APPLICATION ESSAYS

Can molecular data validate morphometric identification of faecal pellets in Tanzanian forest antelope species?

Andrew E. Bowkett · Trevor Jones ·
Richard L. Laizzer · Amy B. Plowman ·
Jamie R. Stevens

Abstract Forest-dwelling antelope species are often difficult to detect during surveys due to their cryptic behaviour and densely vegetated habitats. Dung counts have traditionally been used to infer forest antelope abundance but genetic identification has shown that visual identification of ungulate dung to species is often unreliable. This study attempted to use easily obtained morphometric data from faecal pellets to statistically assign antelope dung piles to species. We measured pellets from 238 dung piles collected from the Udzungwa Mountains, south-central Tanzania, a largely forested landscape with five forest-associated antelope species including the endangered Abbott's duiker *Cephalophus spadix*. The species identity of sampled dung piles was determined by amplifying a c. 600 bp fragment of the mitochondrial control region and aligning DNA sequences with published references. We found no diagnostic differences in faecal pellet size between antelope species although there were significant differences in mean pellet length and width. We employed a single variable linear discriminant analysis to predict the species of dung piles based on pellet length. Despite significant differentiation between species we obtained an overall accuracy of 58.8 % that did not meet our specified probability threshold ($P < 0.05$). Abbott's duiker dung piles were correctly assigned in the majority of cases (74 %). Overall, morphometric assignment of dung piles to species was not accurate enough to validate dung counts as a survey method for forest antelope although our results do not preclude the development of alternative field identification methods using additional non-molecular characters.

Keywords Bar-coding · Mitochondrial D-loop · Neotragus · Tragelaphus · Philantomba

A. E. Bowkett (✉) · J. R. Stevens
Molecular Ecology and Evolution Group, Biosciences, College of Life and Environmental Science, University of Exeter, Exeter EX4 4QD, UK
e-mail: andrew.bowkett@paigntonzoo.org.uk

A. E. Bowkett · A. B. Plowman
Field Conservation and Research Department, Whitley Wildlife Conservation Trust, Paignton Zoo, Totnes Road, Paignton TQ4 7EU, UK

T. Jones · R. L. Laizzer
Udzungwa Elephant Project, P.O. Box 99, Mang'ula, Tanzania

T. Jones
Animal and Environmental Research Group, Department of Life Sciences, Anglia Ruskin University, Cambridge CB1 1PT, UK

Introduction

Encounter rates and density estimates of faecal deposits are often used to infer the abundance of ungulates for research or management (Campbell et al. 2004; Lunt et al. 2007; Plumptre and Harris 1995). However, genetic identification of ungulate faeces (referred to here as dung), by amplifying species-specific fragments of mitochondrial DNA, has revealed that dung piles are often mistakenly identified in the field (Faria et al. 2011; van Vliet et al. 2008; Yamashiro et al. 2010). This is a particular problem when species of conservation concern, for example the mountain bongo *Tragelaphus eurycerus isaaci* in Kenya (Faria et al. 2011), occupy the same areas as other ungulate species with similar dung morphology.

In common with the previously cited studies, Bowkett et al. (2009) found that identification of forest antelope dung to species in the field was unreliable, in this case in the Udzungwa Mountains, Tanzania. This result cast doubt on the validity of using dung counts as a survey and monitoring tool for the endangered Tanzanian endemic Abbott's duiker *Cephalophus spadix* and sympatric species. Abbott's duiker is found in only a small number of highland forests and is threatened by continuing habitat disturbance and hunting for bushmeat (Moyer et al. 2008). Like other forest antelopes, this species appears to be largely solitary and nocturnal (Rovero et al. 2005) and therefore unsuitable for survey methods based on sightings.

Surveys using genetic identification of dung and automatic camera-traps have been successful in documenting the presence of Abbott's duiker (Jones and Bowkett 2012). However, these methods are relatively expensive and not easily replicated as part of routine monitoring by wildlife authorities. Hibert et al. (2008) demonstrated the potential for discriminant analysis to identify savannah antelope species based on faecal pellet morphometric data. Bowkett et al. (2009) attempted to statistically assign dung piles to species based on the size of faecal pellets, a method that could easily be carried out in the field by local survey teams. However, their preliminary morphometric analysis only sampled a small part of the Udzungwa landscape and was limited by small sample size. For this study we employ a larger data set of genetically verified antelope dung piles from across the Udzungwa Mountains to assess whether dung from forest antelope can be reliably determined to species using simple size measurements.

Specifically, we aimed to test (1) whether or not diagnostic differences in faecal pellet measurements exist between target antelope species, (2) whether a discriminant analysis can assign pellets to their correct species group with a high statistical probability ($P < 0.05$), and (3) whether these approaches can distinguish Abbott's duiker pellets even if other species are unresolved. In addition, we test some of the assumptions of reliability inherent in applying such a method under field conditions.

Materials and methods

Antelope faecal pellets were collected from all but one of the major forest fragments (>10 km^2) in the Udzungwa Mountains, south-central Tanzania, between 2006 and 2009. General landscape and habitat details are given in Marshall et al. (2010). Each pile was visually assigned to species in order to test the accuracy of identification in the field. A subset of faecal pellets from each dung pile was stored in RNAlater (Ambion Ltd, Huntington, UK) for later DNA extraction (QIAamp DNA Stool Mini Kit, QIAGEN,

Crawley, UK). Molecular identification of samples involved amplifying c. 600 bp of the mitochondrial control region using the primer combinations and PCR conditions detailed by Ntie et al. (2010). We opted not to use the cytochrome c oxidase subunit 1 (*COX1*) as this marker fails to delineate several duiker species under standard DNA bar-coding criteria (Johnston et al. 2011). Sequences were aligned with published references for all potential antelope species using MUSCLE (Edgar 2004) in SEAVIEW (Gouy et al. 2010). We constructed a neighbour-joining phylogeny based on Kimura 2-parameter corrected distances (Kimura 1980) in PAUP* (Swofford 2001) to show that all haplotypes formed monophyletic species clades. The tree was rooted using the bushbuck *Tragelaphus scriptus* as an out-group. A highly variable 103 bp section was removed before tree-building as it proved difficult to align consistently.

For each dung pile, we measured up to twenty randomly selected pellets with calipers (to the nearest 0.01 mm) and subsequently treated the mean value as our sampling unit. We applied a bootstrapping procedure on a subset of samples, similar to that of Hibert et al. (2008), in order to ensure that the variance of the pellets measured approached that of the overall within-group variance. This procedure involved 1,000 permutations of each number of measured pellets between two and twenty (see Online Resource 1).

All measured dung piles were estimated to be less than 24 h old due to their shiny moist surface. Forest antelope faecal pellets are generally cylindrical often forming a point at one end with a shallow indentation at the other. Faecal pellet illustrations and measurements for various antelope species are given in Chame (2003), including three of the species found in the Udzungwa Mountains. We measured length as the maximal distance between pellet ends and width as the diameter at the widest point. We also calculated the length to width ratio following Bowkett et al. (2009). Pellets were stored in tied plastic bags and either measured during the field survey or kept refrigerated (c. 4 °C) and measured at a later date (mean delay in days of 138 recorded dung piles = 6, maximum = 21, see below). Pellets were not dried before measurement as we wanted to test a method that could be easily replicated by protected area staff under humid forest conditions.

As well as testing for absolute inter-specific differences in mean pellet size we also attempted a linear discriminant analysis for statistically assigning dung piles to species. All three measured variables were normally distributed but highly inter-correlated ($r = 0.35$–0.73, $P < 0.001$). Length was identified as the variable with the highest discriminatory power following one-way ANOVA tests using species as a factor: length ($R^2 = 0.67$), width ($R^2 = 0.63$), ratio ($R^2 = 0.03$) where R^2 measures the proportion of the observed variance explained by the independent factor

species. The discriminant analysis was therefore carried out on the single variable Length (log_{10} transformed to improve homogeneity of variance). Box's M test was used to test for homogeneity in the variance–covariance matrices and the significance of discrimination between species tested with Wilks' Λ.

We also tested two assumptions in measuring faecal pellets that could potentially be violated under field conditions. Firstly, we tested for the effect of a delay between collecting and measuring pellets by measuring individual pellets on the day of collection and again after 31 days under two different storage conditions: refrigerated (n = 135) and ambient temperature (n = 119). Secondly, we checked for inter-observer bias (different survey personnel) by two of us (AEB and RLL) measuring the same 220 individual pellets. All statistical tests were carried out in SPSS 16 (SPSS Inc., Chicago, IL, USA).

Results

We measured pellets from 238 dung piles (4,654 individual pellets) from which we successfully obtained control region sequences (two additional piles did not yield readable sequences). This dataset included 48 dung piles measured for Bowkett et al. (2009). Sixty-eight dung piles (28.6 %) were incorrectly assigned to species by the survey teams.

All control region haplotypes recovered from dung samples aligned with published sequences for forest antelope species known from the Udzungwa Mountains and formed monophyletic clades with high bootstrap support (97–100 %, Fig. 1): Harvey's duiker *Cephalophus harveyi* (36 haplotypes), blue duiker *Philantomba monticola* (4), suni *Neotragus moschatus* (27), bushbuck (2) and Abbott's duiker (6). All sequence reads were at least 520 bp except for one Abbott's duiker sample (310 bp). Two sequences aligned closely to reference species but contained too many ambiguous bases to assign to a specific haplotype and so were excluded from the phylogeny shown here.

There was a statistically significant difference across the five species detected in untransformed mean pellet length (One way ANOVA: $F_{4, 233} = 120.3$, $P < 0.001$) and width ($F_{4, 233} = 98.3$, $P < 0.001$) but not in length: width ratio ($F_{4, 233} = 1.67$, $P = 0.16$). Similar P values were obtained with Welch's ANOVA which was carried out for width ($P < 0.001$) and ratio ($P = 0.15$) which did not meet the assumption of homogeneity of variance. However, the range of all variables overlapped between at least two species, thus no diagnostic differences in pellet size were apparent. Abbott's duiker pellets overlapped with those of Harvey's duiker and bushbuck (Fig. 2).

Discriminant analysis was carried out for faecal pellet length with equal prior probability assumed for all species. Significant discrimination was found between species (Wilks' Λ = 0.32, $\chi^2 = 265.28$, $P < 0.001$) and 58.8 % of samples were correctly classified. There was no significant deviation from homogeneity (Box's M test: $P = 0.131$). Correct identification varied by species (Table 1) with 74 % of Abbott's duiker dung piles correctly assigned with the remainder predicted to be either Harvey's duiker or bushbuck. Conversely, 16 % of Harvey's duiker dung piles were incorrectly predicted to be Abbott's duiker.

We found no statistical differences in the three pellet variables following 31 days storage in plastic bags (Independent t tests and Mann–Whitney tests depending on data distribution, all tests = $P > 0.05$). However, there was a greater decrease in size for pellets stored at ambient temperature (e.g. mean length [log_{10}]: day 1 = 9.07 mm, day 31 = 8.66 mm, $U = 6,159$, $P = 0.083$) compared to refrigerated (mean length [log_{10}]: day 1 = 9.15, day 31 = 9.11, $t_{268} = 0.196$, $P = 0.845$). We therefore assume our storage protocol had a minimal effect on this study's results (as pellets were refrigerated and usually measured within 7 days). We note that the difference in mean pellet length at ambient temperature approached statistical significance ($P < 0.1$).

There was a significant difference in mean length when two observers measured the same individual pellets (Paired t test: $t_{219} = 6.514$, $P < 0.001$, with similar results for width and ratio). However, the mean difference was only 0.09 mm, less than the within-group variance recorded for any dung pile measured for this study (mean variance = 1.62 ± 0.15). Therefore, we included measurements taken by either observer in our final analysis but we also note that a discriminant analysis with only AEB measurements (207 dung piles) produced highly similar results (AEB only: 58.9 % correct, Wilks' Λ = 0.326, $\chi^2 = 227.49$, $P < 0.001$, results for full data-set shown above).

Discussion

Overall our results indicate that morphometric assignment of antelope dung piles to species is not a valid method for identification using the characters measured during this study. Our results also further emphasize the inaccuracy of visual field identification (van Vliet et al. 2008; Bowkett et al. 2009).

We did not find diagnostic differences in pellet size between species and the discriminant analysis was unable to correctly identify species with high probability. The lack of absolute difference between species is perhaps not surprising given the likely individual variation due to age, sex
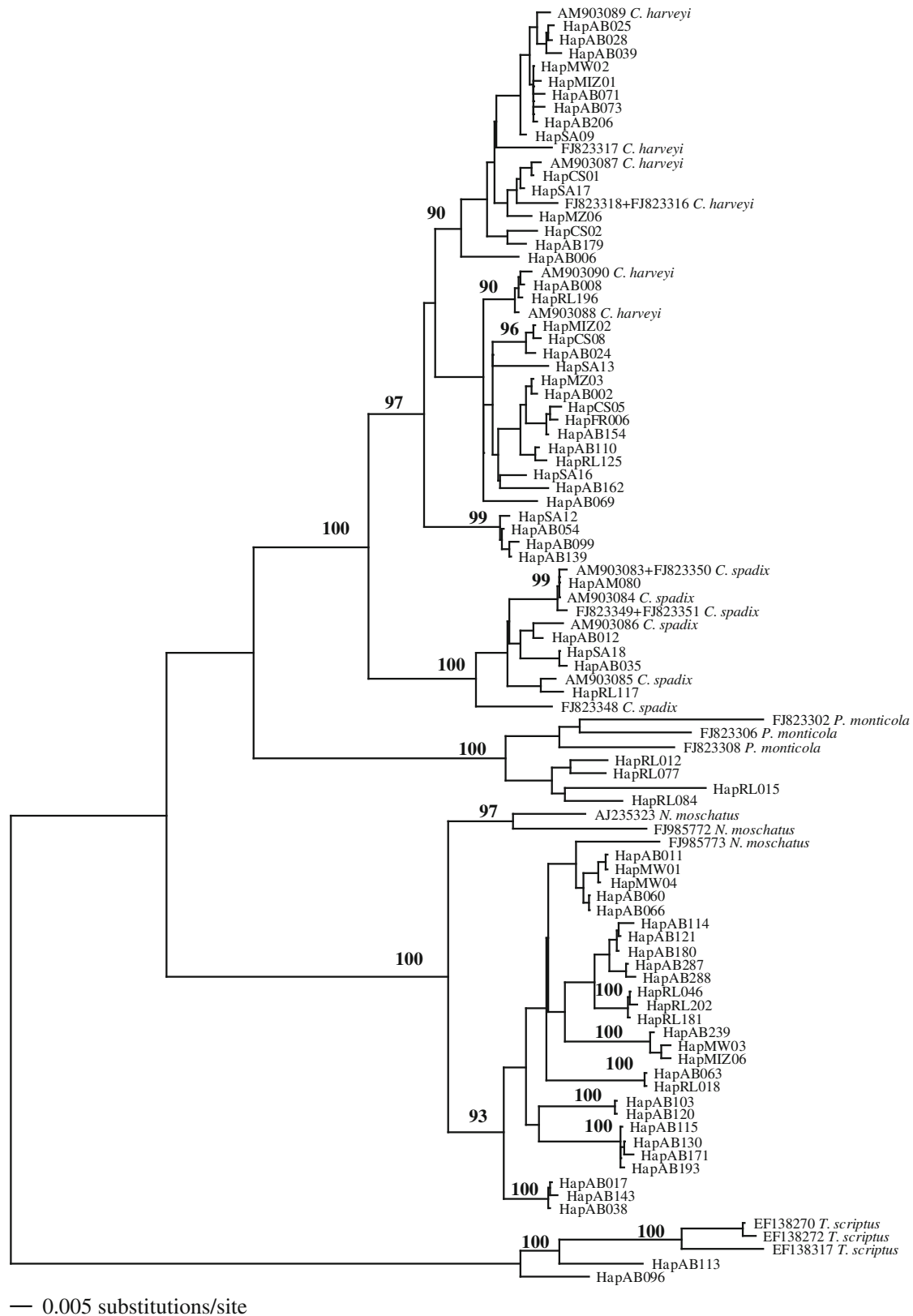
Fig. 1 Control region neighbour-joining bootstrap consensus phylogeny for faecal DNA sequences from the Udzungwa Mountains, Tanzania (c. 530 bp). Bootstrap values are percentages of 1,000 iterations (values below 90 % not shown). Gen Bank accession numbers and species names are given for published sequences

Fig. 2 Box-and-whisker plot of faecal pellet length for five species of forest antelope from the Udzungwa Mountains, Tanzania. Species identity determined from mitochondrial control region sequences
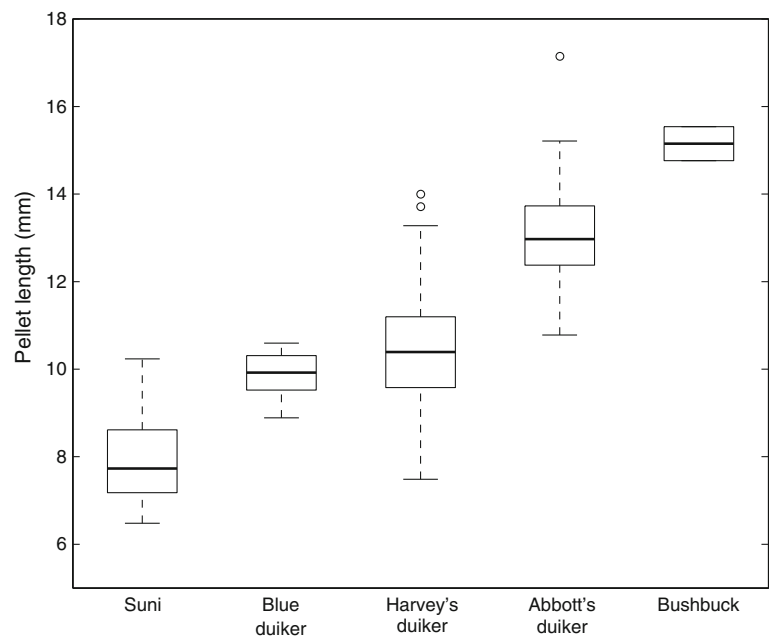

Table 1 Species identity of antelope dung piles from the Udzungwa Mountains, Tanzania, as established by mitochondrial control region sequencing compared to predicted species from a linear discriminant analysis of faecal pellet length

	mtDNA	Predicted species				
		Harvey's duiker	Blue duiker	Suni	Bushbuck	Abbott's duiker
Harvey's duiker	136	60 (44 %)	40 (29 %)	14 (10 %)	0	22 (16 %)
Blue duiker	5	2 (40 %)	3 (60 %)	0	0	0
Suni	57	1 (2 %)	9 (16 %)	47 (83 %)	0	0
Bushbuck	2	0	0	0	2 (100 %)	0
Abbott's duiker	38	5 (13 %)	0	0	5 (13 %)	28 (74 %)

and diet (Lunt and Mhlanga 2011). However, prior to this analysis it seemed likely that a statistical assignment based on mean pellet size would be valuable given the difference in adult body size between some species. For example, Harvey's duiker weighs 13–16 kg compared to 50–60 kg for Abbott's duiker (Kingdon 1997) and yet we found substantial overlap in faecal pellet size between the two species (Fig. 2).

One limitation of our discriminant approach was the small sample size included for the less frequently encountered species (bushbuck and blue duiker) even though these species clearly overlap with at least one other species. Given that blue duiker are absent from some areas of the Udzungwa Mountains (Rovero and Marshall 2009) and that bushbuck appear to be rare in our study area it could be argued that these species should be excluded from our analysis in order to improve accuracy. Indeed, the percentage of correctly assigned dung piles does improve if we run the analysis with only three species (77.5 % correct, Wilks' $\Lambda = 0.332$, $\chi^2 = 251.18$, $P < 0.001$). However, we do not

recommend this approach because the probability of correct identification still falls far short of our specified 95 %, the analysis would not be valid for large areas where blue duiker may be present, and camera-trap surveys reveal that Harvey's duiker and bushbuck are almost always found in the same forests as Abbott's duiker in the Udzungwa Mountains (Rovero and Marshall 2009; TJ unpubl. data).

The accuracy of field identification of forest antelope dung could potentially be improved by the recording of additional characters such as pellet weight, colour, smell or the presence of known food-plant fragments. For example, although we did not quantify pellet colour for each sample, several dung piles were noted as containing yellow pellets in the field and all of these were subsequently identified as suni. However, in our experience the aforementioned characters are difficult to measure under varied field conditions with multiple observers and most could not easily be incorporated into routine protected area patrols.

Although we cannot recommend the morphometric approach to dung identification tested here, this study does

demonstrate the value of genetic identification as a non-invasive approach to detecting rare species in remote areas (see also Jones and Bowkett 2012). Camera-trapping is another survey tool that has proven useful for Abbott's duiker (Rovero et al. 2005; Rovero and Marshall 2009; Jones and Bowkett 2012) but similarly to genetic identification of dung it requires significant financial and scientific input. These resources are not normally available to protected area staff in Tanzania but such investment in the relatively few areas where viable Abbott's duiker populations remain is recommended for the conservation of what may well be Africa's rarest duiker species.

References

Bowkett AE, Plowman AB, Stevens JR, Davenport TRB, van Vuuren BJ (2009) Genetic testing of dung identification for antelope surveys in the Udzungwa Mountains Tanzania. Conserv Genet 10(1):251–255. doi:10.1007/s10592-008-9564-7

Campbell D, Swanson GM, Sales J (2004) Comparing the precision and cost-effectiveness of faecal pellet group count methods. J Appl Ecol 41(6):1185–1196

Chame M (2003) Terrestrial mammal feces: a morphometric summary and description. Mem Inst Oswaldo Cruz 98(suppl. I):71–94

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5(19):1–19

Faria PJ, Kavembe GD, Jung'a JO, Kimwele CN, Estes LD, Reillo PR, Mwangi AG, Bruford MW (2011) The use of non-invasive molecular techniques to confirm the presence of mountain bongo *Tragelaphus eurycerus isaaci* populations in Kenya and preliminary inference of their mitochondrial genetic variation. Conserv Genet 12(3):745–751. doi:10.1007/s10592-011-0181-5

Gouy M, Guindon S, Gascuel O (2010) Seaview version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27(2):221–224. doi:10.1093/molbev/msp259

Hibert F, Fritz H, Poilecot P, Abdou HN, Dulieu D (2008) Morphological criteria to identify faecal pellets of sympatric ungulates in West African savanna and estimates of associated error. Afr J Ecol 46(4):523–532. doi:10.1111/j.1365-2028.2007.00889.x

Johnston AR, Morikawa MK, Ntie S, Anthony NM (2011) Evaluating DNA barcoding criteria using African duiker antelope (Cephalophinae) as a test case. Conserv Genet 12(5):1173–1182. doi:10.1007/s10592-011-0220-2

Jones T, Bowkett AE (2012) New populations of an Endangered Tanzanian antelope confirmed using DNA and camera traps. Oryx 46(1):14–15

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. J Mol Evol 16(2):111–120. doi:10.1007/bf01731581

Kingdon J (1997) Field guide to African mammals. Academic Press, Oxford

Lunt N, Mhlanga MR (2011) Defecation rate variability in the common duiker: importance of food quality, season, sex and age. S Afr J Wildl Res 41(1):29–35. doi:10.3957/056.041.0113

Lunt N, Bowkett AE, Plowman AB (2007) Implications of assumption violation in density estimates of antelope from dung-heap counts: a case study on grey duiker (*Sylvicapra grimmia*) in Zimbabwe. Afr J Ecol 45(3):382–389. doi:10.1111/j.1365-2028.2006.00724.x

Marshall AR, Jorgensbye HI, 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(4):325–336. doi:10.1002/ajp.20787

Moyer DC, Jones T, Rovero F (2008) Cephalophus spadix. In: IUCN Red List of Threatened Species. Version 2011.2. www.iucnredlist.org Accessed on 30 May 2012

Ntie S, Johnston AR, Mickala P, Bowkett AE, Jansen van Vuuren B, Colyn M, Telfer P, Maisels F, Hymas O, Rouyer RL, Wallace RA, LeBlanc K, van Vliet N, Sonet G, Verheyen E, Pires D, Wickings EJ, Lahm SA, Anthony NM (2010) A molecular diagnostic for identifying central African forest artiodactyls from faecal pellets. Anim Conserv 13(1):80–93. doi:10.1111/j.1469-1795.2009.00303.x

Plumptre AJ, Harris S (1995) Estimating the biomass of large mammalian herbivores in a tropical montane forest: a method of fecal counting that avoids assuming a steady-state system. J Appl Ecol 32(1):111–120

Rovero F, Marshall AR (2009) Camera trapping photographic rate as an index of density in forest ungulates. J Appl Ecol 46(5):1011–1017. doi:10.1111/j.1365-2664.2009.01705.x

Rovero F, Jones T, Sanderson J (2005) Notes on Abbott's duiker (Cephalophus spadix True 1890) and other forest antelopes of Mwanihana Forest, Udzungwa Mountains, Tanzania, as revealed by camera-trapping and direct observations. Trop Zool 18(1):13–23

Swofford D (2001) PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b2. Sinauer Associates, Sunderland, Massachusetts

van Vliet N, Zundel S, Miquel C, Taberlet P, Nasi R (2008) Distinguishing dung from blue, red and yellow-backed duikers through noninvasive genetic techniques. Afr J Ecol 46(3):411–417. doi:10.1111/j.1365-2028.2007.00879.x

Yamashiro A, Yamashiro T, Baba M, Endo A, Kamada M (2010) Species identification based on the faecal DNA samples of the Japanese serow (*Capricornis crispus*). Conserv Genet Resour 2(S1):409–414. doi:10.1007/s12686-010-9281-7