout measure) in U. Scarp. As the null alleles were not present in all populations, we analyzed the effect that the null alleles had on our data by comparing F_{ST} values without null alleles and taking into account the null allele proportions using FreeNa (Chapuis and Estoup, 2007). Because no difference was observed between the different panels, and corrected F_{ST} values controlling for null alleles were consistent with the uncorrected F_{ST} values, we kept loci D12S67 and c19a in all downstream analyses.

All loci were polymorphic in all forest blocks. The number of alleles per locus across populations ranged from 5 to 8, and per population ranged from 2 to 7 (mean \pm s.d.; 4.94 ± 1.10). Corrected mean allelic richness was $4.42 (\pm 0.31)$. The presence of private alleles was in general low (<0.20 in Magombera, Mwanihana, Matundu and Ndundulu), being more frequent in U. Scarp (0.84). The mean observed heterozygosity across loci and forest blocks was $0.67 (\pm 0.16)$, and the mean unbiased expected heterozygosity was $0.66 (\pm 0.12)$. Allelic richness, observed heterozygosity and expected heterozygosity were similar for all forest blocks (Table 1). Overall F_{IS} was -0.024 (upper/lower 95% confidence interval (CI) = -0.054 – 0.010) and not significantly different from 0. F_{IS} for the different fragments ranged from -0.061 in Magombera to 0.03 in U. Scarp but in no case was it significantly different from zero (Table 1).

Overall F_{ST} was 0.067 (upper/lower 95% CI = 0.041 – 0.095) and was significantly different from zero as were all the pairwise comparisons among the five forest blocks (Table 2). Using STRUCTURE, we identified two different clusters because ΔK (37.09) was maximized at $K=2$. At $K=2$, the results show that Magombera and U. Scarp form two different clusters, while the other three forests show varying levels of admixture between the aforementioned two forests. Because the $\ln P(D)$ (-3173.69) was maximized at $K=5$, and the ΔK has been shown to infer the uppermost level of structure when hierarchical population structure is present (Evanno *et al.*, 2005), we ran each cluster defined by $K=2$ separately (Figure 2). The first cluster reanalyzed included the forest blocks of U. Scarp (mean $Q=0.901$), Matundu (mean $Q=0.508$) and Ndundulu (mean $Q=0.757$). We decided to include Matundu in this cluster because although it is admixed the mean Q value across individuals was >0.5 . The analysis identified two different clusters with ΔK maximized at $K=2$ and separated U. Scarp from the other two forests. The second cluster included the forest blocks of Mwanihana (mean $Q=0.639$) and Magombera (mean $Q=0.851$). The results for the analysis of this cluster also identified two different clusters with ΔK maximized at $K=2$. These clusters separated the forests of Magombera and Mwanihana (see Figure 2). Thus our results suggest the presence of four clusters. STRUCTURE results from analysis of the data set that excluded first-degree relatives as well as results from the DAPC (both including and excluding relatives) are consistent with the existence of a complex hierarchical

structure and four different genetic clusters (Supplementary Material S2).

When analyzing recent gene flow between forest blocks using BayesAss, we found a high proportion of individuals derived from their own population in Magombera, U. Scarp and Ndundulu (Supplementary Table S3). Most of the proportions of migrant individuals detected had 95% CIs that overlapped with zero, thus indicating a lack of significance. The only forests that had a proportion of migrants that deviated significantly from zero were in Mwanihana ($m=0.26$, 95% CI 0.32 – 0.20), Matundu ($m=0.22$, 95% CI 0.32 – 0.12) and Ndundulu ($m=0.09$, 95% CI 0.17 – 0.01), with Magombera being the source of the migrants in all three cases (Supplementary Table S2).

Landscape genetics

After filtering the data set of full sibling and parent offspring dyads, we obtained a sample of 90 individuals across 36 locations from which genetic distances were derived. Overall, we found a significant signature of isolation by distance among individuals, but Euclidean distance explained only 14.1% (Pearson's $r=0.376$) of population structure (Supplementary Figure S4). We found that three resistance surfaces explained substantially more variation in the pairwise genetic data than distance alone (distance to the nearest village, fire density and TPI; Table 3). Ruggedness, a measure of the topographic complexity of the landscape, railroads, and current and historical land cover explained less variation in the pairwise genetic data than distance. There was little difference between historical land cover and contemporary forest cover based on the bootstrap \bar{w} and $\hat{\pi}$, while the density of fires and distance to the nearest village best explained the observed genetic distances (Table 3). Specifically, resistance increased as the density of fires increased, and resistance decreased as the distance from village increased (Supplementary Figures S3C and S5). Landscape resistance was estimated to be greater in low areas and in the mountains (negative and positive TPI, respectively; Supplementary Figure S5); however, TPI was infrequently the top model in the bootstrap analysis ($\hat{\pi}=0.089$, Table 3), suggesting its influence on gene flow may be limited.

We constructed composite resistance surfaces using the three surfaces that individually explained the genetic data substantially better than Euclidean distance: distance to the nearest village, density of fire, and TPI (Table 4). None of these layers were correlated with one another as the Spearman coefficient never exceeded $\rho \pm 0.04$. Following bootstrap model selection, there was clear support for a single top model (Composite C) that was the combination of fire density and distance to the nearest village (Figure 3; Table 4). As in the single surface optimization, resistance increased as the density of fires increased and resistance decreased as the distance to the nearest village increased. This composite surface had the highest average rank, greatest average weight \bar{w} and was the top model in 71% of bootstrap iterations (Table 4). The cumulative average weight of the three surfaces included in the composite analysis further reinforces that fire density and distance to the nearest village are the most important surfaces affecting gene flow of red colobus monkeys across the landscape (Sum \bar{w} : distance from village = 0.760 , fire density = 0.734 , TPI = 0.477). The cumulative average weight of models that included TPI were 1.5–1.6 times lower. Our visualization of current density according to Composite Surface C highlights several regions where dispersal is most likely to occur (Figure 4). Although the current flow in the unbiased assessment provides a visualization of the current density across the landscape (Figure 4a), the biased assessment (Figure 4b) is more reflective of current density across our sampling areas. Current density on the landscape is the greatest in areas that are

Table 2 Pairwise F_{ST} on the lower half and Euclidean distance (km) on the upper half

	U. Scarp	Matundu	Ndundulu	Mwanihana	Magombera
U. Scarp	0	60.85	83.29	119.13	131.54
Matundu	0.113	0	22.52	58.51	74.35
Ndundulu	0.074	0.046	0	36.15	54.96
Mwanihana	0.091	0.026	0.040	0	27.67
Magombera	0.155	0.048	0.087	0.031	0

Significant pairwise comparisons are in bold ($P < 0.005$).

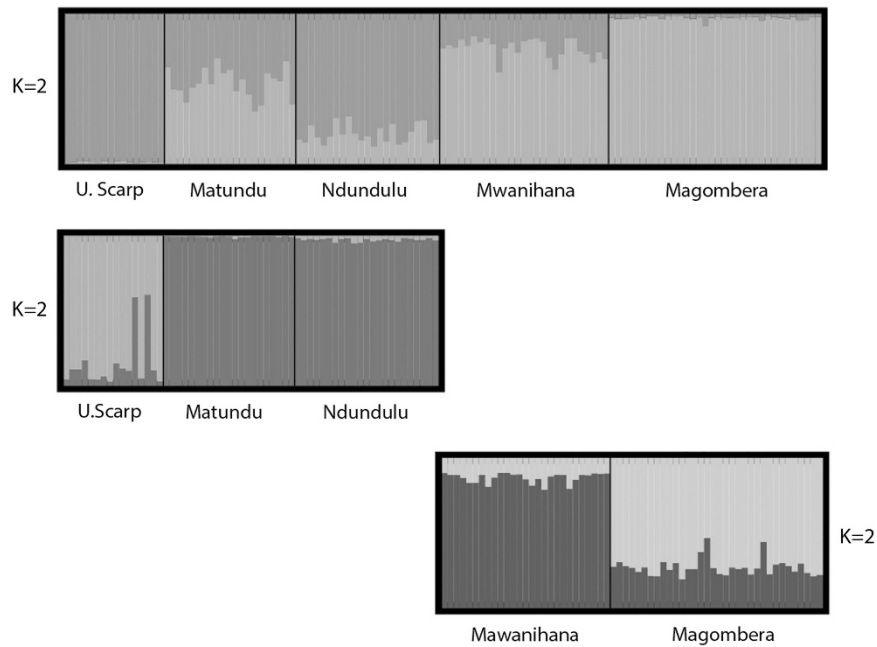


Figure 2 Hierarchical analyses of structure for the different forest blocks. Individuals are organized by the forests in which they are found (Udzungwa Scarp, Matundu, Ndundulu, Mwanihana, Magombera).

Table 3 Model selection results for linear mixed-effects models optimized on pairwise genetic differentiation

Surface	K	Avg. rank	\bar{w}	$\hat{\pi}$
Village	4	1.865	0.375	0.592
Fire density	4	2.393	0.243	0.288
TPI	4	3.537	0.109	0.089
Distance	2	4.470	0.048	0.016
Rugged	4	4.611	0.085	0.014
Historical forest	7	5.428	0.047	0.001
Railroad	3	6.280	0.047	0.000
Current forest	4	7.417	0.046	0.000

Abbreviations: Avg. rank, average model ranking from 10 000 bootstrap iterations; K, number of parameters in the transformation of continuous surfaces or number of land cover classes in categorical surfaces; TPI, topographic position index; \bar{w} , Akaike weight representing the probability that a model is the best in the model set, averaged over 10 000 bootstrap iterations; $\hat{\pi}$, proportion of 10 000 bootstrap iterations that a model was the top model. Only surfaces that were selected as the top model more frequently than distance alone were used to develop composite models.

Table 4 Model selection results for individual and all composite resistance surfaces evaluated

Surface	K	Avg. rank	\bar{w}	$\hat{\pi}$	Parameters
Composite C	7	1.632	0.297	0.708	Fire density, village
Fire	4	4.012	0.094	0.085	Fire density
TPI	4	5.956	0.060	0.062	TPI
Composite A	10	2.737	0.257	0.061	Fire density, TPI, village
Village	4	3.607	0.132	0.053	Village
Composite D	7	4.921	0.073	0.031	Fire density, TPI
Composite B	7	5.136	0.085	0.000	TPI, village

Abbreviations: Avg. rank, the average model ranking from 10 000 bootstrap iterations; K, number of parameters in the transformation of continuous surfaces or number of land cover classes in categorical surfaces; TPI, topographic position index; \bar{w} , Akaike weight representing the probability that a model is the best in the model set, averaged over 10 000 bootstrap iterations; $\hat{\pi}$, proportion of 10 000 bootstrap iterations that a model was the top model. Each surface consists of a combination of distance to the nearest village, fire density and TPI, which were the best-supported individual resistance surfaces based on 10 000 bootstrap iterations.

furthest from villages and have the lowest densities of fire. Because this composite resistance surface does not explicitly include current forest cover, we ran a generalized linear model with a binomial error distribution to further understand how estimated current density relates to existing forest coverage. Results showed that the probability of the landscape being forested increases significantly as current density increases ($z_{1,498} = 6.079$, $P < 0.0001$; Supplementary Figure S6). Thus there is a tendency for more current to be directed through forested habitat. Parameter estimates from linear mixed effects models fit to each optimized surface are in Supplementary Tables S4 and S5.

DISCUSSION

In this study, we implemented a new landscape genetic approach to assess whether forest fragmentation has affected genetic diversity and gene flow in the Udzungwa red colobus monkey and whether natural or human-related landscape features best explain genetic differentiation. In doing so, we present the first empirical application of 'ResistanceGA' to simultaneously optimize multiple resistance surfaces and create novel composite surfaces describing resistance to gene flow, and our flexible optimization procedure improves upon previous methods of assigning resistance values to surfaces as it makes no *a priori* assumptions about the direction or scale of the resistance relationship. Our findings suggest that genetic diversity is similar across the different forest blocks studied despite differences in forest size and species abundance. However, genetic differentiation among these blocks is present, and the landscape analyses suggest that a combination of human-related landscape features best explains this differentiation. Given the value of the Udzungwa red colobus monkey as an indicator species endemic to a global biodiversity hotspot, these results have important implications for understanding the impacts of human activity in the tropics.

Patterns of genetic diversity and genetic differentiation

Despite dramatic differences in forest area (Marshall *et al.*, 2010) and red colobus monkey population densities (Araldi *et al.*, 2014), we

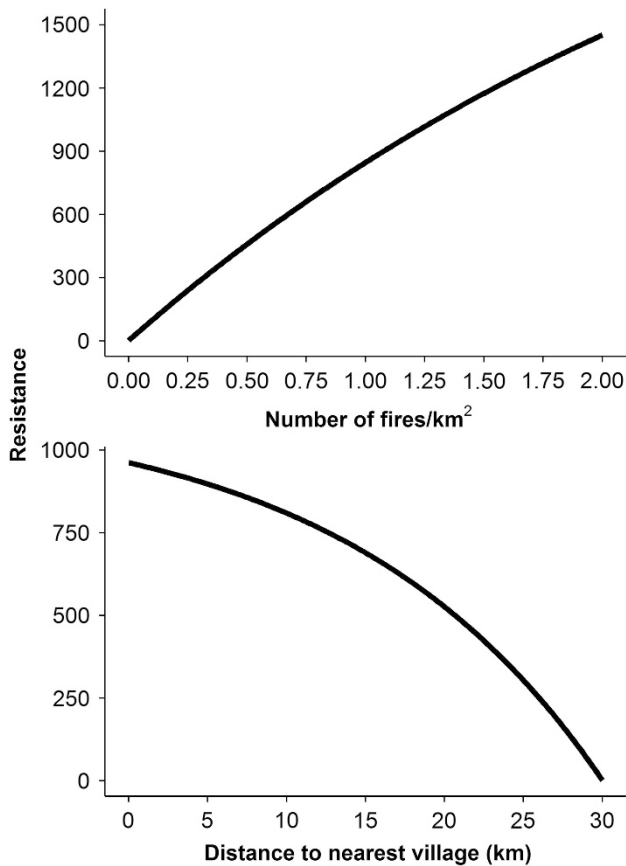


Figure 3 Response curves demonstrating the contribution and relationship of fire density and distance to the nearest village to landscape resistance within the best-supported composite resistance surface, Composite C.

found no strong differences in various measures of neutral genetic diversity across the five sampled forest blocks. This includes heterozygosity and allelic richness corrected for sample size. It is thus likely that habitat fragmentation in this system has been recent relative to red colobus monkey generation time (6–8 years) and/or effective population sizes, rendering any potential negative effects of genetic drift and inbreeding undetectable using standard measures of genetic diversity. Recent gene flow among some of the forest blocks (see below) may have also aided in maintaining similar levels of genetic diversity across the forests.

Despite a lack of differences in genetic diversity, we found significant levels of genetic differentiation among the five forest blocks and the existence of a complicated pattern of hierarchical structure. In general, at the highest hierarchical level, U. Scarp and Magombera form two differentiated clusters with the populations within the national park showing varying levels of admixture. When the clusters are further analyzed, the two aforementioned forest blocks again separate from the other forests, confirming that they are the most differentiated ones. In particular, U. Scarp is the most isolated forest as it forms the most differentiated population cluster and displays the highest proportion of private alleles after sample size correction, the highest pairwise F_{ST} values and a lack of recent migration. Magombera, in contrast, forms a distinct population cluster but displays lower pairwise F_{ST} values and possesses no private alleles. There is also evidence of recent asymmetrical migration (within the past several generations) flowing from Magombera into Matundu, Ndundulu and Mwanihana. This is consistent with the notion that red colobus

monkeys in small fragments such as Magombera suffer from population compression owing to recent isolation, clearance of nearby forests and extraction of commercial timber resources (Marshall, 2007). Although these results demonstrate that varying levels of genetic differentiation exist among red colobus monkey populations inhabiting the Udzungwa Mountain forests, they do not shed light on the exact causes of this differentiation. Furthermore, because of the significant relationship between genetic distance and geographical distance (Supplementary Figure S4), it is difficult to attribute genetic differentiation to naturally occurring forest blocks and/or human disturbance. This signal of isolation by distance, where populations in geographic proximity tend to be more similar than those that are far apart, commonly confounds studies such as this. It is thus important to implement a landscape approach that allows us to understand how species react to natural and anthropogenic factors.

Human-related landscape features explain genetic differentiation

Our use of 'ResistanceGA' to simultaneously optimize multiple resistance surfaces and create composite resistance surfaces demonstrated that a composite surface comprised of fire density and distance to the nearest village best explains the variation in pairwise genetic data. Although some fires on the landscape may be natural, the vast majority are maintained by humans, and previous studies indicate that bushfires are a major threat to these forests (Dinesen *et al.*, 2001). There is a long tradition of annual burning in this region, which maintains access to the forest, creates open areas for cattle grazing and agriculture and initiates hunting and honey collection. Ultimately, these fires inhibit forest regeneration, reduce forest cover and quality and create major barriers to dispersal for forest-dependent animals (Dinesen *et al.*, 2001). The strong influence of the presence of villages is also a major threat likely related to increased human disturbance surrounding such areas. Clearing for agricultural land, commercial and subsistence logging and hunting are all frequently associated with the increasing human population density adjacent to forests (Rovero *et al.*, 2014).

Railroad by itself was found to have relatively minimal ability to explain differences in pairwise genetic differentiation, but this is most likely because only a limited number of pairwise comparisons were directly separated by the railroad (Supplementary Figure S1). Surprisingly, neither current nor historical forest cover were good predictors of pairwise genetic differentiation. Although altitude has a strong effect on the abundance of this species (Barelli *et al.*, 2014), there was only moderate support for the effect of topographic position when assessed by itself (Table 4), and very little support when assessed in combination with fire density and distance to the nearest village. As such, it does not seem that altitude is acting as a barrier to gene flow. These results emphasize the effects of human disturbance and the ability of a landscape approach to detect the impacts of relatively recent landscape features on a long-lived species.

Although it is tempting to conclude from these results that human-mediated forest fragmentation is the primary driving force of Udzungwa red colobus monkey genetic differentiation, this is not necessarily the case. In fact, it is possible that some of the fragmentation is natural and quite old, but in such cases human settlement and activity in the intervening matrix habitats has likely maintained fragmentation and is contributing substantially to driving genetic differences among forests. For example, fire density might be the main driver of genetic differentiation, because it is negatively and significantly correlated with forest coverage ($\rho = -0.403$, $P < 0.001$), and it thus accounts for both forest coverage and other types of human disturbance that adversely affect red colobus gene flow. This is

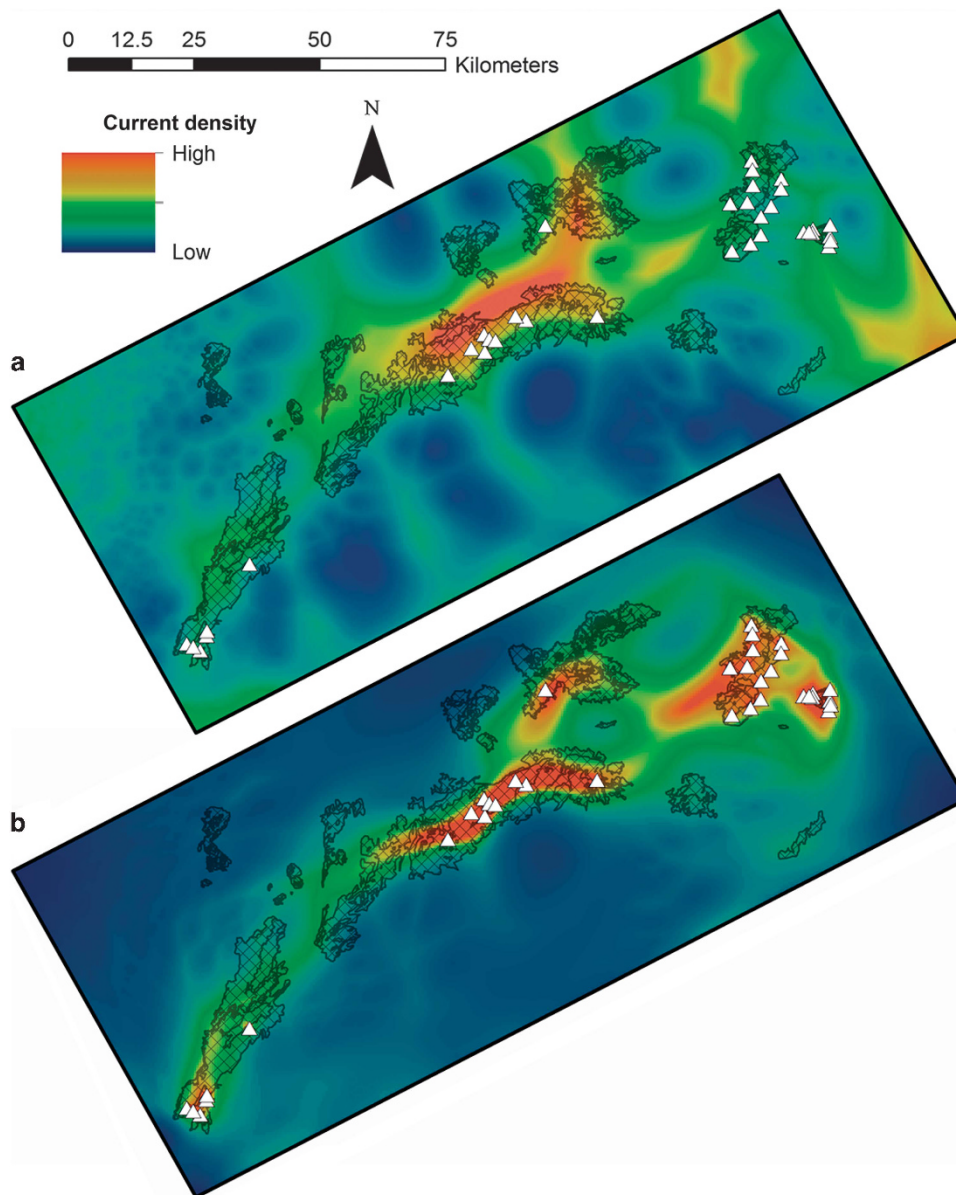


Figure 4 Maps showing current flow and demonstrating areas of potential red colobus monkey gene flow across the Udzungwa landscape. The underlying resistance surface for these maps is derived from the best composite resistance surface that combined fire density and distance to the nearest village. White triangles indicate locations where genetic samples were collected and the hatched polygons are forested habitat on the landscape. (a) Unbiased current flow map where current flows originates from points outside the focal landscape but flows across it. (b) Biased current map where current flows between sample locations.

consistent with the notion that human disturbance can overwhelm other factors more commonly associated with distribution and abundance in this species (Rovero *et al.*, 2012). These results thus demonstrate the strong influence anthropogenic activities have had on the movement of red colobus monkeys, but they should not undermine the importance of maintaining forested habitat.

Conservation implications

The Udzungwa Mountains are a conservation priority in a global biodiversity hotspot. They harbor high levels of biological endemism, and consistent discovery of new plants and animals, including mammals, indicates a high level of cryptic species diversity (Dinesen *et al.*, 2001). A landscape genetic study of an Udzungwa indicator

species is thus important for assessing the health of this important ecosystem. Our genetic study of the Udzungwa red colobus monkey complements previous ecological and demographic studies (Struhsaker *et al.*, 2004; Marshall *et al.*, 2010; Rovero *et al.*, 2012; Araldi *et al.*, 2014) and indicates that this ecosystem, while relatively healthy and intact, is presently in a precarious state. Despite no current deleterious effects of genetic drift or inbreeding on genetic diversity among Udzungwa red colobus monkey populations, our method for generating composite landscape surfaces demonstrates that human-related features are combining to drive genetic differentiation among forest block populations. Although the maintenance of genetic diversity in the face of forest fragmentation is encouraging, efforts should be made to reduce human activities that are driving genetic differentiation in

the Udzungwa red colobus. In particular, corridors can be established by reducing the prevalence of forest burning and allowing for forest regeneration along areas of potential gene flow directly between forest blocks in our current flow map (Figure 4). If this action is not taken, it is possible that genetic drift and/or inbreeding will begin to affect wildlife populations in the smaller forests, which will increase the probability of local extinction. With regard to the sampled forest blocks, the U. Scarp forest is the most isolated and needs to be reconnected to the other forests via the Mngeta corridor (Rovero *et al.*, 2012; Araldi *et al.*, 2014), which coincides with an area of potential dispersal in the current flow map. Furthermore, small and isolated forest fragments such as Magombera require better protection to counter increasing degradation and population compression. Implementing such plans now will prevent future losses in genetic diversity and species extinction across taxa in this biodiversity hotspot.

CONCLUSIONS

Studying the effects of habitat loss in the tropics has generated much interest over the years owing to the high levels of biodiversity in tropical forests. Although genetic studies have had difficulty in detecting the effects of fragmentation in tropical forest-adapted taxa (Radespiel and Bruford, 2014), landscape genetics approaches are rarely used in such regions (Manel and Holderegger, 2013). Meanwhile, assignment of resistance values to landscape features has been a challenge for the field of landscape genetics as different methods contain biological assumptions that are not appropriate for all circumstances. We presented an empirical example of a novel landscape genetic approach that improves upon previous methods by creating composite surfaces and optimizing multiple resistance surfaces without making *a priori* assumptions regarding the direction or scale of the resistance relationship. Application of this method to the Udzungwa red colobus monkey facilitated a deeper understanding of the consequences of human activities in a fragmented tropical forest environment. Our results reinforce the notion that recent modifications and activities on the landscape can drive genetic differentiation, even in a species with a relatively long generation time, and a variety of human activities are combining to exert a strong effect on the genetic differentiation of tropical forest-adapted taxa.

DATA ACCESSIBILITY

The STR genotypes are available on Dryad at <https://datadryad.org/resource/doi:10.5061/dryad.rr611>.

CONFLICT OF INTEREST

There authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the Tanzania Commission for Science and Technology (COSTECH), Tanzania Wildlife Research Institute (TAWIRI) and Tanzania National Parks (TANAPA) for granting us permission to conduct the study. We also thank the warden and staff of Udzungwa Mountains National Park, the Tanzanian field assistants, A Araldi, JF Gallardo Palacios and T Wolf. This manuscript was greatly improved through the feedback of the editor and three anonymous reviewers. Financial support for this research was provided by the Provincia Autonoma di Trento and the EU (Marie Curie Actions COFUND, postdoctoral grant to CB), MUSE—Museo delle Scienze to FR and CB, Rufford Small Grants Foundation (1033-C to FR), the German Primate Centre (DPZ) and in part by NIH grant TW009237 as part of the joint NIH—NSF Ecology of Infectious Disease program and the UK Economic and Social Research Council.

- Akaike H (1974). A new look at the statistical model identification. *IEEE Trans Autom Control* **19**: 716–723.
- Araldi A, Barelli C, Hodges K, Rovero F (2014). Density estimation of the endangered Udzungwa Red Colobus (*Procolobus gordonorum*) and other arboreal primates in the Udzungwa Mountains using systematic distance sampling. *Int J Primatol* **35**: 941–956.
- Barelli C, Gallardo Palacios J, Rovero F (2014). Variation in primate abundance along an elevational gradient in the Udzungwa Mountains of Tanzania. In: Grow NB, Gursky-Doyen S, Krzton A (eds). *High Altitude Primates*. Springer: New York, USA. pp 211–226.
- Bates D, Maechler M, Bolker B, Walker S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. Available from <http://CRAN.R-project.org/package=lme4>.
- Burnham KP, Anderson DR (eds) (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag: New York, USA.
- Chapuis M-P, Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* **24**: 621–631.
- Dilts T. (2010). Topography Tools for ArcGIS v. 9.3. Available from <http://arcscripts.esri.com/details.asp?dbid=15996>.
- Dinesen L, Lehmer T, Rahner MC, Fjeldsø J (2001). Conservation priorities for the forests of the Udzungwa Mountains, Tanzania, based on primates, duikers and birds. *Biol Conserv* **99**: 223–236.
- Dudaniec RY, Rhodes JR, Worthington Wilmer J, Lyons M, Lee KE, McAlpine CA *et al.* (2013). Using multilevel models to identify drivers of landscape-genetic structure among management areas. *Mol Ecol* **22**: 3752–3765.
- 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.
- Goudet J (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* **86**: 485–486.
- Hansen M, DeFries R, Townshend JRG, Sohlberg R (1998). *UMD Global Land Cover Classification, 1 Kilometer, 1.0, 1981–1994*. Department of Geography, University of Maryland: College Park, Maryland, USA.
- Hardy OJ, Vekemans X (2002). spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* **2**: 618–620.
- Holderegger R, Wagner HH (2008). Landscape genetics. *Bioscience* **58**: 199–207.
- Jombart T, Devillard S, Balloux F (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* **11**: 94.
- Kalinowski ST (2004). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Resour* **5**: 187–189.
- Kalinowski ST, Wagner AP, Taper ML (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* **6**: 576–579.
- Koen EL, Bowman J, Sadowski C, Walpole AA (2014). Landscape connectivity for wildlife: development and validation of multispecies linkage maps. *Methods Ecol Evol* **5**: 626–633.
- Manel S, Holderegger R (2013). Ten years of landscape genetics. *Trends Ecol Evol* **28**: 614–621.
- Marshall AR (2007). *Disturbance in the Udzungwas: Responses of monkeys and trees to forest degradation*. PhD thesis University of York: UK.
- Marshall AR, Jorgensby H, Rovero F, Platts PJ, White PC, Lovett JC (2010). The species-area relationship and confounding variables in a threatened monkey community. *Am J Primatol* **72**: 325–336.
- McRae BH (2006). Isolation by resistance. *Evolution* **60**: 1551–1561.
- Nsubuga AM, Robbins MM, Roeder AD, Morin PA., Boesch C, Vigilant L. (2004). Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Mol Ecol* **13**: 2089–2094.
- Peakall ROD, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.
- Peterman WE (2014). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *bioRxiv*; doi:10.1101/007575.
- Peterman WE, Connette GM, Semlitsch RD, Eggert LS (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol Ecol* **23**: 2402–2413.
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Radespiel U, Bruford MW (2014). Fragmentation genetics of rainforest animals: insights from recent studies. *Conserv Genet* **15**: 245–260.
- Rice W (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
- Rousset (2000). Genetic differentiation between individuals. *J Evol Biol* **13**: 58–62.
- Rousset F. (2008). Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* **8**: 103–106.
- Rovero F, Menegon M, Fjeldsø J, Collett L, Daggart N, Leonard C *et al.* (2014). Targeted vertebrate surveys enhance the faunal importance and improve explanatory models within the Eastern Arc Mountains of Kenya and Tanzania. *Divers Distrib* **20**: 1438–1449.
- Rovero F, Mtui AS, Kitegile AS, Nielsen MR (2012). Hunting or habitat degradation? Decline of primate populations in Udzungwa Mountains, Tanzania: an analysis of threats. *Biol Conserv* **146**: 89–96.
- Scrucca L (2013). GA: a package for genetic algorithms in R. *J Stat Softw* **53**: 1–37.
- Struhsaker T, Marshall A, Detwiler K, Siex K, Ehardt C, Lisbjerg D *et al.* (2004). Demographic variation among Udzungwa red colobus in relation to gross ecological and sociological parameters. *Int J Primatol* **25**: 615–658.

- Struhsaker TT, Butynski TM, Ehardt C. (2008). *Procolobus gordonorum*. The IUCN Red List of Threatened Species. Version 2014.2 <www.iucnredlist.org>. Downloaded on 01 September 2014.
- Van Oosterhout C, Van Heuven MK, Brakefield PM (2004). On the neutrality of molecular genetic markers: pedigree analysis of genetic variation in fragmented populations. *Mol Ecol* **13**: 1025–1034.
- Van Strien MJ, Keller D, Holderegger R (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* **21**: 4010–4023.
- Waits LP, Luikart G, Taberlet P (2001). Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Mol Ecol* **10**: 249–256.
- Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**: 1177–1191.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)