AN ABSTRACT OF THE THESIS OF

Rachel Selena Crowhurst for the degree of Master of Science in Wildlife Science presented on February 27, 2012.

Title: Landscape Features Affecting Genetic Diversity and Structure in East African Ungulate Species.

Abstract approved:

_____

Clinton W. Epps

Habitat loss and fragmentation is a crisis affecting wildlife worldwide. In Tanzania, East Africa, a dramatic and recent (<80 years) expansion in human settlement and agriculture threatens to reduce gene flow among protected areas for many species of large mammals. Wildlife linkages can mitigate population isolation, but linkage designs lacking empirical justification may be controversial and ineffective. Connectivity conservation requires an understanding of how biogeographic factors shaped gene flow prior to habitat loss or fragmentation, however the history of interaction among populations is rarely known. The goal of my study was to provide context for connectivity conservation in central and southern Tanzania by identifying landscape features that have shaped gene flow for three ungulate species with different dispersal capabilities.

I investigated historical patterns of connectivity for Maasai giraffe (*Giraffa camelopardalis tippelskirchi*), impala (*Aepyceros melampus*), and eland (*Tragelaphus oryx*) by estimating genetic structure among four to eight protected areas per species. Genetic structure changes very slowly among large populations and thus is likely to reflect historical processes instead of recent anthropogenic influences. I collected noninvasive DNA samples and generated microsatellite genotypes at 8 to 15 loci per species, then estimated genetic diversity metrics (allelic richness, A_R, and expected

heterozygosity, H_E) for each population (defined by reserve). I also calculated genetic distance (F_{ST} and Nei's unbiased genetic distance, D_{hat}) and an estimate of gene flow (*Nm*) between all population pairs for each species.

To elucidate the possible causes of genetic structure between these populations, I tested for isolation by distance and isolation by resistance based on a suite of biogeographic factors hypothesized to affect gene flow. To do this, I created GIS-based resistance surfaces that assigned different costs of movement to landscape features. I created one or more resistance surfaces for each hypothesis of landscape effect. I used circuit theory to estimate the cumulative resistance between each pair of reserves for each weighting scheme, and then performed Mantel tests to calculate the correlation between these resistances and the observed population pairwise genetic distances (D_{hat}). I chose the optimal resistance model for each species as the model that was most highly correlated with observed genetic patterns. To verify that the correlation of resistance models with genetic distance was not an artefact of geographic distance, I performed partial Mantel tests to calculate correlation while controlling for the effect of geographic distance. Finally, I compared historical gene flow patterns to the distribution of contemporary human activity to predict areas that are at risk of a loss of connectivity.

Indices of genetic diversity were moderate for all three species and comparable to previously reported values for other savannah ungulates. Diversity (both H_E and A_R) was highest in eland and lowest in giraffe for these populations, and was not consistently correlated with reserve size as has been reported for other species in East Africa. Although patterns in genetic distance were broadly similar across all three species there were also striking differences in connectivity, highlighting the importance of cross-species comparisons in connectivity conservation.

At this scale, resistance models based on slope strongly predicted population structure for all three species; distance to water was also highly correlated with genetic distance in eland. For all three species, the greatest genetic distances occurred between populations separated by the Eastern Arc Mountains, suggesting that the topography

of this area has long acted as a barrier to gene flow, but this effect is present in varying degrees for each species. I observed high levels of historical gene flow between centrally located populations (Ruaha National Park and Rungwa Game Reserve) and those in the southwest (Katavi National Park and Rukwa Game Reserve). Although human settlement in this area has been low relative to other areas, the connection between the Katavi/Rukwa and Ruaha ecosystems may be threatened by increased human activity and warrants conservation.

High levels of historical gene flow were also seen between reserves in the northeast (Tarangire National Park, Swagaswaga Game Reserve) and the central and southwest populations. These connections appear highly threatened due to current land use practices, and may have already suffered a loss of gene flow. Field surveys in the lands surrounding the northeastern reserves are needed to quantify current levels of connectivity and determine whether corridors could be established to maintain or restore gene flow with other reserves.

Landscape Features Affecting Genetic Diversity and Structure in East African Ungulate Species

by
Rachel Selena Crowhurst

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented February 27, 2012
Commencement June 2012

Master of Science thesis of Rachel Selena Crowhurst presented on February 27, 2012.

APPROVED:

_____

Major Professor, representing Wildlife Science

_____

Head of the Department of Fisheries and Wildlife

_____

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

_____

Rachel Selena Crowhurst, Author

ACKNOWLEDGEMENTS

To my advisor, Clint Epps, thank you for allowing me to be a part of such an incredible and worthwhile project. Your devotion to this field and your unwavering commitment to excellence are both inspiring and humbling. Dagaa aside, I'd be honoured to work with you again. I also thank you for teaching me the meaning of 'self-defenestration' and for regularly encouraging me not to do it. I thank my committee members, Sue Haig and Mike Blouin for their helpful suggestions and encouragement throughout this process. Thank you to Alix Gitelman for serving as my graduate council representative.

This project would not have been possible without the assistance of many park managers, wardens, rangers, and staff at the beautiful national parks and game reserves in Tanzania. Thank you for your interest in this project and for all of the helpful information. I thank TANAPA, TAWIRI, and COSTECH for granting permission to conduct this research in the reserves.

I was fortunate to work with many dedicated students and assistants on this project. In Tanzania, Adam Setembo, Aklei Albert, Joachim Sindano, and Msafiri Mgumba were instrumental in collecting samples, keeping the camp and vehicle running smoothly, and prepping samples for extraction -- asanteni sana na samahani kidogo kwa ndorobo! Phillip Grenz and Keely Lopez spent many hours in the laboratory at Oregon State University scraping and extracting samples with unwavering good nature. I am grateful to Tom Mullins for conducting the Illumina sequencing to develop the giraffe primers, and Caprice Rosato for taking time to help me troubleshoot some of my early microsatellite runs.

I am honoured to have received funding assistance from a variety of sources: the Natural Sciences and Engineering Research Council of Canada, American Society of Mammalogists, the Bill Burtness Fellowship from the Rocky Mountain Goats Foundation, David B. and Georgia Leupold, Gamma Sigma Delta, the Savery Fund, and the Oregon lottery. Thank you for your confidence in me.

In Tanzania I was fortunate to meet many wonderful people who made me feel at home and helped me find my way (both literally and figuratively). I thank the Mtengeti family for their kindness, Peter Marwa Ezra for his friendship and dagaa-cooking tips, and Chantal Stoner for her compassion and infectious optimism. Professor Benezeth Mutayoba graciously allowed me to utilize his laboratory facilities at the Sokoine University of Agriculture, and assisted with the daunting logistics of the project. I had the experience of a lifetime in Tanzania, and you all shared no small part in making it so enjoyable.

The department of fisheries and wildlife has been like a family to me for the past three years; there are too many people to list but I say a heartfelt thanks to you all. Without the front office ladies (Jan, Jane, Trish, Kim, Nancy, Leighann), I'd still be buried under reams of paperwork; you ladies hide your frustration well! The IT personnel (Aaron and Scott) have been my knights in shining armour, rescuing me more times than I'd like to admit. Dana, thank you for being such a great listener and for your thoughtful advice. Thank you also to the Epps lab members (Katie, Mark, Jessica, Anne, Tyler, Brandon, and our honourary member Kendra) for the informative discussions and laughter - may you have no shortage of success or samples in your future careers.

Last but not least, I've been blessed to have had a wonderful group of supporters throughout my years at OSU. Aurélie, I thank you for your friendship. Pat, I appreciate your seemingly unending patience (now make that funny!). Matt – I thank you for all of the hot meals and for dragging me away from my computer and out to the woods when I needed it. And finally, to my parents, Peter and Kathy, and my sister Sarah – you made this all possible. You encouraged me when I didn't know if I had it in me. Thank you for the moral support, extended periods of doggy daycare, for your understanding when I went 'missing' for weeks on end, and for being so proud of a daughter/sister who collects dung for a living. I love you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

LIST OF APPENDICES

To my Mom and Dad,

Who strong-armed me into my first year of university
And cheered when I went back for more.

Landscape Features Affecting Genetic Diversity and Structure in East African Ungulate Species

INTRODUCTION

The survival of our wildlife is a matter of grave concern to all of us in Africa. These wild creatures amid the wild places they inhabit are not only important as a source of wonder and inspiration, but are an integral part of our natural resources and our future livelihood and well being.

- President Julius K. Nyerere, Arusha Manifesto, Tanzania, 1961

Wildlife species worldwide are at risk due to a suite of threats including habitat loss and fragmentation (Bender *et al*. 1998), overexploitation (Broad *et al*. 2001, Rosen and Smith 2010), infectious diseases (Aguirre and Tabor 2008), competition from invasive species (Clavero and García-Berthou 2005), and climate change (Mantyka-pringle *et al*. 2012). In many developing areas, including much of Sub-Saharan Africa, the main threats to wildlife persistence are habitat loss and fragmentation, which come as a result of growing human populations and an associated increase in the consumption of natural resources (Newmark 2008). Loss or fragmentation of habitat may lead to a loss of connectivity for wildlife, whereby populations that previously experienced gene flow become isolated from one another. Isolated populations may be more vulnerable to stochastic events such as drought, disease and fire (Shaffer 1981).

Isolation can have less obvious but equally detrimental impacts on the genetic viability of a population. As distance between patches increases, immigration from outside populations becomes less frequent (MacArthur and Wilson 1967). Small populations may quickly lose rare alleles through genetic drift, and mutation is unlikely to replace this diversity at the rate at which it is lost. Low genetic diversity can affect population persistence through an increase in inbreeding or the accumulation of deleterious alleles (Frankham *et al*. 2002).

Genetic diversity is a critically important resource that warrants conservation attention (Frankel and Soulé 1981). The International Union for the Conservation of Nature (IUCN) lists the preservation of genetic diversity as a necessity "in order to

maintain biological interactions, ecological processes and function" (IUCN 2002). However, in many cases, little to nothing is known about the basic genetic diversity of a species, let alone the population structure that exists within an area. This baseline information on variability and structure is important for management strategies, as it can identify populations that are at risk of inbreeding depression, or conversely, those sufficiently differentiated so as to warrant management as a distinct unit. In the face of climate change, where intraspecific variation may play a key role in the survival of a species in a changing environment, knowledge of genetic diversity and structure may help managers to better preserve the evolutionary potential of species at risk (Frankel and Soulé 1981; Walther *et al*. 2002). Genetic information can also identify populations that have historically interacted with high levels of gene flow, but that may become isolated due to increasing alteration of the landscape matrix between them.

With its legendary wealth of flora and fauna, Tanzania has repeatedly been shown to be a trove of biodiversity (Burgess *et al.* 1998, Myers *et al*. 2000, Robertson 2002, Pettorelli *et al*. 2010). The IUCN's critical goal of preserving genetic diversity is shared by the Tanzanian National Parks Authority (TANAPA), one of whose mandates is to "strive to protect the full range of genetic types native to plant and animal populations … by perpetuating natural evolutionary processes and minimizing human interference with evolving genetic diversity" (TANAPA 1994).

There is a long history of wildlife conservation in Tanzania; accordingly, some level of protective status has been conferred on over one-quarter of the country's area, through the gazetting of national parks, game and forest reserves, and wildlife management areas (Leader-Williams 2000). Unlike many other African countries, the boundaries of Tanzania's parks and reserves are not fenced. Historically this fluid system made it possible for individuals to disperse from their natal territory into new areas where they could augment existing populations, establish new populations, or recolonize previously occupied areas.

In addition to ecological benefits this commitment to conservation has great economic impact. Kweka *et al*. (2003) estimated that, in 1992, international tourism accounted for approximately 5.8% of Tanzania's GDP directly; in 2007 it climbed to almost 11% ($1.6 billion US) (Mitchell *et al*. 2009). However, the protection of wildlife resources for tourism comes at a cost. Revenue generated through tourism is rarely disseminated to local people, thus they often lose access to wildlife and land while seeing few benefits of conservation (Leader-Williams 2000). In 2001 an estimated 35.7% of people in Tanzania lived below the poverty line (World Bank, accessed 13 January 2011), and the national average income was less than $2 per day. As well, land targeted for new protected areas often occurs in areas close to or overlapping human settlements and that proximity can cause conflict with local peoples. Like many developing countries, Tanzania is facing increasing population and economic pressures, and human-wildlife conflict will continue to occur as human populations increase and there is competition for land. Human settlement and activity in areas surrounding protected areas creates the opportunity for human-wildlife conflict through crop-raiding, and can also result in a loss of connectivity for wildlife populations within the reserve (Newmark 1996, Graham *et al*. 2009). To restore connectivity, wildlife management strategies in these areas may benefit from the creation of linkages (tracts of land selected to facilitate movement of multiple species; Beier *et* al. 2008). But, for any hope of success, conservation activities must be founded in strong empirical evidence.

Linking wildlife populations is an increasingly important conservation goal worldwide. At the core of linkage design is the need to accurately identify the landscape attributes that affect gene flow between populations. However, the methods used to identify locations for linkages can greatly impact their effectiveness. The use of expert opinion to prioritize lands for conservation is a common practice, but expert opinion may fail to appreciate the nuances of habitat selection (for example, see Clevenger *et al.* 2002) and could result in linkage placement in inappropriate areas. Sawyer *et al*. (2011) demonstrated that areas highlighted as optimal linkages can vary

widely depending on habitat variables that are included in connectivity models, and reported that linkage design studies rarely validate their opinion models using empirical data.

In lieu of expert opinion, some studies use animal movement data (e.g., VHF and GPS telemetry) to investigate animal movements over large landscapes. Unfortunately, although GPS telemetry can provide invaluable information about the movement patterns of an individual, collar deployment is costly and studies often suffer small sample sizes (Hebblewhite and Haydon 2010). Data on gene flow are seldom captured because telemetry tracks daily habitat use, and dispersal followed by reproduction (i.e., gene flow) can be difficult to detect.

This issue is increasingly addressed by using genetic data to optimize GIS-based resistance surfaces that test hypotheses of the relative effects of landscape features (e.g., Cushman *et al*. 2006, Epps *et al*. 2007). In this framework, gene flow (inferred from the genetic structure of populations) is used as a measure of connectivity between populations. Landscape features are expressed as raster data (wherein the entire landscape is divided into pixels, and each pixel has a value). From these raster data, features are assigned weights that reflect the cost of traversal by the focal species, resulting in the creation of a 'resistance surface' (Cushman *et al*. 2006, Shah and McRae 2008). To quantify the landscape resistance to movement between two populations, cumulative resistance between two areas can be calculated (Shah and McRae 2008). Frequently, resistance surfaces have been parameterized by expert opinion or information on habitat use. However, by confronting hypotheses about landscape resistance (in the form of resistance models) with empirical evidence (e.g., genetic data), different resistance models can be compared to identify the one that most closely mirrors observed patterns of gene flow (Cushman *et al*. 2006). In this way, one can test the role of major landscape elements in shaping gene flow, and then optimize the model of the relationship by testing many different variants of the resistance weighting scheme (Shirk *et al*. 2010).

Although data on genetic structure provide an excellent way of empirically testing hypotheses of landscape effects on historical gene flow, there are some caveats to this method. Population structure is not necessarily indicative of levels of gene flow and can be generated by several other processes, for instance, colonization without subsequent gene flow between populations, or vicariance (Bohonak 1999). During the Pleistocene, many ungulate species may have survived in refugia that offered habitat in an increasingly arid landscape (Lorenzen *et al*. 2006, Lorenzen *et al*. 2010). In the absence of migration between subpopulations, population structure would not be indicative of levels of gene flow, but of colonization of these patches and, subsequently, genetic drift. Another process that could cause population structure is vicariance (the creation of a physical barrier which isolates subpopulations and prohibits gene flow) (Avise 2004). Vicariance can occur with tectonic plate uplifting or the creation of mountains or rivers that subdivide a previously intact population. Lastly, population structure can be caused by historically low levels of gene flow between subpopulations.

In this study, I gathered genetic data to investigate historical connectivity levels for three wide-ranging ungulate species across southern and central Tanzania: Maasai giraffe (*Giraffa camelopardalis tippelskirchi*), common impala (*Aepyceros melampus melampus*), and common eland (*Tragelaphus oryx*, formerly *Taurotragus oryx*). I used these genetic data to contrast models of landscape resistance and investigate the relative effects of different landscape features on gene flow for these species.

The three focal species of this study were historically distributed across much of Tanzania (Tanzania Mammal Atlas Project, accessed 15 February 2012), although the largest populations persist in reserves at present time (Newmark 2008). Large population sizes and long generation time (as seen in these large herbivores) can result in a time lag before changes in population structure are seen, thus it is important to note that current patterns of genetic structure will represent effects of historical processes (i.e., "*historical connectivity*"), not estimates of recent gene flow (Anderson

et al. 2010; Cushman and Landguth 2010). Because much of the expansion in human settlement and agriculture has occurred within the last 80 years it is unlikely that the impacts of these activities on population structure will be detectable at this time (Holzhauer *et al*. 2006). I believe that historically, species-specific habitat requirements dictated the broad patterns of population structure and gene flow for many ungulate species in south and central Tanzania. Thus, population structure will be indicative of how much gene flow occurred between these populations in the past, not colonization of refugia and drift. Population structure is commonly reported using the fixation metric F_{ST} (Wright 1921), which ranges from 0 to 1 and quantifies the proportion of the total variation observed in the subpopulations, relative to the variation in the total population. Larger F_{ST} values suggest greater differentiation between populations (Hartl and Clark 1997).

Like genetic structure, a detectability of a change in genetic diversity is also subject to a time lag. This time lag depends on the effective size of the population being studied, and in populations with moderate to large numbers of individuals the resultant decrease in genetic diversity may not be evident for multiple generations after the event occurs (Cushman and Landguth 2010). The effective population size (N_e) is the number of individuals in an idealized population that shows the same extent of genetic drift as the sampled population (Hartl and Clark 1997). Most natural populations do not fit the assumptions of an idealized population (i.e., nonoverlapping generations, equal population sizes, and an absence of migration, mutation, and selection); however, because N_e is closely linked to measures of diversity and distance, it is a more informative measure of the genetic health of a population than census size (N_c, or commonly N). The relationship between F_{ST} and N_e is given by the equation

$$F_{ST} = 1 - e^{t/2Ne} \quad \text{[Wright 1943]}$$

where t is the time in generations. Thus, with smaller effective population sizes, F_{ST} increases more rapidly. Since many of the populations of ungulates in Tanzanian reserves are in the thousands or tens of thousands (Caro 2008, Waltert *et al*. 2008), recent changes in connectivity may take many generations to become detectable.

However, by conducting estimates of population structure now, we can elucidate what the patterns of connectivity were before human activities increased on this landscape.

The purposes of this study are fourfold. First, I estimate genetic diversity and population structure of three ungulate species spanning a wide body-size gradient in a large and understudied region of Tanzania. Secondly, I use spatially-explicit GIS models to investigate the effects of landscape features on gene flow for each species and test hypotheses about the historical processes that may have shaped the current genetic structure of these populations. Thirdly, I use these analyses to contrast the scales at which gene flow (and thus dispersal) has occurred, and the relative impact of landscape features on gene flow for species with varying dispersal capabilities. I predict that genetic diversity within reserves will be highest for smaller-bodied animals, due to larger population sizes. Lastly, I identify areas where trends of increasing human activity may sever or limit gene flow. At this time, it is critical to identify wildlife populations that are most likely to suffer a loss of connectivity, while minimizing the impact of land conservation on human populations. These results will allow us to focus the broad discussion of connectivity to a scale and location at which it is biologically relevant for these species.

MATERIALS AND METHODS

Description of study area

Tanzania is renowned for the size and number of its national parks, as well as the diverse community of large mammals and endemic species that they possess (Newmark 1996, East 1998). At the highest level of protection, national parks cover approximately 5% of the country's landmass. Within national parks hunting, resource extraction, livestock grazing and settlement are prohibited; parks are maintained solely for conservation and tourism and human activity is strictly regulated (TANAPA 1994). In contrast, game reserves allow selective harvest of certain species in 'trophy' hunts, and the Tanzanian Wildlife Division sets quotas for each hunting block. Additional areas are protected to a lesser degree than parks and game reserves. For instance, forest reserves offer limited resource extraction opportunities (e.g., honey-collecting, firewood and timber) for local peoples. Wildlife management areas are another category of land use established as a collaborative conservation effort between local villages and the Wildlife Division that allows limited human settlement while managing off-take of wildlife (Hurt and Ravn 2000). Game and forest reserves in Tanzania often abut national parks, creating 'reserve complexes' that provide large tracts of contiguous habitat for wildlife.

Variations in soil type and levels of annual precipitation have created several different ecosystems across the Tanzanian landscape (McNaughton and Georgiadis 1986, Caro *et al*. 1998). In the northern and eastern regions, fertile volcanic soils and two annual rainy seasons ('long' rains from March through May and highly variable 'short' rains from October through December) contribute to the growth of *Acacia-Commiphora* woodlands and bush cover (Leader-Williams 2000). Serengeti NP, Ngorongoro Conservation Area and Tarangire NP fall within the area of bimodal rains; Tarangire NP receives an average of 529mm of rainfall per year and is approximately 1200m above sea level (Peterson 1978 *in* Borner 1985).

However, in southern and western areas, nutrient-poor soil and a single annual rainy season (October through May) promote the growth of *Brachystegia-*

Julbernardia miombo woodlands (Leader-Williams 2000). The Ruaha NP complex, Selous GR, Mikumi NP, and Katavi NP complex all fall within the region of unimodal rains. Katavi NP is located at 800-1100m above sea level and receives an average of 923mm of rainfall each year. The eastern arm of the Selous GR receives approximately 760mm per year (Rodgers 1974) and ranges from 100-750m in elevation; it consists mainly of woodlands, *Combretum* thickets, and wooded savannah. Ruaha NP is a mixture of *Brachystegia miombo* on the west, *Commiphora-Combretum* shrublands on the east, and *Acacia* shrublands in the northeast. Ruaha NP receives approximately 580mm of rainfall annually (Barnes and Douglas-Hamilton 1982).

While raising cattle is an important source of sustenance for pastoral tribes in northern areas, proliferation of the tsetse fly (*Glossina sp.*) in southern regions precludes this activity because the insect is a vector for trypanosomiasis, or 'sleeping sickness' (Leader-Williams 2000). Indeed, sections of both the Selous GR and Katavi NP (which both lie within *miombo* and the tsetse fly range) were gazetted after thousands of villagers were removed and relocated as part of trypanosomiasis-control measures (Leader-Williams 2000, Mulder *et al.* 2007). *Miombo* woodlands generally harbour lower human population densities because agriculture is limited by poor soil quality and prolonged droughts (WWF 2001).

African ecosystems are greatly influenced by the distribution and amount of rainfall. Recent studies suggest that climate change may result in prolonged drought conditions in Africa, which may in turn influence available habitat and the composition and range of sub-Saharan species (Zinyowera *et al*. 1998, Low 2005). Ungulate species vary in their dependence on water: some are able to obtain stored water through browse, while others are reliant on free water sources. However, if recent changes in precipitation are any indication of future water availability, water sources will become much more difficult to find. Zinyowera *et al*. (1998) caution that, although climate forecast models suggest a small increase in rainfall over much of Africa, this benefit will be negated as higher temperatures (approximately 1.4 degrees

Celsius by 2050) will also cause increased evapotranspiration. Dry season droughts may increase in severity, causing changes in vegetative communities; the savannah grasslands of East Africa may be particularly vulnerable to these changes (Zinyowera *et al*. 1998). Vanacker *et al*. (2005) suggested that even short-term variation in precipitation levels can impact vegetation cover, and that grassland and shrub savannahs are particularly sensitive to the timing and duration of rainfall events. In turn, alterations in vegetation type and cover can affect species distribution or communities, survival rates, and migration patterns. Thus, if predictions of wetter rainy seasons and more severe droughts in East Africa are realized, then current delineations of national parks may be unable to sustain the wildlife they were created to protect.

Much of the land that is currently gazetted as reserves originally supported human settlement at varying (but probably low) densities; many people were relocated under health-related or social programs, or when park boundaries were established. As a result, land adjacent to the parks often was more densely settled and human population densities increased in these areas by 5-15% (Newmark 1996). Tanzania's population was estimated at 43.7 million in 2009, with an annual growth rate of 2.5% (World Bank, accessed 13 January 2011). Estes *et al*. (2012) estimated that human population growth rate and the conversion of habitat to agriculture around Serengeti NP were higher than the national average, due to a paucity of arable land elsewhere. Thus, the insularization of wildlife populations in reserves is likely to continue to accelerate as the nation's human population increases.

Several of the smaller protected areas in Tanzania have already suffered species losses consistent with the predictions of island biogeography theory (Newmark 1996). Many wildlife species in Tanzania are also at risk due to poaching for bushmeat (Caro 2008), collection for the wildlife trade (Roe *et al*. 2002), diseases such as rinderpest and anthrax (Prins and Weyerhaeuser 1987) and canine distemper (Munson *et* al. 2008), as well as increased land conversion for agriculture (Newmark 2008).

Focal species

In general, ungulates are excellent species to use for population connectivity analyses for several reasons: first, they can be very abundant and conspicuous on a landscape. Unlike carnivores, ungulates tend to be diurnal and gregarious, and many species gather in large numbers in riverine areas during the dry season. Thus, collection of faecal samples from many individuals is often relatively easy. Secondly, they may also serve as 'focal species' for conservation: present-day surveys of village lands between Ruaha NP and Mikumi NP have shown that the presence of large ungulates (specifically elephants) is correlated with species richness in an area (Epps *et al*. 2011). This suggests that identifying and preserving historical travel routes for large herbivores such as elephants may benefit many other species.

In this study I used non-invasive sampling methods to collect genetic samples and generate microsatellite genotypes for three wide-ranging ungulate species: Maasai giraffe, common eland, and common impala. These species were chosen to cover a spectrum of body masses, dietary and water requirements, habitat preferences, home range sizes, and social systems (Table 1). Each of these factors could potentially have a pronounced impact on the dispersal capability (and thus the potential for gene flow) of a species. Although the majority of studies investigating the effect of landscape attributes on gene flow focus on only one species, this multi-species approach will allow comparisons across species with varying habitat requirements and levels of vagility. The purpose of this study is to help inform efforts to conserve connectivity in Tanzania, where managers are challenged with protecting an extremely diverse array of species.

Giraffe

Male giraffe range in size from 800-1930 kg, while females typically weigh between 550-1180 kg (Estes 1992). Although they have been known to eat over 100 different tree species they preferentially forage on *Acacia*, *Commiphora*, and *Terminalia* trees. Due to an increased amount of stomach surface area, giraffe are extremely efficient browsers and do not require as much food as other ungulates,

though they tend to choose higher quality material (Kingdon 1997). Giraffe obtain most of their moisture through leaves, and thus are not reliant on permanent water sources, however, they will drink during the rainy season.

Although gregarious, giraffe are assumed to have a very fluid social system with no clear hierarchical structure (Estes 1992). Herds of up to 50 individuals may collect in an area, but they form no lasting associations. Individuals of both sexes and all ages may be present in these herds at any time. Female giraffes remain close to their natal site, while males form bachelor herds and emigrate (van der Jeugd and Prins 2000). Giraffe are non-territorial, however, during the breeding season bulls may attempt to monopolize the females in an area by driving off other males. Giraffes have been shown to have home ranges that vary from 5-600km^2 (Estes 1992); the average within Tsavo NP was approximately 160 km^2 (Leuthold and Leuthold 1978), but only 5.2-8.6 km^2 in Lake Manyara NP (van der Jeugd and Prins 2000). Van der Jeugd and Prins (2000) postulated that variation in habitat quality was responsible for the differences observed between these two studies. Savannahs, woodlands and floodplains provide suitable habitat, though Ginnett and Demment (1999) found that bulls and cows without calves preferred densely wooded areas, while cows with calves were more often found in floodplains and mixed-shrub locations with higher visibility. Giraffe may even be found at elevations of up to 2000m if the terrain is not overly steep (Kingdon 1997). Like many other ungulate species they concentrate in riverine areas during the dry season, but disperse to areas with better forage during the rainy season (Estes 1992).

Brown *et al*. (2007) analysed population structure of giraffes across their range in Africa using microsatellite and mitochondrial DNA analysis and suggested that there are at least six genetically distinct species, with the Maasai species (*Giraffa tippelskirchi*) found in Tanzania and southern Kenya. However, the taxonomic distinction of giraffes is still under debate, thus I will refer to Maasai giraffe by the more widely accepted subspecies denomination (*Giraffa camelopardalis tippelskirchi*). Although once common throughout much of Africa, giraffe populations

in western Africa have declined and the species is mostly restricted to national parks and reserves in East Africa. Giraffe are listed as a species of 'least concern' on the IUCN red list (IUCN 2011) but overall trends show a decrease in abundance and a range reduction due to increased anthropogenic pressure (Fennessy and Brown 2010); the Maasai subspecies was not assessed individually. Major threats to giraffe persistence include habitat fragmentation and poaching (Fennessy and Brown 2010).

Brown *et al*. (2007) estimated divergence times of the six giraffe lineages and suggested that the Maasai giraffe was initially part of a southern clade that also included the Angolan (*G. c. angolensis*) and South African (*G. c. giraffa*) subspecies, but that it diverged 0.13 to 0.37 Ma, during the mid-Pleistocene. They propose that fluctuations in aridity and temperature fragmented habitat, causing allopatric speciation. Although several giraffe subspecies have ranges that do not appear to be separated by biogeographic barriers (e.g., Maasai and reticulated, *G. c. reticulata*), Brown *et al*. (2007) found little evidence of gene flow and suggest that the subspecies are maintained by reproductive isolation due to pelage appearance or differences in mating season.

Impala

Impala are medium-sized (45-60 kg), gregarious antelope common to many Tanzanian parks and reserves. There are two recognized subspecies of impala: the common impala (*Aepyceros melampus melampus*) which ranges from Kenya to South Africa, and the black-faced impala (*A. m. petersi*) of northwest Namibia and southwest Angola. Common impala are listed as a species of least concern on the IUCN red list (IUCN 2011) and most populations are stable or increasing (IUCN SSC Antelope Specialist Group 2008a).

Impala can exploit a wide variety of food sources: during the wet season, they preferentially graze on fresh grass, but switch to browse in the dry season. As a result of this dietary generalism, impala are sedentary and have small home ranges relative to other antelopes (1.3-5.8 km^2, Kingdon 1997). They are found in both woodland and savannah ecosystems. However, impala are dependent on a permanent source of free

water and prefer habitat with modest slope at most; thus although widely distributed across sub-Saharan Africa they are often found in patchy distribution at a smaller scale (Estes 1992). Where they do occur, impala congregate in high densities in clans of females and juveniles. Adult males are territorial and use threatening displays to defend a small area during the breeding season, while subadult males emigrate into neighbouring clans. Using mark-recapture methods, Murray (1982) observed dispersal by 26 subadult male impala and estimated that the average distance emigrated was 1.2 km (range 0.4-3.2km), which placed the individual outside the range of its natal clan.

Lorenzen *et al*. (2006) used microsatellites to analyse genetic structure of common impala across their range, and found a significant difference between samples from southern countries (Botswana, Zimbabwe and Zambia) and East Africa (Uganda and Kenya). Individuals from Tanzanian reserves were genetically 'intermediate' and were assigned with less certainty to both the eastern and southern clusters. Upon observing higher levels of genetic diversity in southern populations, Lorenzen *et al*. (2006) suggested that impala may have colonized East Africa from South Africa, and that the current genetic structure within the species is a result of isolation by distance.

Eland

Eland are large (*ca*. 300-600 kg for females, 400-942 kg for males), ox-like, sexually size-dimorphic members of the tribe Tragelaphini (Estes 1992). There are two recognized species, the giant eland (*Tragelaphus derbianus*, formerly *Taurotragus derbianus*) which ranges from west to central Africa, and the common eland, *T. oryx*, whose range extends from Kenya to southern Africa.

Eland are highly gregarious and form groups larger than most other bovids (herds of >100 individuals are not uncommon). However, the social structure of herds is quite dynamic, and individuals form no lasting associations. Young cow eland are often highly nomadic and non-territorial, with reported home ranges of up to 1500 km^2. Males, especially mature bulls, tend to be more sedentary and may form small

bachelor groups or remain independent. Group size tends to increase during the rainy season, when females with young congregate to forage on newly sprouting grasses and benefit from increased vigilance against predators (Kingdon 1997). Eland generally inhabit areas of light woodland and woodland-savannah but are considered habitat generalists and are the only tragelaphines known to utilize open savannah (Estes 1992). Eland have also been found at elevations of up to 4,900 m (Hillman 1974). They are generalist browsers and can consume lower quality foliage and herbs than other ungulate species (Kingdon 1997), but may also consume new grasses during the rainy season. Eland are thought to be less reliant on permanent water sources than other species, as they are able to obtain moisture from their food (Thouless 2008). Because they are vulnerable to rinderpest and the loss of habitat to agriculture, as well as being considered a high quality species for bushmeat, *T. oryx* has disappeared over parts of its range (Kingdon 1997). Although eland are listed as a species of least concern on the IUCN red list (IUCN 2011), that listing has the caveat that habitat loss may adversely affect wild populations and species persistence will increasingly rely on populations in private facilities such as the intensively-managed, fenced private game reserves common to southern Africa (IUCN SSC Antelope Specialist Group 2008b).

Lorenzen *et al*. (2010) analysed mitochondrial DNA sequences for common eland across their range and found a more recent common ancestor for populations in East Africa (0.21 Ma) than in South Africa (0.35Ma). East African reserves also showed lower genetic diversity and more structure, suggesting that these populations were colonized from older populations in the south, possibly after localized extinctions during the Pleistocene.

Sample collection

From June-December 2009 I collected noninvasive (dung and carcass) samples from eight national parks and game reserves across Tanzania (Figure 1, Table 2):

Katavi National Park (KNP)
Mikumi National Park (MNP)
Ruaha National Park (RNP)

Rukwa Game Reserve (RkGR)
Rungwa Game Reserve (RgGR)
Selous Game Reserve (SeGR)
Swagaswaga Game Reserve (SwGR)
Tarangire National Park (TNP)

Within this thesis, I use *reserve* as a general term to refer to a protected area of any classification (i.e., game reserve, national park, forest reserve). I use the terms *national park* and *game reserve* when the distinction is important.

These eight reserves were chosen to complement limited sampling done by Clinton Epps in 2006-2007 at Mikumi NP, Ruaha NP, and the Idodi-Pawaga Management Area, as well as village lands (VILL) located in Morogoro, Dodoma, and Iringa regions between Mikumi NP and Ruaha NP (Figure 1). Surveys of current wildlife connectivity suggested that the major concentrations of these three species in southern Tanzania are largely restricted to protected areas (Epps *et al*. 2011); hence I did not do any additional sampling of unprotected areas intervening the reserves.

Although all three of these species undergo seasonal movements to exploit resources available at different times of the year, their movements are relatively localized and they do not make the large-scale annual migrations seen in wildebeest (*Connochaetes taurinus*) and plains zebra (*Equus quagga*). Although technically possible, long-distance dispersal between the reserve complexes in this system is expected to be relatively rare due to the scale of the study area. I considered samples from each reserve to be an independent population for genetic analysis because sampling areas were relatively small and discrete within each reserve. Thus, for this study I considered any individual sampled within a reserve to be representative of that population's allelic frequencies (i.e., not an immigrant from a neighbouring reserve).

The eight reserves I sampled are isolated from one another to varying degrees by geographic distance, as well as possible biogeographic, topographical, and anthropogenic features. As such, they provide comparisons as to the relative effect of each of these features on genetic structure. They also differ widely in size and time since establishment (Table 2). Some sites (e.g., Ruaha NP and Rungwa GR, or Katavi

NP and Rukwa GR) occur within a 'complex' of protected areas (i.e., adjacent reserves that offer contiguous habitat, though often with varying protective status; see Figure 1). Although all three species did not co-occur in all areas, the overall distribution of sampling locations was similar for the three species and was designed to allow comparative analyses of genetic variation, genetic structure, and the effects of intervening landscape attributes on different species.

I located animals via diurnal driving transects on improved and unimproved roads in each reserve. I attempted to collect samples across a wide geographic area within each reserve to observe as much of the inherent genetic variation as possible, and also to prevent sampling the same individual multiple times. Swagaswaga GR was much smaller than the other reserves and only recently gazetted, so it had few roads. Thus, in this reserve animals were located using walking surveys in addition to driving. When possible, I confirmed species identity and sex of the animal by observing pellet deposition. If pellets were found without observing the animal, I attempted to locate tracks to minimize the possibility of misidentification, but also used traits such as the size and shape, odour, and composition of the pellets (Stuart and Stuart 2000).

I chose fresh samples where available, but also collected recently dried samples; older pellets that appeared faded or cracked or had been rained on were not collected because DNA quality was assumed to degrade rapidly as pellets weathered (Murphy *et al*. 2007, Brinkman *et al*. 2010). Each pellet pile was assumed to have originated from a single individual; pellets were collected using twigs and a 'chopstick' method to prevent contamination between samples or from the person collecting. Dung samples were placed in paper envelopes and dried thoroughly within two days of collection to prevent DNA degradation and the proliferation of bacteria and fungi (Wehausen *et al*. 2004). I also collected tissue samples from carcasses in the field or hunter trophies in game reserves. Tissue samples were also stored in paper envelopes and dried thoroughly at ambient temperature before extraction.

For each sample I recorded a point location using a Garmin Map60cs GPS linked to a Hewlett-Packard IPAQ hand-held computer. I designed a field survey database in Program CYBERTRACKER (http://www.cybertracker.org/) to collect data on date and time, estimated sample age, sample type (dung or carcass), sample condition, species, and certainty of species identity. In the laboratory, samples were stored at ambient temperature in a sealed bucket with desiccant until processed.

Laboratory techniques

Sample processing and extraction

To obtain epithelial cells from dung for DNA extraction, I scraped the outer surfaces of pellets with a razor blade (Wehausen *et al*. 2004). Total genomic DNA was extracted from scrapings using a modified protocol (Appendix 1) of the AquaGenomic stool and soil extraction kit (Multitarget Pharmaceuticals, Salt Lake City, UT, USA).

I extracted genomic DNA from tissue samples using a DNEasy Tissue Kit™ (Qiagen, Valencia, CA, USA) with a minimum of 3 hours of incubation for lysing. Particularly tough tissue samples were lysed overnight at 60ºC. After extraction, I precipitated the DNA for transport to the USA (Oregon State University), where I rehydrated it with 200µL of 1x TE buffer and stored it at –80ºC. I conducted all sample scraping and extraction in a dedicated low-DNA area where no PCR product had been handled and used barrier tips to minimize contamination between samples.

Microsatellite genotyping

Microsatellites are biparentally inherited tandem repeats of 1-6 base pair length, found in non-coding nuclear DNA. As a result of base pair 'slippage' and copying errors during DNA replication, they can accumulate a large amount of variation over a relatively short period of evolutionary time. Because these replication errors occur in non-coding DNA, this variation occurs without detrimental effects to the organism, and is believed not to be affected by natural selection (Frankham *et al*. 2002). Although mutation rates can vary by species, locus, sex, age, and repeat

length, the average mutation frequencies suggested for microsatellites (~10^{-4} mutations per locus, per generation) are much higher than for the whole genome in eukaryotes (~10^{-9} mutations per nucleotide, per generation, Ellegren 2000). Microsatellite loci typically exhibit differences in allelic frequencies from population to population, and for this reason they are useful for examining patterns of fine-scale variation across a species (Bruford *et al*. 1996).

Where possible, I screened species-specific microsatellite loci (e.g., Huebinger *et al*. 2002, Huebinger *et al*. 2006), or markers that had previously been shown to amplify in our study species (Lorenzen *et al*. 2006). I tested bovine markers for amplification in eland (see Table 4c). I used the software Primer3 (Rozen and Skaletsky 2000) to redesign primers from Huebinger *et al*. (2002) that did not amplify cleanly. I identified 10 polymorphic loci for eland, 6 for giraffe and 11 for impala that consistently amplified and were scoreable (Table 4).

To complement published giraffe primers, I collaborated with Tom Mullins (Forest and Rangeland Ecosystem Science Center, USGS), to conduct an Illumina sequencing run (see Jennings *et al*. 2011 for methods) on a Maasai giraffe tissue sample. I used Primer3 (Rozen and Skaletsky 2000) to identify primer sites for microsatellites within from these sequences. From this process I tested 36 new primer pairs for amplification and polymorphism using fluorescein-labelled dCTP (Perkin-Elmer, Boston, MA, USA). I screened eight samples from across multiple reserves to verify polymorphism at each new locus, and retained loci for population genetic analysis if they exhibited two or more alleles across the range of eight samples and had clean peak profiles. I identified nine novel polymorphic loci (Table 4a); all new loci consist of dinucleotide repeat motifs. These nine loci were subsequently labelled with fluorescent dyes and reanalyzed for all giraffe samples in each population. Giraffe samples were thus genotyped at fifteen loci (six previously published, nine designed for this study).

For all samples I used the polymerase chain reaction (PCR) to amplify microsatellite loci in 15μL reactions consisting of 1x magnesium-free PCR buffer,

2.25mM $MgCl_2$, 160µM of each dNTP, 7.5µg bovine serum albumin (New England Biolabs, Ipswich, MA, USA), 0.1µM of each primer, 0.7*U* Apex Hot Start *Taq* polymerase (Genesee Scientific, San Diego, CA, USA), and 0.6µL of genomic DNA. Thermalcycling conditions were as follows: an initial denaturation of 15 minutes at 95ºC, followed by 40 cycles of: denaturation at 95ºC for 30s, 45s of annealing at a locus-specific temperature (see Table 4), and extension at 72ºC for 30s. The PCR had a final elongation step of five minutes at 72ºC. For each primer pair one primer was fluorescently tagged on the 5' end with NED, PET, VIC (ABI, Carlsbad, CA, USA) or 6-FAM (Sigma-Aldrich, St. Louis, MO, USA).

I assessed amplification quality by electrophoresing 4µL of the PCR product on a 2% agarose gel prestained with GelRed (Biotium, Hayward, CA, USA). Samples were diluted as necessary, multiplexed with 2-4 other loci, and then processed on an ABI3730 DNA analyser at the Center for Genome Research and Biocomputing (Oregon State University), with LIZ500 sizing standard.

Following the 'multiple tubes approach' (Navidi *et al.* 1992, Taberlet *et al.* 1996), I amplified and scored all samples a minimum of three times to obtain consensus genotypes and minimize error due to allelic dropout. For an allele to be verified it must have appeared in two separate replicates; samples which did not generate reliable consensus genotypes or that appeared aberrant were reamplified in an additional three replicates. I included one sample with a known genotype in all ABI3730 runs as a positive control to standardize allele size, and also included a negative control in each run to detect contamination. I used Genemapper 4.1 to verify allele sizes and generate genotypes. Because eland and giraffe pellets can be similar size and shape, for samples of uncertain species origin I distinguished between giraffe and eland using a diagnostic panel of microsatellite loci that either consistently failed to amplify in giraffe but worked in eland (TCRBV62) or were fixed for alleles in giraffe that were not seen in eland (MMP9, OarFCB193, and OarAE129).

Genetic analyses

Defining sample populations

Within each protected area I combined all samples from each species into a 'population' for genetic analysis. Although some protected areas are found within a large protected complex (e.g., Ruaha NP and Rungwa GR in the greater Ruaha ecosystem), I analysed each reserve as a separate population where sample sizes permitted. This allowed me to test genetic differentiation of populations separated by tracts of protected land versus unprotected habitats; it also increased the number of comparisons possible and my power to detect an effect if one was present. For eland, however, sample sizes in Katavi NP and Rukwa GR were too small to be analysed separately so I combined the two; neither reserve showed population-specific alleles.

I chose to do population-based analyses instead of individual-based because my sampling scheme was concentrated on roads and riverine areas, and did not uniformly sample across the landscape within each reserve (Cushman *et al.* 2006). Because I collected samples in the dry season when animals congregate along riverine areas, inter-individual geographic distances are probably not representative of year-round spatial partitioning of genetic diversity. On a larger scale, by comparing genetic structure and diversity between protected areas, my overall sampling distribution was necessarily clumped and was better suited for a population-based approach (Cushman *et al*. 2006).

Identifying unique individuals

After combining allele calls across all runs into one consensus genotype per individual, I used GENALEX 6.41 (Peakall and Smouse 2006) to calculate global allelic frequencies and determine the probability of identification ($P_{IDunr,}$ the probability that two unrelated individuals have the same genotype) and probability of identification for siblings (P_{IDsib}) for each population of each species. I used $P_{IDunr} = 0.001$ and $P_{IDsib} = 0.05$ as the acceptable maximum probability that two unique individuals would have an identical genotype. I conducted multiple trials in which locus order was randomized to determine the mean number of loci required to reach suitable P_{ID} values

for each species (GIR: n = 6 loci; IMP: n = 6 loci, ELAND: n = 6 loci). I then removed individuals that had amplified at fewer than this number of loci.

I used CERVUS 3.0.3 (Marshall *et al.* 1998, Kalinowski *et al.* 2007) to identify duplicate genotypes. I incrementally decreased the stringency of matching conditions (i.e., requiring fewer matching loci for two individuals to be considered a duplicate) until the program began returning pairs of samples that were highly unlikely to be duplicates due to vast geographic distances or multiple mismatches at different loci. When likely duplicates were identified, I removed whichever sample had a less complete genotype than the other.

Linkage disequilibrium and Hardy-Weinberg equilibrium

I used a Markov chain Monte Carlo (McMC) approximation in GENEPOP v.4 (Raymond and Rousset 1995) to test for evidence of linkage disequilibrium between all pairs of loci across all populations for each species. I also performed exact tests for departures from Hardy-Weinberg equilibrium (HWE) in GENEPOP, with a dememorization of 10,000, and 1000 batches of 10,000 iterations each. I estimated 95% confidence intervals for null allele frequencies (Pemberton *et al*. 1995) for each locus in each population in GENEPOP. For each of these comparisons (LD, HWE, null alleles) I applied a Bonferroni adjustment for multiple comparisons (Rice 1989) to maintain an overall significance level of 0.05 while conducting multiple tests on the data set. I quantified within-population allelic diversity (number of alleles, A), observed heterozygosity (H_O), and expected heterozygosity (H_E) for each reserve with GENALEX. I averaged H_E across all loci for each population.

Genetic diversity and distances

I used a rarefaction procedure in HP-RARE 1.1 (Kalinowski 2005) to compare allelic richness (A_R) between reserves while accounting for uneven sample sizes. In the rarefaction procedure, I subsampled all populations to equal the minimum number of alleles observed at any locus in any population (giraffe, 14 alleles; impala, 10 alleles; eland, 8 alleles).

I calculated population pairwise genetic differentiation (Weir and Cockerham's θ_{ST} analogue of F_{ST}, hereafter referred to as F_{ST}; Weir and Cockerham 1984) values using GENEPOP, with a dememorization of 10,000, and 1,000 batches of 10,000 iterations each. F_{ST} measures the degree of divergence in allele frequencies between subpopulations. Although F_{ST} is widely reported in genetic studies, it has been shown to inadequately describe interpopulation differentiation derived from microsatellite genotypes in some circumstances, in part because F_{ST} was derived for use with diallelic loci (Hedrick 1999). F_{ST} scales from 0 to 1 in systems with two alleles, but higher numbers of alleles at a locus reduce the maximum F_{ST} values that can be observed (Hedrick 1999). There are also unrealistic assumptions inherent in F_{ST} metrics, including large and equal population sizes between all populations for all generations (Weir and Cockerham 1984). F_{ST} values describe the level of fixation of allelic frequencies across all populations of a species (not just those sampled).

Although F_{ST} values are commonly reported in population genetic studies, this metric is not appropriate for interspecies comparisons because different species will have different evolutionary histories and will have violated the assumptions in different ways (Nei 1986). As an alternative, in populations with moderate to high average allelic richness, Nei's standard (Nei 1972) and unbiased (Nei 1978) genetic distances (D and D_{hat}, respectively) provide a more suitable measure of differentiation because they are not sensitive to allelic richness and are calculated based only on populations sampled (not the entire suite of populations). Nei's D is linear with time since divergence, thus it is suitable for use in systems where genetic drift and mutation are the main drivers of differentiation (Weir 1996). Although mutation is likely occurring in this system, I expect genetic drift to have a much larger impact on allele frequencies than mutation in these populations, because large mammals typically have relatively low effective population sizes. For these reasons I used Nei's unbiased genetic distance for inter-species comparisons of gene flow, as calculated by GENALEX. I chose to use the unbiased version of Nei's standard genetic distance distances (D_{hat}) because it includes a correction for small sample sizes (Nei 1978).

Effective population size

Because the effects of isolation and drift are dependent on effective population size (N_e), I used approximate Bayesian computation in ONeSAMP (Tallmon *et al*. 2008) to estimate N_e for each population of each species. From the observed genotypes for each population, ONeSAMP calculates eight summary statistics that are a function of N_e, including the number of alleles per locus, expected heterozygosity, and the number of alleles per locus divided by the range of allele sizes at that locus (Tallmon *et al*. 2008). It then simulates 50,000 populations and compares these simulated populations to the observed population. ONeSAMP requires that all loci be polymorphic within a population, thus I removed loci from populations in which they showed only one allele. ONeSAMP is also sensitive to missing data, and so I removed individuals that did not have genotypes at two or more loci. Lastly, ONeSAMP requires a prior minimum and maximum bound for the estimate of N_e. Census sizes for Katavi NP/Rukwa GR and Ruaha NP/Rungwa GR respectively were in the thousands for giraffe and eland, and in the tens of thousands for impala (Waltert *et al*. 2008, Barnes and Douglas-Hamilton 1982). Frankham (1995) suggested that N_e is typically only one-tenth the census size for wild populations. Therefore, I set these priors at 2 – 1000 for giraffe, 2 – 1500 for eland, and 2 – 5000 for impala.

Interspecies comparisons

To compare patterns of genetic diversity across species I calculated the correlation (Pearson's correlation coefficient, r) of corrected allelic richness for each pair of species in the reserves in which I had samples for both species (i.e., giraffe and impala were compared in seven populations; giraffe and eland and impala and eland were compared over the three populations where eland were sampled). I also calculated the correlation of matrices of population-pairwise genetic distances (D_{hat}) between each pair of species using a Mantel test (Mantel 1967) with 10,000 permutations.

Soulé *et al*. (1979) suggested that, in the absence of connectivity, East African reserves will function as islands and lose genetic diversity, and that smaller reserves

will be affected more rapidly because of small population and effective population sizes. I tested the hypothesis that larger reserves host higher levels of genetic variation (Heller *et al*. 2010) by regressing allelic richness (A_R) on reserve size [measured as $\log_{10}(km^2)$] and sampling polygon size [$\log_{10}(km^2)$] for each species using XLSTAT, version 2011.4.02 (Addinsoft, New York, USA). I tested the relationship with total reserve size and also with sampling polygon size, because in some cases (e.g., Selous GR) the sampled area was considerably smaller than the reserve area (Table 3).

As a relative index of gene flow among populations for each species, I also calculated the approximate number of migrants per generation (*Nm*) from the F_{ST} values, using the equation:

$Nm = (1-F_{ST}) / (4* F_{ST})$ [Wright 1921]

This approach to estimating *Nm* has assumptions that likely are violated in many systems, including the assumption of equal migration between all populations, and mutation-drift equilibrium (Whitlock and McCauley 1999). However, *Nm* estimates can provide insight into relative rates of connectivity if used for comparison purposes among many populations instead of being interpreted as absolute migration values (e.g., Epps *et al*. 2005).

Evaluating population structure with assignment tests

For frequency analyses I considered all samples collected within a particular national park or game reserve to be a population. However, I also wanted to investigate structure within each species without imposing predefined population clusters. I used the Bayesian clustering program STRUCTURE 2.3.3 (Pritchard *et al*. 2000), which requires no prior information regarding sampling location. STRUCTURE investigates population structure by grouping genetically similar individuals into the number of clusters that minimize deviations from HWE and linkage disequilibrium. I set an initial burn-in period of 100,000 iterations (to allow convergence of the Markov chains), with 500,000 iterations thereafter. I ran all samples from each species separately, with a range of K = 1-11 clusters, with 10 replicates per cluster value (K). I assumed population admixture was present and that allelic frequencies between

genetic clusters were correlated. I used Evanno *et al.*'s (2005) ad hoc Δ*K* statistic to infer the most plausible number of genetic clusters. However, because this statistic is based on the second-order rate of change of the log likelihood of the number of clusters, it cannot be estimated for a *K* value of 1, thus I also plotted *K* versus the average log likelihood for all replicates of each *K* value. Once the highest order of population structure was identified, I re-ran samples in each of those clusters, to detect any hierarchical structure patterns. These additional analyses were done with the same burn-in and number of iterations as the initial runs, but with a range of *K* = 1-8 clusters.

GIS-Based landscape analyses

To investigate landscape effects on connectivity I created models of landscape resistance (resistance surfaces) between populations. I summarized cumulative resistance between populations for each hypothesis of landscape resistance and tested those values against genetic data. Details for this process are as follows.

Protected areas, habitat layers, and sampling polygons

I delineated boundaries of Tanzanian protected areas in a Geographic Information System (GIS; ArcGIS 9.3.1, ESRI, Redlands, CA, USA; used for all subsequent spatial analyses) using publically available data sets (UNEP-WCMC 2007, FAO 2005), but corrected obvious errors and omissions. I used a digital elevation model (DEM) to characterize slope and topography across the study area. I also obtained data layers of vegetation classification, agriculture, and water courses generated by classification of remotely-sensed data *c*. 1997 (FAO 2005).

I plotted collection locations for all samples and drew minimum convex polygons around all sampling locations for each species in each reserve. I measured geographic (Euclidean) distance between all pairs of polygons for each species as the shortest distance between polygon edges, and constructed matrices of pairwise geographic distances.

Least-cost paths and circuit theory

In order to elucidate the effects of landscape elements on gene flow, many studies use least-cost paths (LCP, Adriaensen *et al*. 2003; for example see Coulon *et al*. 2004, Lada *et al*. 2008) to identify the path of least resistance that an individual could take to travel from one habitat patch to another. Resistance layers are built from raster layers (a GIS representation of the landscape divided into cells, where each cell has a single value). Each cell across the landscape is assigned a cost of traversal according to the landscape elements within it; costs are based on some hypothesized resistance of these landscape factors, and can vary within each resistance layer. Then, assuming that an individual will choose the path with the lowest cost between patches (i.e., the path with least resistance), a single optimal pathway across the landscape can be identified. However, least-cost paths assume that an animal has complete knowledge of the landscape and will choose the path with the least resistance. Also, LCP models do not account for gene flow between patches that occurs across multiple generations; these assumptions are likely violated in many species.

In contrast, circuit theoretic modelling (McRae 2006; McRae *et al*. 2008) also uses resistance layers based on GIS rasters, but instead of identifying a single pathway it simultaneously considers all possible paths across a landscape to describe cumulative habitat connectivity. In contrast to least cost modelling, circuit theory does not assume animals have prior knowledge of the habitat, nor that movements are made in a single generation. Circuit theory is based on electrical resistance theory and random walk modelling, and may predict connectivity much more accurately than least-cost paths (McRae and Beier 2007). Thus, I chose to analyse landscape resistance using circuit theory instead of least cost modelling.

Resistance surfaces

First, I generated a raster layer in which each cell was assigned a resistance value of one; lakes were classified as ‘no data’ (i.e., prohibitive to movement) because they were not likely traversed by these species. This ‘unity’ raster was used for tests of isolation by distance (IBD; Wright 1943), that is, assuming that geographic distance

alone (no landscape variables) affected gene flow. I generated this raster so that geographic distance and resistance hypotheses (see below) could be evaluated comparably using circuit theory.

From vector layers of woodland density, grassland density, and water, I generated raster maps with $1km^2$ cell size. Although raster cells of this size forfeit some information regarding fine-scale landscape patterns, the FAO vector layers were relatively coarse themselves, and smaller raster cells increase computation time prohibitively when addressing a landscape of this size. Because water presence is seasonally variable and many streams decrease in flow or disappear entirely during the dry season, I restricted my analysis of water availability to permanent rivers. I then created resistance surfaces by assigning cost values to different landscape features or habitat elements to estimate the effect of each feature on the connectivity and gene flow of a species (isolation by resistance; e.g., Cushman *et al*. 2006, Epps *et al*. 2007). For instance, to test a hypothesis that movement through open areas is ten times more costly than through wooded areas, I assigned densely wooded cells in the vegetation map a value of one, and non-wooded (i.e., grassland) areas a value of ten. For each hypothesis that I tested, I created one or more resistance surfaces with different weighting schemes. I incorporated information from published studies on species-specific requirements such as a dependence on reliable water sources or specific habitat ecotypes (Table 5). In giraffe, I predicted that gene flow would be correlated with geographic distance and slope, but not distance to water (Table 5a). For impala, I predicted that geographic distance, slope, and distance to water would all be correlated with gene flow (Table 5b). For eland, I predicted that gene flow would be correlated with geographic distance but not distance to water and slope (Table 5c).

For categorical variables (i.e., grassland density and woodland density), I constructed several resistance surfaces with different weighting schemes to test species-specific hypotheses (Table 5a-c). For continuous variables (i.e., slope and minimum distance to water), I tested multiple transformations to identify the one with the highest correlation with genetic data. I developed a series of rasters wherein slope

(in degrees) or minimum distance to water (in km) was exponentiated to values ranging from 0.2-2.5 and 0.5-2, respectively. Although I made predictions regarding the effect of slope and distance to water on each species (Table 5) and did not expect both variables to be correlated with gene flow for all three species, I optimized my slope and water models for each species by testing the entire range of transformations against observed gene flow. I selected the weighting scheme that was most highly correlated with genetic distance, and then compared this correlation with the null model of isolation by distance (Shirk *et al*. 2010).

Circuit theoretic modelling

I applied circuit theory modelling (McRae 2006; McRae and Beier 2007) to quantify the cumulative landscape resistance between all population pairs for each weighting scheme assigned in ArcGIS. Sample polygons and landscape raster maps were exported into CIRCUITSCAPE (v.3.5; Shah and McRae 2008); this program uses random walk theory to identify all possible paths that an individual could use to cross a landscape, and then calculates the cumulative resistance of the surface between two areas. Paths were allowed to travel into any of the eight neighbouring cells (i.e., movement was not restricted to the four cardinal directions). I used these resistance estimates to create matrices of population-pairwise cumulative resistance for each hypothesis; this exercise was repeated for all three species.

Tests of isolation by distance and isolation by resistance

Matrices of cumulative resistance were used to test hypotheses of landscape effects. However, I tested a simple isolation-by-distance relationship first, by assessing the correlation of cumulative resistance across a unity landscape (where all cells have a cost of one) with genetic distance (D_{hat}) via a simple Mantel test (Mantel 1967) in XLSTAT, with 10,000 permutations. The Mantel test assesses correlation between distance matrices; unlike multiple linear regression it is not constrained by the requirement that all pairwise distances in the matrix be independent (Smouse *et al*., 1986). The Mantel test computes the Pearson correlation coefficient (r_{YX}) between a

"predictor" matrix (X) and a "response" matrix (Y). It then randomizes full rows or columns (not individual observations) of one matrix and recalculates the correlation between the matrices. Significance of the test statistic is assessed by comparing the observed correlation coefficient to the values obtained through random permutation of the matrices. For each hypothesis of landscape effect I calculated the correlation of the matrix of resistance generated by CIRCUITSCAPE against a matrix of D_{hat} values. For each species, I tested only the habitat factors that I had *a priori* hypothesized to affect gene flow (i.e., woodland density for eland and giraffe and grassland density for impala). Although I expected that slope and distance to water would affect some species more than others (Table 5), I used a univariate optimization approach to identify the best transformation (Shirk *et al*. 2010). I tested cumulative resistance matrices for the entire range of slope and water schemes, to identify the model that was most strongly correlated with gene flow. In this way, I tested hypotheses of the effect of slope and distance to water, but was able to simultaneously optimize my model by comparing a range of potential parameters (Cushman *et al*. 2006, Epps *et al*. 2007, Shirk *et al*. 2010). Within each species, I ranked models by the Mantel test Pearson's correlation coefficient (*r*, which ranges from -1 to 1), and chose the 'optimal' model as the one with the strongest correlation with the matrix of genetic distance. By confronting landscape hypotheses with genetic data, I had an objective method of assessing a variety of models.

Under a causal modelling framework (Cushman *et al*. 2006, Shirk *et al*. 2010), resistance models that performed better than IBD models were then tested with a partial Mantel test (Smouse *et al*. 1986). The partial Mantel test determines the partial correlation ($r_{Y1|2}$) between two matrices (Y and X_1) while controlling for the effect of a third matrix (X_2) that may be correlated with either one or both of the other matrices. If the additional explanatory variable (X_1) is non-spuriously correlated with the dependent variable (Y) then I would expect it to still be significant in a partial Mantel test, after the effects of the initial explanatory variable (X_2) had been controlled. In this case, I controlled for the effect of geographic distance to determine whether a

given resistance model explained variation in gene flow beyond the effects of geographic distance. Geographic distance has a strong influence on cumulative resistance for some resistance models because it is significantly correlated with several explanatory variables (Appendix 2). In the event that more than one landscape variable was significantly correlated with genetic distance, I tested different combinations of the standardized resistance values for the two landscape rasters (e.g., average value for the two cell resistances, maximum value of the two cell resistances) to create additional models and explore whether the two variables in combination would explain a greater proportion of the genetic distance (i.e., multivariate modelling, Shirk *et al.* 2010). Shirk *et al*. (2010) found that this method of systematically varying resistance values for each parameter and then building multivariate models (if necessary) generated models that were highly correlated with genetic patterns.

Mantel and partial Mantel tests have been the focus of some debate: Raufaste and Rousset (2001) suggested that the permutation procedure for calculating significance is invalid in partial Mantel tests and may underestimate the possibility of a Type I error (i.e., rejecting a null hypothesis when in fact it is true). Castellano and Balletto (2002) defend the use of partial Mantel tests but warn that high levels of multicollinearity between explanatory variables can make regression coefficients unreliable. In contrast, Cushman and Landguth (2010) suggest that although simple Mantel tests are prone to spurious correlations, partial Mantel tests and causal modelling parse out the underlying processes and reject incorrect relationships much more reliably when testing resistance models.

RESULTS

Sample sizes and numbers of protected areas sampled per species

I visited eight protected areas to collect samples. Although I attempted to obtain twenty samples per species in each reserve, this was not possible in all cases (Table 3). Due to habitat preferences and varying resource availability these species were not evenly distributed across the reserves, and the network of available roads limited the areas that I could survey. I collected only a few giraffe samples and no impala samples in Rungwa GR because I was only able to access the densely wooded western portion of the reserve; I expect that the eastern and southern areas (which offer grassland and *Acacia* shrubland) would harbour higher densities of those two species. Swagaswaga GR in particular was difficult to survey, in part because of the limited road access but also because of the dense *miombo* woodlands (which make spotting animals difficult) and high levels of poaching. In Selous GR I was forced to discontinue sampling earlier than anticipated due to the onset of the rainy season.

Protected area status also influenced my ability to collect samples. In game reserves, wildlife are often much more skittish as a result of hunting pressure; this made it difficult to detect animals. In contrast to eland and impala, giraffe are protected from hunting in game reserves; sample sizes are probably higher for this species because they were less likely to flee as the vehicle approached. Overall, eland were the most difficult species to sample because they are highly sensitive to human presence and have the greatest flight distances of any African ungulate (Estes 1992). Unfortunately, eland pellets often look quite similar to giraffe pellets, and thus I had higher percent misidentification of dung for this species than others. However, sample sizes at four reserve complexes were sufficient to examine the broader genetic patterns for eland across Tanzania, and I was also able to confirm cross-amplification of several bovine microsatellites in this species. In total, I sampled eight populations of giraffes, seven populations of impala and four of eland. The small number of comparisons possible for eland limited my power to detect differences, however since

this is the first population genetic microsatellite study of this species any inferences may be useful.

Hardy-Weinberg and linkage disequilibrium

After removing individuals with <6 loci successfully genotyped and any duplicate individuals as inferred by CERVUS, my final dataset included 186 giraffe, 133 impala, and 70 eland genotypes (Table 3). After a Bonferroni correction, populations of giraffe exhibited no evidence of linkage disequilibrium or departure from Hardy-Weinberg proportions; however, locus Gica7401 had high suspected null allele frequencies (estimates ranged from 0.40-0.76 across all populations) and it was discarded from further analyses. For the impala samples there was no evidence of linkage disequilibrium between any locus pairs or departure from HWE in any population, nor was there any consistent pattern of null allele presence over locus and reserve, thus all loci and populations were retained for further analyses. Village sample sizes for giraffe (n = 4) and impala (n = 2) were too low for inclusion as separate populations in the analyses of frequency statistics, thus they were used only in the assignment tests. I combined eland samples for Katavi NP (n = 3) with Rukwa GR (n = 5) due to low sample sizes and because of the proximity of sampling locations within those reserves. I excluded Rungwa GR (n = 3), Selous GR (n = 1), and Swagaswaga GR (n = 4) eland samples from frequency analyses because of low sample sizes, but included all samples in assignment tests. Eland samples showed no evidence of linkage disequilibrium, however both TCRBV62 and MAF209 were out of Hardy Weinberg equilibrium (heterozygote deficiency) in Ruaha NP and Tarangire NP. TCRBV62 also showed heterozygote deficiency in Mikumi NP and thus I excluded these two loci from further analyses. There was no consistent pattern of null allele occurrence over eland loci and populations.

Genetic diversity and distances

Giraffe

Over all populations the total number of alleles per locus ranged from 3 (11HDZ073, Gica16120) to 18 (Gica16160) and averaged 7.7 (Table 4a). On a population-by-population basis the average allelic diversity (*A*) was 4.6. Expected heterozygosity (H_E) ranged from 0.50-0.61, with an average H_E of 0.57 (Table 3a). After rarefaction to a sample size of seven (14 alleles), average allelic richness (A_R) per population ranged from 3.1-4.2 (mean = 3.8). Allelic richness was lowest in Selous GR and highest in Rungwa GR. Pairwise F_{ST} values ranged from <0.01-0.17 (mean = 0.08) and D_{hat} values from <0.01-0.27 (mean = 0.12) (Table 6a).

Impala

The number of alleles per locus ranged from 5 (32HDZ707) to 12 (32HDZ688, MCM58) and averaged 8.4 when combined across all populations (Table 4b). The average allelic diversity (*A*) in each population was 5.2. Expected heterozygosity ranged from 0.60-0.71, with an average H_E of 0.67 (Table 3b). After rarefaction to a sample size of five (ten alleles), allelic richness per population ranged from 3.7-4.2 (mean = 3.9). Allelic richness was lowest in Selous GR and highest in Rukwa GR. Pairwise F_{ST} values ranged from 0.01-0.15 (mean = 0.09) and D_{hat} values from 0.05-0.48 (mean = 0.24) (Table 6b).

Eland

When combined across all populations, the number of alleles per locus ranged from 4 (MMP9) to 17 (BM757) and averaged 10.1 (Table 4c). The average allelic diversity (*A*) in each population was 6.7. Expected heterozygosity ranged from 0.61-0.75, with an average H_E of 0.71 (Table 3c). After rarefaction to a sample size of four (eight alleles), allelic richness per locus ranged from 3.2-4.3. Allelic richness was lowest in Mikumi NP and highest in the Katavi NP/Rukwa GR population. Pairwise F_{ST} values ranged from <0.01-0.07 (mean = 0.03) and D_{hat} values from 0.01-0.27 (mean = 0.12) (Table 6c).

Comparison of all three species

The smallest-bodied species did not have the highest genetic diversity as I had predicted. Both allelic richness and expected heterozygosity were highest in eland and lowest in giraffe (Figure 2). Average genetic distance over all population pairs was lowest for eland (mean $D_{hat} = 0.12$) and giraffe (mean $D_{hat} = 0.12$), and highest for impala (mean $D_{hat} = 0.24$). The greatest genetic distances for all three species were between the southeastern populations (Mikumi NP and Selous GR) and the other reserves (Figure 5), although giraffe also showed genetic differentiation between Katavi NP and Tarangire NP that was much higher than that observed in eland and impala.

On a population by population basis, mean genetic distance values for each reserve were similar between eland (D_{hat} range 0.06-0.17) and giraffe (D_{hat} range 0.09-0.19), impala differentiation values were much higher (D_{hat} range 0.18-0.34) (Table 6). In all three species, Ruaha NP had the lowest mean interpopulation pairwise distance (D_{hat} ranging from 0.09 in giraffe to 0.18 in impala). Selous GR had the highest mean pairwise genetic distance of any reserve in both giraffe (mean $D_{hat} = 0.19$) and impala (mean $D_{hat} = 0.34$) and Mikumi NP showed the highest distance among eland populations (mean $D_{hat} = 0.17$).

To test whether patterns of genetic distance were correlated between species, I performed Mantel tests to compare patterns of genetic distance (D_{hat}) between species at the reserves where both were sampled, and found a strong positive correlation between giraffe and impala genetic distances ($r = 0.595$, $P < 0.0001$). I also found strong correlation between interpopulation genetic distances of giraffe and eland, although in this case it was a negative correlation ($r = -0.999$, $P = 0.017$). The negative correlation appears to be mainly driven by the genetic distances between Mikumi NP and Tarangire NP, which are high in eland ($D_{hat} = 0.27$) yet among the lowest observed in giraffe ($D_{hat} = 0.07$). Genetic distances were not significantly correlated between impala and eland ($r = 0.312$, $P = 0.646$).

I investigated the effect of reserve and sampling polygon size on allelic richness and found varying correlations between species. For giraffe and impala, neither reserve size (log-transformed) nor sample polygon size (log-transformed) was correlated with allelic richness ($P > 0.05$ for all comparisons). For eland, the logarithm of the reserve area was not significantly correlated with allelic richness ($R^2 = 0.29$, $P > 0.05$) but the logarithm of the sampling polygon area was correlated ($R^2 = 0.94$, $P < 0.05$). This relationship appeared to be driven by Mikumi NP, which had much lower allelic richness and a sampling polygon that was considerably smaller than the rest of the polygons. The other three reserves had similar allelic richness values despite differences in sample polygon area (Table 3c).

Evaluating population structure with assignment tests

Giraffe

According to Evanno's ΔK method as well as inspection of the average log likelihoods of the different K values, two is the most plausible number of clusters for the giraffe samples (Figure 3a). One cluster contained individuals from Selous GR, Village lands, Mikumi NP, Tarangire NP, and several samples collected from each of Ruaha NP, Rukwa GR, and Swagaswaga GR. The second cluster contained individuals from Katavi NP, Rungwa GR, and the remaining samples from Ruaha NP, Rukwa GR, and Swagaswaga GR (Figure 4a). When the first cluster was reexamined in STRUCTURE for the presence of hierarchical structuring, all individuals from Selous GR (the southeastern-most cluster) and several from Rukwa GR formed a separate cluster (data not shown). When the second original cluster was reexamined, only individuals from Katavi NP (the northwestern-most cluster) separated from the rest. These results suggest that instead of strong population structure across the landscape, giraffe show a genetic gradient of isolation by distance. Because STRUCTURE is not recommended for use with populations exhibiting strong patterns of IBD (Pritchard *et al*. 2007), I did not attempt to elucidate any further hierarchical substructuring.

Impala

The Evanno Δ*K* method and the average log likelihood of the different *K* values suggested that impala are also broadly grouped into two genetic clusters (Figure 3b). One cluster contained individuals from Selous GR and Mikumi NP; the other cluster contained individuals from Katavi NP, Rukwa GR, Rungwa GR, Ruaha NP, Tarangire NP, Swagaswaga GR, and Village lands (Figure 4b). These clusters are separated from one another by the Eastern Arc Mountains. When the second cluster was reexamined for hierarchical substructuring, these samples grouped into two additional clusters: a northern cluster (Tarangire NP and Swagaswaga GR) and a west-central cluster (Katavi NP, Rukwa GR, Rungwa GR, Ruaha NP, and Village lands). Subsequent analysis of the first cluster found that Selous GR and Mikumi NP were not sufficiently substructured to form additional genetic clusters.

Eland

The average log likelihood was highest for K=1, suggesting that the eland populations I sampled are one genetic cluster (Figure 3c).

Estimates of effective population size (N_e)

From ONeSAMP, estimated mean effective population sizes for giraffe ranged from 10-44 (Table 7). Giraffe N_e was lowest in Rungwa GR and highest in Ruaha NP. Mean effective population size for impala was lowest in Swagaswaga GR (N_e = 7) and highest in Tarangire NP (N_e = 130). Estimated mean N_e for eland ranged from 7-135 and was lowest in Mikumi NP and highest in Tarangire NP (Table 7).

Landscape analyses

Giraffe

I compared hypotheses regarding the impact of distance, slope, woodland, and minimum distance to water on gene flow (Table 5a), and found that geographic distance was a very good predictor of genetic distance (r = 0.72; P < 0.0001; Figure 6a). After testing multiple transformations of slope degrees, I found that a resistance

layer based on slope (in degrees, nontransformed) was more strongly correlated with genetic distance ($r = 0.80$; $P < 0.0001$; Figure 7). However, because geographic distance strongly influences cumulative resistance estimates regardless of the resistance model (B. McRae, pers. comm), I applied a partial Mantel test and found that the effect of slope was still significant ($r = 0.52$; $P = 0.003$) even after accounting for the effect of geographic distance (Table 8a). Neither of the woodland-based resistance models performed better than the null (IBD) model, nor did any of the models based on transformations of the minimum distance to water. Cumulative resistance estimates for the slope (nontransformed) and geographic distance models were highly correlated ($r = 0.90$, $P < 0.0001$; Appendix 2).

Impala

For impala, geographic distance alone was a relatively poor predictor of genetic distance ($r = 0.48$; $P = 0.026$; Figure 6b). I tested hypotheses regarding the impact of slope, grassland and minimum distance to water (Table 5b) and found that a resistance model based on slope (in degrees, squared) was a much better predictor ($r = 0.91$; $P < 0.0001$; Figure 7). Even after geographic distance was accounted for by partial Mantel test, slope-squared remained highly significant ($r = 0.88$, $P < 0.0001$). In combination with the population structure suggested by STRUCTURE, this suggests that steeper slopes strongly impede dispersal in impala, and that the Eastern Arc Mountains are a barrier to gene flow. I also tested the effect of grassland and found that a model ranking high grassland percentage as low resistance and areas with lower levels of grassland as highly resistant, with woodland being the most costly to traverse (hypothesis "Grass1", Table 5b) was a better predictor of genetic distance than geographic distance alone ($r = 0.49$; $P = 0.024$); however this improvement was not significant once geographic distance was removed (partial Mantel test: $r = 0.14$; $P = 0.544$; Table 8b). Although it has been suggested that impala distribution is moderately constrained by distance to water (Estes 1992), I found no evidence that distance to water influenced gene flow. The best water model (distance to water exponentiated to a factor of 1.5) was not significant after geographic distance was

removed (partial Mantel test; $r = 0.19$, $P = 0.409$; Table 8b). Cumulative resistance estimates for the slope (squared) model and the geographic distance model were correlated ($r = 0.54$, $P = 0.013$; Appendix 2).

Eland

Geographic distance was a poor predictor of genetic distance ($r = 0.47$; $P = 0.372$; Figure 6c), although statistical power was lower for this species due to the small number of populations for comparison. Models of landscape resistance incorporating slope and distance to water both performed better than the null IBD model (Table 8c). For slope models, the highest correlation coefficient was obtained for the resistance model incorporating nontransformed slope ($r = 0.78$; $P = 0.029$; Figure 7). This model still approached significance after geographic distance was accounted for ($r = 0.72$, $P = 0.06$; Table 8c). A model that included the minimum distance to water (in kilometres) squared also performed much better than the null IBD model ($r = 0.92$, $P = 0.007$) and was still significant after accounting for geographic distance ($r = 0.92$, $P = 0.007$; Table 8c). I tested several models that examined combinations of standardized slope and distance to water (e.g., one model assigning each raster cell the maximum value of the two variables, another model in which the values for both variables were averaged, and additional models in which the two variables were weighted as one-third and two-thirds the final value of each cell, and vice versa). None of these combination models performed better than the simpler models with only distance to water squared or nontransformed slope (data not shown). For eland, distance to water (squared) was not correlated with geographic distance ($r = 0.66$, $P = 0.175$; Appendix 2), nor was slope (nontransformed) ($r = 0.74$, $P = 0.146$; Appendix 2). However, distance to water (squared) was highly correlated with nontransformed slope ($r = 0.94$, $P = 0.007$; Appendix 2).

Scale of gene flow

Although I expected genetic diversity to scale with body size, this pattern was not seen. Eland had the highest allelic diversity and expected heterozygosity, and

these measures were the lowest among populations of giraffe. Geographic distance was a relatively poor predictor of genetic distance for impala and eland (Figure 6); evidence of IBD was observed in giraffe however correlation was higher when landscape factors (specifically slope) were included. Resistance values showed much higher correlation with genetic distance; however, the optimal models were different for all three species. Interestingly, resistance values were highly correlated with geographic distance in giraffe ($r = 0.90$), while less so in impala ($r = 0.54$) and eland ($r = 0.66$). As a metric of relative historical gene flow between populations, *Nm* estimates showed that connectivity declines with increasing resistance (Figure 9).

Nm ranged from 3.2-12.3 for giraffe populations. Some of the highest values of *Nm* were found between adjacent reserves, for instance Ruaha NP-Rungwa GR (*Nm* = 9.3) and Katavi NP-Rukwa GR (*Nm* = 8.2). Some comparisons between nonadjacent reserves also showed high values of gene flow (e.g., Ruaha NP-Rukwa GR: *Nm* = 11.7, Rukwa GR-Rungwa GR: *Nm* = 12.3, Swagaswaga GR-Ruaha NP: *Nm* = 6.7, Swagaswaga GR-Tarangire NP: *Nm* = 6.3; Figure 10a) and highlight areas that were historically quite connected. As expected for giraffe, the lowest *Nm* values were observed between reserves that are separated by large geographic distances (e.g., Katavi NP-Tarangire NP: *Nm* = 3.2; Figure 10a) or terrain with a lot of slope (e.g., Tarangire NP-Selous GR: *Nm* = 3.7; Figure 10a).

In impala, *Nm* ranged from 2.1-9.8 (Figure 10b). As in giraffe, *Nm* values between connected reserves were high (e.g., Katavi NP-Rukwa GR: *Nm* = 9.8). I also observed high historical levels of gene flow between Ruaha NP-Rukwa GR (*Nm* = 8.7), Ruaha NP –Tarangire NP (*Nm* = 6.7), and Katavi NP-Tarangire NP (*Nm* = 5.8). Unlike giraffe, impala in Swagaswaga GR and Ruaha NP had only moderate historical gene flow (*Nm* = 4.3), though this is probably an artefact of genetic drift acting on a small population size in Swagaswaga GR, as Tarangire NP is farther away but shows higher connectivity with Ruaha NP. Some of the smallest *Nm* values are found between Tarangire NP-Selous GR (*Nm* = 2.6; Figure 10b) and Mikumi NP-Ruaha NP (*Nm* = 2.8); these populations are separated by large slopes.

Lastly, *Nm* ranged from 3.3-11.1 for eland (Figure 10c). The highest level of connectivity was observed between Ruaha NP and Tarangire NP (*Nm* = 11.1; Figure 10c) and Katavi NP and Ruaha NP (*Nm* = 9.0); the lowest values were between Mikumi NP and Tarangire NP (*Nm* = 3.3) and Katavi NP and Mikumi NP (*Nm* = 3.5).

DISCUSSION

Genetic diversity

The first objective of this study was to obtain baseline genetic information for three wide-ranging ungulate species, in a region of Tanzania that has hitherto had little population genetic research on ungulates. Few if any in-depth population genetic studies have been conducted on these species. Brown *et al*. (2007) investigated microsatellite variation in giraffe across Africa but did not present fine-scale patterns of variation and structure, nor did they investigate the effects of biogeographic elements. Lorenzen *et al*. (2006) conducted a similarly broad genetic study of impala. To my knowledge this is the first population genetic study of eland that uses microsatellites. Most studies that include Tanzania sample northern populations only (e.g., van Hooft *et al*. 2000, Brown *et al*. 2007), yet I sampled most (or all) of the major populations of these species across a wide region of Tanzania. Moreover, this study provides the first clear comparison of nuclear genetic structure for these species (or possibly any other large African mammals) on a single, systematically-sampled landscape.

Mean number of alleles per locus was moderate across all three species (range 4.6-6.7; Table 4); studies have found similar values for roan antelope (*Hippotragus equinus*; Alpers *et al*. 2004: $A = 3.4$), common impala (Lorenzen *et al*. 2006: $A = 5.0$), Swayne's hartebeest (*Alcelaphus buselaphus swaynei*; Flagstad *et al*. 2000: $A = 6.9$), and African buffalo (*Syncerus caffer*; van Hooft *et al*. 2000: $A = 7.1$). Despite its vast area and counter to the expectation of the island biogeography model of reserve isolation (Soulé *et al*. 1979, Newmark 1996) Selous GR had the lowest corrected allelic richness for all three species (Table 3). Unfortunately, I was only able to access a small portion of the reserve to sample, but regardless I would have expected greater diversity in this area because larger reserves are expected to have larger population sizes and therefore be less affected by genetic drift. Even Swagaswaga GR, which was gazetted fewer than two decades ago and is much smaller than the other reserves, has allelic diversity comparable to larger and older reserves (Table 3). After

rarefaction to a sample of eight alleles, eland had the highest allelic richness, followed by impala and then giraffe (Figure 2).

Average expected heterozygosity was also moderate (range 0.57-0.71; Table 3), and was similar to studies of other African ungulates including roan antelope (Alpers *et al*. 2004: H_E = 0.46), Swayne's hartebeest (Flagstad *et al*. 2000: H_E = 0.70), and African buffalo (Heller *et al*. 2008: H_E = 0.79). Expected heterozygosity for the three species followed the same patterns as allelic richness, in that H_E was also highest in eland and lowest in giraffe (Figure 2).

Eblate *et al*. (2011) tested 38 microsatellite primer pairs for amplification in common eland: 22 isolated from bovine, ovine and caprine sources, and 16 from blue wildebeest (*Connochaetes taurinus*). Of those 38, 20 successfully amplified. I screened seven of the same primer pairs and found that only two of them were scoreable and polymorphic in my samples (BM804 and MAF209, the latter of which was discarded due to homozygote deficiency). This is surprising because that study used only 15 individuals from a single protected area (Serengeti NP), while my samples (n = 70; 4 protected areas) covered a much broader geographic range. Although I found homozygote deficiency at locus MAF209 in my samples, they did not see the same result (H_O = 0.60). However, one locus that did not amplify in their samples did amplify in mine (BM757) and had moderate levels of expected heterozygosity (H_E = 0.69) and a high number of alleles (A = 17). The one locus that both studies shared (BM804) had similar values of expected heterozygosity between studies (Eblate *et al*. 2011: H_E = 0.85; this study: H_E = 0.81; Table 3c) and total number of alleles (Eblate *et al*. 2011: n = 11 alleles; this study: n = 13 alleles). Grand mean values of observed and expected heterozygosities were also similar among studies (Eblate *et al*. 2011: H_O = 0.69, H_E = 0.76; this study: H_O = 0.70, H_E = 0.71). Together, these results suggest that these three species currently have a moderate amount of genetic diversity in the reserves studied, and that they do not show signs of genetic depauperacy due to isolation.

Genetic distance and population structure

Lorenzen *et al*. (2010) conducted a phylogenetic study of common eland across east and southern Africa using mitochondrial DNA. They found that samples from East Africa (Tanzania, Kenya, Uganda and Ethiopia) exhibited strong population differentiation, low nucleotide diversity, and a more recent common ancestor relative to southern African (Zambia, Zimbabwe, Namibia, Botswana and South Africa) populations. Lorenzen *et al*. (2010) also found that Rukwa GR (misidentified as Rungwa GR) had some of the highest nucleotide diversities of any population in their study, and reported that individuals that clustered with both the eastern and southern lineages were found within that population. In my study the Katavi NP/Rukwa GR complex has the highest corrected allelic richness (A_R = 4.32) for eland populations although it does not appear to be statistically different than other reserve complexes (e.g., Ruaha NP; A_R = 4.22). Interestingly, their study found evidence of IBD among East African populations, but not within the southern African populations. I failed to see evidence of isolation by distance among my populations and overall structure appeared to be much lower than for giraffe and impala (Table 6), with moderate to high levels of genetic diversity (Figure 2). However, my sample sizes were quite low for several populations (Table 3c) and I may not have been able to detect structure between these reserves. Also, IBD tends to be seen at larger scales, such as the range-wide study in Lorenzen *et al*. (2010). Future studies should compare microsatellite data for eland on a much larger scale and with larger sample sizes, to see whether the same pattern that Lorenzen *et al*. (2010) report is evident in nuclear data.

Brown *et al.* (2007) found significant substructuring among Maasai giraffe populations within Serengeti NP; their maximum reported F_{ST} value within Serengeti subpopulations (Kirawira versus Seronera; F_{ST} =0.13) is almost as large as the largest differentiation that I observed across my entire study range (Katavi NP versus Selous GR; F_{ST} =0.17; Table 6a), even though their Serengeti NP populations are at most 130 km apart and are not separated by any major topographical features. Brown *et al.* (2007) suggest that despite close proximity of sampling locations and putative fluidity

of social groups, populations of Maasai giraffe are reproductively isolated by ecological or behavioural factors, even within continuous habitat like the Serengeti NP. I did not observe these patterns in my sampling localities; F_{ST} values between reserves within the same complex (e.g., Ruaha NP-Rungwa GR, Katavi NP-Rukwa GR) were among the lowest F_{ST} values observed for this species (0.02 and 0.04, respectively; Table 6a). Brown *et al*. (2007) also suggest that the Maasai giraffe is actually comprised of two species, separated by the Rift Valley. While I did not sample this area, their discovery of strong genetic differentiation in an area separated by increased slope is not surprising. However, the great difference in the range of F_{ST} values observed between our studies cautions against subdividing the Maasai subspecies without further investigation.

Brown *et al*. (2007) also tested for IBD patterns in giraffe; they found significant correlations between linearized F_{ST} and geographic distance for all Maasai populations and also within the five Serengeti Maasai populations, but not for the reticulated (*G. c. reticulata*) or Angolan (*G. c. angolensis*) populations, nor across all subspecies. My dataset has an array of populations separated by a wide range of geographic distances and I also found strong evidence of IBD in the Maasai subspecies.

Schwab *et al*. (2011) investigated genetic structure of common impala in two populations in South Africa, using mitochondrial sequencing and seven microsatellite loci (of which four are shared with this study). They sampled populations that are among the southernmost portion of the species' range. To test the hypothesis of decreasing genetic diversity from southern to East African populations, Schwab *et al*. (2011) compared their results with those of Lorenzen *et al*. (2006), who sampled impala across much of their range, from South Africa to Kenya. Both studies found high levels of genetic diversity in South African populations (e.g., Schwab *et al*. 2011: $H_E = 0.72$); Lorenzen *et al*. (2006) also found high diversity in both Tanzanian populations sampled (e.g., Burko Forest Reserve: $H_E = 0.74$). Interestingly, I found that Selous GR had the lowest allelic richness ($A_R = 3.71$; Table 3b) and second

lowest expected heterozygosity (H_E = 0.63) of sampled reserves, but Lorenzen *et al*. (2006) reported genetic diversity for this game reserve (H_E = 0.73) comparable to that of South African populations. This would suggest that the samples in my study were of too restricted a range to characterize the diversity of the reserve as a whole, however it is interesting to note that my sample size for Selous GR (n = 26) was twice that of Lorenzen *et al.* (2006) (n = 13). Unfortunately, Lorenzen *et al*. (2006) do not present a map of sampling areas within each reserve, and thus I could not compare the geographic range of our samples. The Kenyan and Ugandan populations in Lorenzen *et al*. (2006) had the lowest genetic differentiation of all of the populations they sampled (Samburu National Forest Reserve, Kenya: H_E = 0.56, and Lake Mburu NP, Uganda: H_E = 0.61), although the authors had <15 samples for each of the eight common impala populations they surveyed.

Nersting and Arctander (2001) suggested that higher diversity in southern impala populations relative to East African populations could be a result of the species colonizing East Africa from South Africa. My results show that genetic diversity of impala populations in Tanzania (mean H_E = 0.67; Table 3b) is intermediate to that of South African and more northern (i.e., Kenyan and Ugandan) populations, however only two populations in Kenya and Uganda have been studied to date (by Lorenzen *et al.*, 2006), and sample sizes were small. Alternatively, Schwab *et al*. (2011) proposed the existence of several Pleistocene refugia (some in East Africa), from which regional differentiation and recolonization occurred. Additional sampling in the northern periphery of the species' range will be needed to resolve this debate.

Effective population size

The estimates of effective population size seemed to be rather small for all three species (Table 7). In giraffe, Ruaha NP had the largest mean N_e (44), while Rungwa GR had the smallest (10). Effective population size for impala ranged from 7 (Swagaswaga GR) to 130 (Tarangire NP) and had a similar range of values as the N_e observed for eland (7 in Mikumi NP to 135 in Tarangire NP). Frankham (1995) suggested that the ratio of N_e to the population census size (N_c) is typically on the

order of 0.1, thus we would expect census population sizes to be ten times as large as these estimates. Even with this adjustment, these estimates are considerably lower than published population sizes based on aerial surveys and walking transects. Barnes and Douglas-Hamilton (1982) used aerial survey data from 1977 to estimate total population sizes for a suite of ungulates in Rungwa GR and Ruaha NP. In Ruaha NP, they estimated that there were 3478 (+1861 SE) giraffe, 9075 (+ 6598 SE) impala, and 1755 (+1155 SE) eland. Even after taking into account the fact that N_e values are often one-tenth of the census size (Frankham 1995), the estimates generated for Ruaha NP by ONeSAMP for giraffe ($N_e = 44$), impala ($N_e = 25$) and eland ($N_e = 22$) were much smaller. Barnes and Douglas-Hamilton's (1982) estimates of N_c for Rungwa GR were also considerably larger than those I generated with ONeSAMP. Waltert *et al*. (2008) used walking surveys to count wildlife in Katavi NP and Rukwa GR and estimated that there were over 5000 giraffe, 2000 - 5000 eland, and 30000 impala in those reserves. Again, my estimates of giraffe, impala, and eland population sizes for these reserves were much smaller (Table 7). However, I do not believe that this is indicative of a true change in the census size for these species in all of these populations. ONeSAMP is sensitive to the number of samples and loci used (Tallmon *et al*. 2008), and with the relatively small sample sizes of this study, the N_e estimates must be regarded with caution. It is interesting to note, however, that N_e estimates were not concordant with reserve size: despite being the largest reserve, Selous GR had only moderate effective population size for both giraffe and impala (Table 2, Table 7). Also, although Tarangire NP was the second smallest reserve, it had the highest estimated effective population sizes for both impala and eland. Thus it appears that, at least with a limited number of samples and loci, reserve size does not accurately predict effective population size for these species.

Effective population size did not correspond with allelic richness in any of the species either. Although Rungwa GR had the lowest effective population size for giraffe ($N_e = 10$; Table 7), it was the reserve with the highest allelic richness for the species ($A_R = 4.24$; Table 3a). In impala, the effective population size of Tarangire NP

was over ten times that of Swagaswaga GR (N_e = 130 versus N_e = 7; Table 7) but the allelic richnesses of these reserves was very similar (Swagaswaga GR: A_R = 3.80; Tarangire NP: A_R = 3.85; Table 3b). It is possible that the correlation between allelic richness and effective population size was not detectable with the number of samples and loci in this study.

Effect of landscape on genetic structure

The second objective of this study was to link genetic differentiation with landscape attributes. Life history traits and species-specific habitat requirements can have a pronounced impact on dispersal capability. By inferring historical connectivity across a variety of landscape features we can elucidate the relative impacts of different landscapes on gene flow. It is possible to contrast these patterns across multiple species to identify areas that may serve as wildlife linkages in an increasingly fragmented habitat. Among studies that utilize GIS-based resistance surfaces to investigate the effect of landscape on genetic structure, very few are conducted in a multi-species framework. This approach allows for valuable comparisons across species with different dispersal capabilities, and offers a robust way to identify biogeographic attributes that may limit gene flow for a suite of species across the landscape.

I created hypotheses for each species based on published literature regarding habitat preferences (Table 5) but also on personal observation of species distributions while conducting my field work. However, the fact that habitat models (i.e., those generated from woodland and grassland rasters) were not included in optimal models for any of the species is not surprising. There seems to be no consensus of what habitat is preferable for these species, and studies that seek to address this issue use habitat categories that are often difficult to reconcile between studies. As a result, 'optimal' habitat for a species can vary dramatically between studies. For instance, van Bommel *et al*. (2006) used satellite imagery to categorize vegetation based on its Normalized Difference Vegetation Index (NDVI) values. They found that impala preferred woodland with intermediate greenness values and were found at or below expected

densities for woodland with high and low greenness values, grassland, and grass on floodplains. In contrast, Averbeck (2002) found season-specific habitat preferences: during the rainy period, impala preferred flat valley bottoms with short grass and <20% canopy cover, but in the dry season they moved into areas with more slope, taller grass, and 20-80% canopy cover. For giraffe, Leuthold and Leuthold (1978) observed a preference for densely wooded areas, with some use of riverine areas during the dry season. Conversely, Riginos and Grace (2008) discovered giraffe dung three times as often more often in sparsely wooded areas than in thicker forests. Sampling methods varied widely among these studies but the contrasting conclusions highlight the lack of consensus for habitat preferences for these species, and illustrate how my hypotheses of habitat effects could be improved. However, the strong correlations observed between slope and genetic distance for all three species evince the impact that this landscape element has on gene flow.

Although widespread and sometimes detected outside of protected areas (Epps *et al*. 2011), impala show strong population structure across Tanzania. It has been suggested that impala are moderately constrained by a reliance on surface water (Estes 1992) and that their distribution across the landscape is patchy for this reason. However, my models that included distance to water did not adequately explain patterns in genetic structure, nor did models addressing percent woodland and grassland. For this species, genetic distance is strongly explained by a model based on the squared exponential power of slope; the major genetic break observed in this species separated Selous GR and Mikumi NP from the other populations, thus high slopes (especially the Eastern Arc Mountains) act as a barrier to dispersal. Averbeck (2002) found that impala preferred areas with low levels of slope (< 5 degrees) in the rainy season, but were found in areas with higher level of slope (~5-15 degrees) during the dry season. This suggests that impala are able to traverse areas with moderate amounts of slope; indeed there may be thresholds of tolerance of slope steepness that have not been exposed by the exponential weighting scheme I applied. When the intervening habitat between the populations in my study had a high cumulative

resistance because of steep slopes, impala exhibited relatively high levels of genetic distance even at small geographic distances (e.g., D_{hat} = 0.30 between Ruaha NP and Mikumi NP, with a geographic distance of <200km). Conversely, the five populations north and west of the Eastern Arc Mountains showed much lower pairwise genetic distances (D_{hat} = 0.05–0.24, average = 0.14; Table 6b) even at greater geographic distances; intervening habitat between those reserves is not steeply sloped.

Impala populations in Katavi NP and Ruaha NP, which are separated by >500km, have a D_{hat} of 0.08, suggesting moderate levels of gene flow historically. The intervening landscape between these reserves currently has relatively little human settlement (Figure 1), and thus it is possible that some level of connectivity has persisted. Tarangire NP and Ruaha NP also show low levels of historic differentiation (D_{hat} = 0.11; Table 6b). In contrast, the land between these reserves has undergone considerable transformation from woodland to agriculture and human settlement in the last fifty years (Figure 1), and current levels of connectivity between these populations may now be quite lower.

The effect of mountainous terrain on impala was also hypothesized by Schwab *et al*. (2011). They characterized genetic structure of impala in two provinces in South Africa that are separated by the Drakensberg Mountain Range, and found significant differentiation with few admixed individuals between the provinces. The average F_{ST} they observed across all population pairs that were separated by the mountains was 0.12, which is very similar to the average F_{ST} that I found over all comparisons of Mikumi NP and Selous GR with populations north and west of the Eastern Arc Mountains (average F_{ST} = 0.13; Table 6b). Many reserves in South Africa are fenced, making gene flow between populations unlikely. Schwab *et al*. (2011) attribute admixed individuals to translocation efforts between reserves, and suggest that future actions preserve the genetic differentiation by only relocating impala to populations within their natal province. In Tanzania the national parks and game reserves are not fenced, allowing migration between reserves; translocation for hunting purposes is not commonplace. I observed remarkably little admixture within impala individuals

(Figure 4b), and most 'misassigned' samples appeared to be individuals that grouped in the northern and western cluster but were sampled south of the Eastern Arc Mountains.

Low D_{hat} values between distant populations also suggest that eland historically sustained significant gene flow over long distances, and thus might be particularly vulnerable to loss of connectivity. As seen in giraffe and impala, eland populations in Ruaha NP and Tarangire NP had high historical levels of gene flow (D_{hat} = 0.07; Table 6c). Katavi NP/Rukwa GR and Ruaha NP also had high gene flow historically (D_{hat} = 0.01). As with impala, the greatest population differentiation in eland occurred between populations on either side of the Eastern Arc Mountains (Mikumi NP-Tarangire NP; D_{hat} = 0.27).

Giraffe were once widespread across Tanzania but are now mostly confined to protected areas (Fennessy and Brown 2010). Like impala, giraffe show patterns of isolation by slope; however, they also appear to be relatively constrained by geographic distance. Thus Katavi NP and Tarangire NP, which showed relatively little differentiation in impala (D_{hat} = 0.13; Table 6b), have historically been more isolated in giraffe (D_{hat} = 0.19; Table 6a). The pattern of isolation by distance is further suggested by population structure tests that assign individuals from centrally located reserves to both eastern and western clusters, but with low assignment probabilities. As well, giraffe do not show a pronounced genetic break at the Eastern Arc Mountains as do impala. Differentiation between Selous GR and Mikumi NP (D_{hat} = 0.13) is actually greater than that observed between Mikumi NP and Ruaha NP (D_{hat} = 0.10) even though Selous GR and Mikumi NP are adjacent (but interlaying areas have quite steep terrain), and Ruaha NP lies on the opposite side of the Eastern Arcs. Although vagility and thus dispersal capability might be expected to be greater for animals with larger body mass, giraffe show higher levels of differentiation than smaller antelope species (i.e., impala) over large geographic distances.

Leuthold and Leuthold (1978) investigated habitat preferences of giraffe in Tsavo National Park, Kenya, and observed that giraffe preferred densely wooded areas

in the rainy season and riverine areas during the dry season. They suggested that human activities that convert wooded areas into more open habitats (e.g., settlement, agriculture) may be detrimental to giraffe populations. They also noted that giraffe densities are higher outside the park during the wet season, and that this species may not be entirely sustained by the national park. It is possible that a similar pattern of land use occurs between many Tanzanian reserves and the surrounding landscape. Multiple historical migration routes that linked Tarangire NP with adjacent land outside the park have now been severed (Borner 1985); wildlife populations may suffer declines in the future if present-day park boundaries are not sufficient to provide year-round forage.

Slope did not affect eland as strongly as it did the other two species. This is consistent with reports of eland sightings at 4900m elevation (Hillman 1974). Hillman (1988) suggested that eland use low elevation plains during the wet season and move into higher altitude bush regions during the dry months; however, that study did not examine the effect of slope. Distance to water showed the greatest correlation with genetic distance. This is surprising, because eland are generally believed to be less limited by water availability than other ungulates, as they obtain a large proportion of their moisture from forage (Thouless 2008). However, due to the small number of eland populations in this study, I had limited statistical power for these comparisons. Nontransformed slope and the squared value of distance to water were highly correlated in eland (partial Mantel test, $r = 0.90$, $P = 0.029$, Appendix 2), and thus it is likely that slope has a large effect on eland connectivity as well.

Effect of reserve size on genetic diversity

Heller *et al*. (2010) used microsatellite loci to examine patterns of genetic diversity and structure among ten African buffalo populations in Kenya and Uganda. They observed that despite previous studies reporting extremely high variation and low structure (e.g., Simonsen *et al*. 1998) across the species' range, populations in Kenyan and Ugandan protected areas show regional genetic structuring. They found correlations between reserve size and genetic diversity (both allelic richness: $R^2 = 0.59$

and mean heterozygosity: $R^2 = 0.60$), and suggest that isolation on a relatively recent (<100 years; ~13 generations) time scale has caused a loss of genetic diversity in smaller reserves. I also calculated the correlations between reserve size (log-transformed) and allelic richness but did not find similar patterns (giraffe: $R^2 = 0.31$, impala: $R^2 = 0.002$, eland: $R^2 = 0.29$), nor did I find significant correlations between the area of the sampling polygon size (log-transformed) and allelic richness for giraffe ($R^2 = 0.01$) or impala ($R^2 = 0.08$). The area of sampling polygons (log-transformed) was significantly correlated with allelic richness in eland ($R^2 = 0.94$). However, comparison was based on only four populations and the relationship was driven mostly by Mikumi NP which has a much lower allelic richness than the other populations (Table 3c).

Heller *et al*. (2010) also found a correlation between logged reserve size and reserve isolation for each population ($R^2 = 0.72$). They suggested that larger reserves are better able to maintain diversity and are less affected by genetic drift. Although I did not test for a correlation between genetic distance and reserve size, I did not find a correlation between allelic richness and reserve size or sampling polygon size for my species. This suggests that reserve size is not the best proxy of genetic diversity in this system, and that smaller reserves have not yet begun to show the effect of genetic drift and small population sizes. However, reserve isolation could still be an important factor within this system.

Predicted impacts of current human activity levels

The fourth and final objective of this study was to identify areas that historically had high levels of connectivity but, due to recent expansion in human settlement and agriculture, may be at risk of isolation. Densities of human settlement are increasing unevenly across Tanzania, with the greatest amount occurring in the northeastern region near Arusha (Figure 1).

In all three species, the Katavi NP/Rukwa GR complex and the Ruaha NP/Rungwa GR complex appeared to historically have high connectivity (Table 6). The slope-based surfaces of current generated by CIRCUITSCAPE (Figure 11a-c) show

broad areas of low resistance between these reserves and do not suggest any bottlenecks (i.e., areas where topography is sufficient to constrain movement to a very narrow passage). The area between Katavi NP/Rukwa GR and Ruaha NP/Rungwa GR is relatively intact and has not undergone the extensive expansion in human settlement that other regions in Tanzania have faced (see agricultural activity, Figure 1). As an area with high historical connectivity and relatively sparse human activity to date, the Katavi NP/Rukwa GR and Ruaha NP/Rungwa GR connection appears to have very high potential for connectivity preservation. However, the layer of agricultural presence that I used as an index of human activity was generated from data collected in 1997, and as such can be assumed to be conservative relative to current patterns.

Another trend that was evident across all three species is the relative lack of connectivity among eastern and southeastern populations (i.e., Tarangire NP/Swagaswaga GR versus Mikumi NP/Selous GR) and central and southeastern populations (i.e., Ruaha NP/Rungwa GR versus Mikumi NP/Selous GR). This is most likely a result of high resistance due to steep slopes in intervening habitat (Figure 11). However, Epps *et al*. (2011) conducted walking surveys across protected and nonprotected land between Ruaha NP and Mikumi NP and found that despite increased human settlement and agriculture, elephants still make long-distance movements across this area. In addition, they detected sign of over 30 additional large mammal (>1kg) species, including spotted hyaena (*Crocuta crocuta*), greater kudu (*Tragelaphus strepsiceros*), and leopard (*Panthera pardus*) in lands outside of the two national parks. These results highlight the conclusion that connectivity is often species-specific, and that interspecies comparisons are important for understanding the relative impact of biogeographic features on connectivity of species with different dispersal capabilities and habitat tolerances.

I was also able to identify regions where human settlement and agriculture are likely to sever areas of historically high gene flow between the study populations. In all three species, populations in the northeast (i.e., Tarangire NP and Swagaswaga GR) had high levels of connectivity with the centrally located Ruaha NP/Rungwa GR

ecosystem (Table 6; Figure 11). The intervening landscape has already undergone some habitat fragmentation due to agriculture and settlement (Figure 1), and isolation of these reserves may have begun although it is still not detectable in genetic structure. Since most of the expansion in human settlement and agriculture in Tanzania has occurred within the last 50-100 years due to the increase in human population observed over that time period, generation times of giraffe (at least 4 years, Dagg 1971), impala (approximately 2 years, Estes 1992) and eland (3 years, Estes 1992) suggest that there have been approximately 25-50 generations since isolation of some of these reserves. With moderate to large population sizes in each of these reserves, it is likely that the impacts of isolation would not yet be detected. However, this is an area where mitigation of habitat loss, through creation of wildlife linkages, could possibly ameliorate some of the expected effects of isolation.

There are also population pairs that showed strong historical connectivity in some species but less in others. For instance, Tarangire NP and Katavi NP had moderate genetic distance in giraffe (probably due to geographic distance) but because of little intervening slope, they had little genetic distance in eland and impala (Table 6). As above, increased settlement in the northeastern region could sever this connection; Borner (1985) called attention to the loss of migration pathways around Tarangire NP more than two decades ago. Additionally, while Tarangire NP and Swagaswaga GR appear to have had little historical connectivity with southeastern reserves (Table 6), agriculture in the lands around Tarangire NP and Swagaswaga GR could isolate these two populations from one another. Moreover, additional sampling should target reserves north of Tarangire NP (e.g., Lake Manyara NP, Serengeti NP, Maswa GR), to quantify genetic structure in this area and determine how strongly Tarangire NP was historically linked to these populations.

CONCLUSIONS

This study provides the first in-depth comparison of genetic diversity and structure for multiple ungulate species across a broad area of Tanzania. Using these empirical data I compared and objectively assessed different hypotheses of biogeographic effects, and identified landscape factors that were most relevant to patterns of historical gene flow. I identified regions where historical connectivity was high, and where current changes in land use could inhibit gene flow. Reserves located in the northeastern portion of the study area appear most likely to become isolated due to increasing human activity; pairwise genetic distances for these populations did not show abnormally low levels of gene flow for any species. However, changes in genetic structure are dependent on effective population sizes, and because the populations I surveyed are relatively large and habitat fragmentation is recent, loss of connectivity is unlikely to be evident in genetic structure at this time. Field surveys across the habitat between these reserves (e.g., as performed in Epps *et al*., 2011) would be a valuable tool for elucidating current connectivity patterns and targeting specific areas for linkage. The western corridor (between Katavi NP/Rukwa GR and Ruaha NP/Rungwa GR) appears to have historically had very high levels of gene flow for all three species, and should be a priority for conservation connectivity.

- Genetic diversity (allelic richness and expected heterozygosity) of giraffe, impala, and eland in reserves in central and southern Tanzania is moderate and within reported ranges for other African ungulate species; within these populations it appears that genetic diversity is highest in eland, then impala, and then giraffe.
- I did not find a correlation between allelic richness and reserve size or sampling polygon size; this suggests that reserve size does not predict present-day genetic diversity in this system. Further, smaller reserves do not appear to have begun to show the predicted effects of genetic drift and small population sizes. However, my study design was not specifically tailored to test these relationships, nor did it look for a correlation between genetic diversity and reserve isolation.

- Population assignment tests showed different patterns for the three species I evaluated. Giraffe show evidence of isolation by distance across this area. Impala are broadly clustered based on topography (east versus west of Eastern Arc Mountains) but also show substructuring that separates the northeastern populations (Tarangire NP/Swagaswaga GR) from the central (Ruaha NP/Rungwa GR) and southwestern (Katavi NP/Rukwa GR) populations. I did not observe any structure in eland populations using assignment tests, but sample sizes and the number of reserves for comparison were low.
- Patterns of genetic distances between populations were broadly similar across all three species. The greatest geographic distances were observed in population pairs that were separated by the Eastern Arc Mountains. However, there were also striking species-specific differences in historical connectivity, highlighting the importance of multiple species comparisons when conducting connectivity conservation.
- Along with IBD, gene flow (and thus movement) among giraffe populations was strongly affected by slope (nontransformed). Gene flow among impala populations was most correlated with slope (squared), and gene flow in eland was affected by both minimum distance to water (squared) and slope (nontransformed). This suggests that impala may be more affected by steeper slopes than are giraffe and eland.
- For all species, gene flow between southwestern (Katavi NP/Rukwa GR) and central (Ruaha NP/Rungwa GR) reserves was historically high, and those reserve complexes are currently separated by limited, low density human settlement. This area is an obvious candidate for connectivity conservation and surveys of current-day connectivity would be useful to further investigate gene flow patterns in this area.
- Northeastern (Tarangire NP/Swagaswaga GR) and southeastern (Mikumi NP/Selous GR) populations historically had little connectivity for all three species, as did central (Ruaha NP/Rungwa GR) and southeastern (Mikumi NP/Selous GR)

populations. However, other species (e.g., elephants; Epps *et al*. 2011) are still able to make long-distance movements between the central and southeastern populations and may be adversely affected by human settlement and agricultural expansion in this area.

- Northeastern (Tarangire NP/Swagaswaga GR) and southwestern (Katavi NP/Rukwa GR) populations of impala and eland appear to have had moderate levels of historical gene flow. Unfortunately, this connectivity is likely to be lost as a result of dense human settlement in northeastern Tanzania. These changes in land use and expansion of human activity will likely also inhibit gene flow between northeastern (Tarangire NP/Swagaswaga GR) and central (Ruaha NP/Rungwa GR) reserves, which were historically connected for all three species.

BIBLIOGRAPHY

Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulinck H, Matthysen E (2003) The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning*, **64**, 233-247.

Aguirre AA, Tabor GM (2008) Global factors driving emerging infectious diseases. *Annals of the New York Academy of Sciences*, **1149**, 1-3.

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, **13**, 1771-1784.

Averbeck C (2002) Population ecology of impala (*Aepyceros melampus*) and community-based wildlife conservation in Uganda. *PhD Dissertation*, Technical University of Munich.

Barnes RFW and Douglas-Hamilton I (1982) The numbers and distribution patterns of large mammals in the Ruaha-Rungwa area of southern Tanzania. *Journal of Applied Ecology*, **19**, 411-425.

Beier P, Majka DR, Spencer WD (2008) Forks in the road: choices in procedures for designing wildland linkages. *Conservation Biology*, **22**, 836-851.

Bender DJ, Contreras TA, Fahrig L (1998) Habitat loss and population decline: a meta-analysis of the patch size effect. *Ecology*, **79**, 517-533.

Bhebhe E, Kogi J, Holder DA, Arevalo E, Derr JN, Linn RA, Ruvuna F, Davis SK, Taylor JF (1994) Caprine microsatellite dinucleotide repeat polymorphisms at the SR-CRSP-6, SR-CRSP-7, SR-CRSP-8, SR-CRSP-9 and SR-CRSP-10 loci. *Animal Genetics*, **25**, 203.

Bishop MD, Kappes SM, Keele JW, Stone RT, Sunden SLF, Hawkins GA, Solinas Toldo S, Fries R, Grosz MD, Yoo J, Beattie CW (1994) A genetic linkage map for cattle. *Genetics*, **136**, 619-639.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21-45.

Borner M (1985) The increasing isolation of Tarangire National Park. *Oryx*, **19**, 91-96.

Brinkman TJ, Schwartz MK, Person DK, Pilgrim KL, Hundertmark KJ (2010) Effects of time and rainfall on PCR success using DNA extracted from deer fecal pellets. *Conservation Genetics*, **11**, 1547-1552.

Broad S, Mulliken T, Roe D (2003) The nature and extent of legal and illegal trade in wildlife. pp 3-22 *In* S. Oldfield (ed.) The Trade in Wildlife. Earthscan Publications, Sterling, VA, USA

Brown DM, Brenneman RA, Koepfli KP, Pollinger JP, Milá B, Georgiadis NJ, Louis Jr EE, Grether GF, Jacobs DK, Wayne RK (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57. doi: 10.1186/1741-7007-5-57.

Bruford MW, Cheesman DJ, Coote T, Green HAA, Haines SA, O'Ryan C, Williams TR (1996) Microsatellites and their application to conservation genetics. pp 278-297 *In* TB Smith and RK Wayne (eds.) Molecular genetic approaches in conservation. Oxford University Press, Inc., New York, New York, USA.

Buchanan FC, Crawford AM (1992) Ovine dinucleotide repeat polymorphism at the MAF209 locus. *Animal Genetics*, **23**, 183.

Buchanan FC, Crawford AM (1993) Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Animal Genetics*, **24**, 145.

Burgess ND, Fjeldsa J, Botterweg R (1998) Faunal importance of the Eastern Arc Mountains of Kenya and Tanzania. *Journal of East African Natural History*, **87**, 37-58.

Caro TM (2008) Decline of large mammals in the Katavi-Rukwa ecosystem of western Tanzania. *African Zoology*, **43**, 99-116.

Caro TM, Pelkey N, Borner M, Campbell KLI, Woodworth BL, Farm BP, Kuwai JO, Huish SA, Severre ELM (1998) Consequences of different forms of conservation for large mammals in Tanzania: preliminary analyses. *African Journal of Ecology*, **36**, 303-320.

Castellano S, Balletto E (2002) Is the partial mantel test inadequate? *Evolution*, **56**, 1871-1873.

Clavero M, García-Berthou E (2005) Invasive species are a leading cause of animal extinctions. *Trends in Ecology and Evolution*, **20**, 110.

Clevenger AP, Wierzchowski J, Chruszcz B, Gunson K (2002) GIS-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. *Conservation Biology*, **16**, 503-514.

Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology*, **13**, 2841-2850.

Cushman SA, Landguth EL (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology*, **19**, 3592-3602.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, **168**, 486-499.

Dagg AI (1971) *Giraffa camelopardalis*. *Mammalian Species*, **5**, 1-8.

East R (1998) African antelope database. IUCN Species survival commission.

Eblate EM, Lughano KJ, Sebastian CD, Peter ML, Knut RH (2011) Polymorphic microsatellite markers for genetic studies of African antelope species. *African Journal of Biotechnology*, **10**, 11817-11820.

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, **16**, 551-558.

Epps CW, Mutayoba BM, Gwin L, Brashares JS (2011) An empirical evaluation of the African elephant as a focal species for connectivity planning in East Africa. *Diversity and Distributions*, **17**, 603-612.

Epps CW, Palsbøll PJ, Wehausen JD, Roderick GK, Ramey II RR, McCullough DR. (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, **8**, 1029-1038**.**

Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology*, **44**, 714-724.

Estes AB, Kuemmerle T, Kushnir H, Radeloff VC, Shugart HH (2012) Land-cover change and human population trends in the greater Serengeti ecosystem from 1984-2003. *Biological Conservation*, **147**, 255-263.

Estes RD (1992) The behavior guide to African mammals. The University of California Press, Berkeley, California.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

FAO (2005) Multipurpose Africover databases on environmental resources. http://www.africover.org/index.htm. Accessed October 8, 2005.

Fennessy J, Brown D (2010) *Giraffa camelopardalis*. In: IUCN 2011. IUCN Red List of threatened species. Version 2011.1. www.iucnredlist.org. Downloaded on 18 July 2011.

Flagstad Ø, Syvertsen PO, Stenseth NC, Stacy JE, Olsaker I, Røed KN, Jakobsen KS (2000) Genetic variability in Swayne's hartebeest, an endangered antelope. *Conservation Biology*, **14**, 254-264.

Frankel OH, Soulé ME (1981) Conservation and evolution. University Press, Cambridge, UK.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genetic Research Cambridge*, **66**, 95-107.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. University Press, Cambridge, UK.

Georges M, Massey JM (1992) Polymorphic DNA markers in *Bovidae*. World Intellectual Property Organization. Geneva. Patent W092/13102.

Ginnett TF, Demment MW (1999) Sexual segregation by Masai giraffes at two spatial scales. *African Journal of Ecology*, **37**, 93-106.

Graham MD, Douglas-Hamilton I, Adams WM, Lee PC (2009) The movement of African elephants in a human-dominated land-use mosaic. *Animal Conservation*, **12**, 445-455.

Hartl DL, Clark AG (1997) Principles of population genetics. Sinauer Associates, Massachusetts.

Hebblewhite M, Haydon DT (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B*, **365**, 2303-2312.

Hedrick PW (1999) Perspective: highly variable genetic loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313-318.

Heller R, Lorenzen ED, Okello JBA, Masembe C, Siegismund HR (2008) Mid-Holocene decline in African buffalos inferred from Bayesian coalescent-based analyses of microsatellites and mitochondrial DNA. *Molecular Ecology*, 17, 4845-4858.

Heller R, Okello JBA, Siegismund H (2010) Can small wildlife conservancies maintain genetically stable populations of large mammals? Evidence for increased genetic drift in geographically restricted populations of Cape buffalo in East Africa. *Molecular Ecology*, **19**, 1324-1334.

Hillman JC (1974) Ecology and behavior of the wild eland. *News – African Wildlife Leadership Foundation of Washington, D.C.*, **9**, 6-9.

Hillman JC (1988) Home range and movement of the common eland (*Taurotragus oryx* Pallas 1766) in Kenya. *African Journal of Ecology*, **26**, 135-148.

Holzhauer SIJ, Ekschmitt K, Sander A, Dauber J, Wolters V (2006) Effect of historic landscape change on the genetic structure of the bush-cricket *Metrioptera roeseli*. *Landscape Ecology*, **21**, 891-899.

Huebinger RM, Pierson DJ, De Maar TW, Brown DM, Brenneman RA, Louis Jr EE (2002) Characterization of 16 microsatellite marker loci in the Maasai giraffe (*Giraffa camelopardalis tippelskirchi*). *Molecular Ecology Notes*, **2**, 531-533.

Huebinger RM, De Maar TWJ, Woodruff LH, Pomp D, Louis Jr EE (2006) Characterization of nine microsatellite loci in impala (*Aepyceros melampus*). *Molecular Ecology Notes*, **6**, 1152-1153.

Hurt R, Ravn P (2000) Hunting and its benefits: an overview of hunting in Africa with special reference to Tanzania. pp. 295-314 *In* HHT Prins, JG Groutenhuis and TT Dolan (eds.) Wildlife conservation by sustainable use. Kluwer Academic Publishers, Boston, Massachusetts, USA.

IUCN (2002) IUCN technical guidelines on the management of ex-situ populations for conservation. *14th meeting of the programme committee of council*, Gland, Switzerland.

IUCN (2011) IUCN Red List of threatened species. Version 2011.1. <www.iucnredlist.org>. Downloaded on 18 July 2011.

IUCN SSC Antelope Specialist Group (2008a) *Aepyceros melampus*. *In*: IUCN Red List of threatened species. Version 2011.1. <www.iucnredlist.org>. Downloaded on 18 July 2011.

IUCN SSC Antelope Specialist Group (2008b) *Tragelaphus oryx*. *In*: IUCN Red List of threatened species. Version 2011.1. <www.iucnredlist.org>. Downloaded on 18 July 2011.

Jennings TN, Knaus BJ, Mullins TD, Haig SM, Cronn RC (2011) Multiplexed microsatellite recovery using massively parallel sequencing. *Molecular Ecology Resources*, **11**, 1060-1067.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187-189.

Kalinowski ST, ML Taper, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099-1106.

Kingdon J (1997) The Kingdon field guide to African mammals. Academic Press, San Diego, California.

Kweka J, Morrissey O, Blake A (2003) The economic potential of tourism in Tanzania. *Journal of International Development*, **15**, 335-351.

Lada H, Thomson JR, Mac Nally R, Taylor AC (2008) Impacts of massive landscape change on a carnivorous marsupial in south-eastern Australia: inferences from landscape genetics analysis. *Journal of Applied Ecology*, **45**, 1732-1741.

Leader-Williams N (2000) The effects of a century of policy and legal change on wildlife conservation and utilisation in Tanzania. pp. 219-246 *in* HHT Prins, JG Groutenhuis and TT Dolan (eds.) Wildlife conservation by sustainable use. Kluwer Academic Publishers, Boston, Massachusetts, USA.

Leuthold BM, Leuthold W (1978) Ecology of the giraffe in Tsavo East National Park, Kenya. *East African Wildlife Journal*, **16**, 1-20.

Lorenzen ED, Arctander P, Siegismund HR (2006) Regional genetic structuring and evolutionary history of the impala *Aepyceros melampus*. *Journal of Heredity*, **97**, 119-132.

Lorenzen ED, Masembe C, Arctander P, Siegsmund HR (2010) A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa:

insights from mtDNA and the common eland antelope. *Journal of Biogeography*, **37**, 571-581.

Low PS (2005) Climate change and Africa. Cambridge University Press, Cambridge UK.

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey USA.

Maddox JF (2001) Mapping the matrix metalloproteinase 9 (MMP-9) gene and the BL 1071 microsatellite to ovine chromosome 13 (OAR13). *Animal Genetics*, **32**, 329-331.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.

Mantyka-pringle CS, Martin TG, Rhodes JR (2012) Interactions between climate and habitat loss effects on biodiversity: a systematic review and meta-analysis. *Global Change Biology*, **18**, 1239-1252.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7,** 639-655.

McNaughton SJ, Georgiadis NJ (1986) Ecology of African grazing and browsing mammals. *Annual Review of Ecology and Systematics*, **17**, 39-65.

McRae BH (2006) Isolation by resistance. *Evolution*, **60**: 1551-1561.

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19885-19890.

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, **89**, 2712-2724.

Mitchell J, Keane J, Laidlaw J (2009) Making success work for the poor: package tourism in Northern Tanzania. Overseas Development Institute report.

Mulder MB, Caro T, Msago OA (2007) The role of research in evaluating conservation strategies in Tanzania: the case of the Katavi-Rukwa ecosystem. *Conservation Biology*, **21**, 647-658.

Munson L, Terio KA, Kock R, Mlengeya T, Roelke ME, Dubovi E, Summers B, Sinclair ARE, Packer C (2008) Climate extremes promote fatal co-infections

during canine distemper epidemics in African lions. *Public Library of Science ONE*, **3**, e2545. doi:10.1371/journal.pone.0002545.

Murphy MA, Kendall KC, Robinson A, Waits LP (2007) The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. *Conservation Genetics*, **8**, 1219-1224.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858.

Navidi W, Arnheim N, Waterman MS (1992) A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American Journal of Human Genetics*, **50**, 347-352.

Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283-292.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **155**, 583-590.

Nei M (1986) Definition and estimation of fixation indices. *Evolution*, **40**, 643-645.

Nersting LG, Arctander P (2001) Phylogeography and conservation of impala and greater kudu. *Molecular Ecology*, **10**, 711-719.

Newmark WD (1996) Insularization of Tanzanian parks and the local extinction of large mammals. *Conservation Biology*, **10**, 1549-1556.

Newmark WD (2008) Isolation of African protected areas. *Frontiers in Ecology and Environment* , **6**, 321-328.

Peakall R, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

Pemberton JM, Slate J, Bancroft DR, Barret JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249-252.

Penty JM, Henry HM, Ede AJ, Crawford AM (1993) Ovine microsatellites at the OarAE16, OarAE54, OarAE57, OarAE119 and OarAE129 loci. *Animal Genetics*, **24**, 219.

Peterson D (1978) Seasonal distribution and interactions of cattle and wild ungulates in Masailand, Tanzania. *MSc Thesis.* Virginia Polytechnic Institute.

Pettorelli N, Lobora AL, Msuha MJ, Foley C, Durant AM (2010) Carnivore biodiversity in Tanzania: revealing the distribution patterns of secretive mammals using camera traps. *Animal Conservation*, **13**, 131-139.

Prins HHT, Weyerhaeuser FJ (1987) Epidemics in populations of wild ruminants: anthrax and impala, rinderpest and buffalo in Lake Manyara National Park, Tanzania. *Oikos*, **49**, 28-38.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Pritchard JK, Wen X, Falush D (2007) Documentation for *structure* software: version 2.2.

Raufaste N, Rousset F (2001) Are partial mantel tests adequate? *Evolution*, **55**, 1703-1705.

Raymond M, Rousset F (1995) GENEPOP (version 3.4): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.

Riginos C, Grace JB (2008) Savanna tree density, herbivores, and the herbaceous community: bottom-up vs. top-down effects. *Ecology*, **89**, 2228-2238.

Robertson HG (2002) Comparison of leaf litter ant communities in woodlands, lowland forests and montane forests of north-eastern Tanzania. *Biodiversity and Conservation*, **11**, 1637-1652.

Rodgers WA (1974) The lion (*Panthera leo*, Linn.) population of the eastern Selous Game Reserve. *African Journal of Ecology*, **12**, 313-317.

Roe D, Mulliken T, Milledge S, Mremi J, Mosha S, Grieg-Gran M (2002) Making a killing or making a living? Wildlife trade, trade controls and rural livelihoods. Biodiversity and livelihoods issues, **6**, *International Institute for Environment and Development*, London, UK

Rosen GE, Smith KF (2010) Summarizing the evidence on the international trade in illegal wildlife. *EcoHealth*, **7**, 24-32.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. *In*: Krawetz S, Misener S (eds.) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa NJ, pp 365-386.

Sawyer SC, Epps CW, Brashares JS (2011) Placing linkages among fragmented habitats: do least-cost models reflect how animals use landscapes? *Journal of Applied Ecology*, **48**, 668-678.

Schwab P, Debes PV, Witt T, Hartl GB, Hmwe SS, Zachos FE, Grobler JP (2011) Genetic structure of the common impala (Aepyceros melampus melampus) in South Africa: phylogeography and implications for conservation. *Journal of Zoological Systematics and Evolutionary Research*, doi: 10.1111/j.1439-0469.2011.00638.x

Shaffer ML (1981) Minimum population sizes for species conservation. *BioScience*, **31**, 131-134.

Shah VB, McRae B (2008) Circuitscape: a tool for landscape ecology. *Proceedings of the 7th Python in science conference*. In Varoquax G, Vaught T, and Millman J (eds.), pp 62-66.

Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI (2010). Inferring landscape effects on gene flow: a new model selection framework. *Molecular Ecology*, **19**, 3603-3619.

Simonsen BT, Siegismund HR, Arctander P (1998). Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. *Molecular Ecology*, **7**, 225-237.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627-632.

Soulé ME, Wilcox BA, Holtby C (1979) Benign neglect: a model of faunal collapse in the game reserves of East Africa. *Biological Conservation*, **15**, 259-272.

Steffen P, Eggen A, Dietz AB, Womack JE, Stranzinger G, Fries R (1993) Isolation and mapping of polymorphic microsatellites in cattle. *Animal Genetics*, **24**, 121-124.

Stuart C, Stuart T (2000) A field guide to the tracks and signs of southern and east African wildlife. Struik Publishers, Cape Town, South Africa.

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **24**, 3189-3194.

Tallmon DA, Koyuk A, Luikart GH, Beaumont MA (2008) ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, **8**, 299-301.

TANAPA (1994) National policies for national parks in Tanzania. Tanzania National Parks National Policy committee.

Tanzania Mammal Atlas Project, accessed 15 February 2012. Tanzania Wildlife Research Institute and Zoological Society of London. http://www.tanzaniamammals.org/index.php

United Nations Environment Programme-World Conservation Monitoring Centre (2007) World Database of Protected Areas. UNEP-WCMC, IUCN.

Vanacker V, Linderman M, Lupo F, Flasse S, Lambin E (2005) Impact of short-term rainfall fluctuation on interannual land cover change in sub-Saharan Africa. *Global Ecology and Biogeography*, **14**, 123-135.

Van Bommel FPJ, Heitkōnig IMA, Epema GF, Ringrose S, Bonyongo C, Veenendaal EM (2006) Remotely sensed habita indicators for predicting distribution of impala (*Aepyceros melampus*) in the Okavango Delta, Botswana. *Journal of Tropical Ecology*, **22**, 101-110.

Van Hooft WF, Groen AF, Prins HHT (2000) Microsatellite analysis of genetic diversity in African buffalo (*Syncerus caffer*) populations throughout Africa. *Molecular Ecology*, **9**, 2017-2025.

Waltert M, Meyer B, Shanyangi MW, Balozi JJ, Kitwara O, Qolli S, Krischke H, Mühlenberg M (2008) Foot surveys of large mammals in woodlands of western Tanzania. *Journal of Wildlife Management*, **72**, 603-610.

Walther GR, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin JM, Hoegh-Guldburg O, Bairlein F (2002) Ecological responses to recent climate change. *Nature*, **416**, 389-395.

Waples RS and Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequlibrium. *Molecular Ecology Resources*, **8**, 753-756.

Wehausen JD, Ramey II RR, Epps CW (2004) Experiments in DNA extraction and PCR amplification from bighorn sheep feces: the importance of DNA extraction method. *Journal of Heredity*, **95**, 503-509.

Weir BS (1996) Genetic Data Analysis II: Methods for discrete population genetic data. Sinauer Associates, Sunderland, MA, USA.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow – F_{ST} does not equal 1/(4Nm+1). *Heredity*, **82**, 117-125.

World Bank, accessed January 13, 2011. www.worldbank.org

Wright S (1921) Systems of mating. *Genetics*, **6**, 111-178.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

WWF (2001) Central Zambezian miombo woodlands. http://www.worldwildlife.org/wildworld/profiles/terrestrial/at/at0704_full.html. Accessed 18 July 2011.

Zinyowera MC, Jallow BP, Maya RS, Okoth-Ogendo HWO (1998) Africa. *In*: Watson RT, Zinyowera MC, and RH Moss (eds.). The regional impacts of climate change. Cambridge University Press, Cambridge UK.

Table 1. Life history information for focal species of population genetic study of southern and central Tanzania.

Species	Average body mass (kg)	Home range (km^2)	Water requirements	Territorial	Degree of diet specialization
Maasai Giraffe, *Giraffa camelopardalis tippelskirchi*	828-1200	160 (up to 600)	low	no	moderate – browser
Common Eland *Tragelaphus oryx*	450-670	298 (up to 1500)	low	no	low-moderate – mostly browser
Common Impala *Aepyceros melampus melampus*	40-55	1.3 - 5.8	moderate	yes	very low – mixed grazer/browser

Table 2. National parks and game reserves sampled in Tanzania, with year gazetted as a protected area and current area (km^2). Year gazetted refers to the earliest record of preservation of the area; many reserves have changed status (e.g., game reserve to national park) since their establishment. Area (km^2) includes only the main reserve and not that of adjacent protected areas, and is the most recent estimate available.

Location	Classification	Year Gazetted	Area (km^2)
Katavi	National Park	1912†	4 270
Mikumi	National Park	1964	3 230
Ruaha	National Park	1964	10 300
Rukwa	Game Reserve	1997	4 194
Rungwa	Game Reserve	1946	4 500
Selous	Game Reserve	1922	45 000
Swagaswaga	Game Reserve	1996	817
Tarangire	National Park	1957‡	2 850

† Areas around Lake Chada and Lake Katavi were gazetted as a game reserve in 1912; Katavi NP was established in 1974

‡Tarangire game reserve was established in 1957; it was reclassified as Tarangire NP in 1970

Table 3. Sample overview and summary genetic statistics for (a) giraffe, (b) impala, and (c) eland across all sampling localities in Tanzania. Statistics were combined over all loci (giraffe: 14 loci; impala: 11 loci; eland: 8 loci). Sample sizes (n), sampling polygon areas (in km^2), average expected heterozygosity (H_E), total numbers of alleles across all loci (Global *A*), average number of alleles per locus (average *A*), and allelic richness (A_R) are shown for each population in each species. For allelic richness, populations were subsampled to the minimum number of alleles observed at any locus in any population.

a) GIRAFFE

Location	Abbreviation	n	Sampling polygon area (km^2)	H_E	Global *A*	Average *A*	A_R (n=7)
Katavi NP	KNP	33	1578	0.526	61	4.36	3.41
Mikumi NP	MNP	22	222	0.612	68	4.86	3.94
Ruaha NP	RNP	40	2853	0.578	71	5.07	3.63
Rukwa GR	RkGR	26	2196	0.566	75	5.36	3.90
Rungwa GR	RgGR	9	567	0.602	63	4.50	4.24
Selous GR	SeGR	23	391	0.499	52	3.71	3.10
Swagaswaga GR	SwGR	14	231	0.579	64	4.57	3.94
Tarangire NP	TNP	15	1503	0.597	63	4.50	4.00
Village Lands	VILL	4	-	-	-	-	-

b) IMPALA

Location	Abbreviation	n	Sampling Polygon Area (km^2)	H_E	Global A	Average A	A_R (n=5)
Katavi NP	KNP	22	1124	0.701	60	5.46	4.08
Mikumi NP	MNP	15	408	0.603	57	5.18	3.85
Ruaha NP	RNP	22	2071	0.692	60	5.46	4.05
Rukwa GR	RkGR	18	376	0.713	59	5.36	4.15
Selous GR	SeGR	26	732	0.634	61	5.54	3.71
Swagaswaga GR	SwGR	7	211	0.663	45	4.09	3.80
Tarangire NP	TNP	21	1075	0.653	61	5.54	3.85
Village Lands	VILL	2	-	-	-	-	-

c) ELAND

Location	Abbreviation	n	Sampling Polygon Area (km^2)	H_E	Global A	Average A	A_R (n=4)
Katavi NP, Rukwa GR‡	KNP/RkGR	8	752	0.728	50	6.25	4.32
Mikumi NP	MNP	7	123	0.613	30	3.75	3.02
Ruaha NP	RNP	22	1142	0.753	66	8.25	4.22
Tarangire NP	TNP	25	870	0.747	69	8.62	4.15
Rungwa GR	RgGR	3	-	-	-	-	-
Selous GR	SeGR	1	-	-	-	-	-
Swagaswaga GR	SwGR	4	-	-	-	-	-

‡Samples from Katavi NP and Rukwa GR were combined in analyses of frequency statistics because of low sample sizes. Sampling locations for these reserves were in close proximity and separated by contiguous protected habitat.

Table 4. Primer sequences, locus-specific annealing temperatures (T_A), observed amplicon sizes, global allelic richnesses and primer references for microsatellite loci for population genetic study of (a) giraffe, (b) impala, and (c) eland.

a) GIRAFFE

Locus	Primer Sequence	T_A (˚C)	Product size (bp)	Number of alleles	Reference
11HDZ443	F: CAT AAA ATT AAA AGG CAC TTG TTC C	54	138-150	7	Huebinger *et al*. 2002
	R: ATG GGG GTC ACA AAG AGT CTG				Huebinger *et al*. 2002
11HDZ550	F: GGA CAG TGG ACT AGG AGA AAA GG	54	172-186	7	Huebinger *et al*. 2002
	R: GCC TGG GAT TCC TGG TAA AC				Huebinger *et al*. 2002
11HDZ665	F: GCC CCT TGC CTA GCT TAA C	60	198-230	7	Huebinger *et al*. 2002
	R: CCG ACT GTA GAA ATG AAG CG				Huebinger *et al*. 2002
11HDZ073	F: AGA CCT AAT GCC ACC AGA ATG	57	191-195	3	Huebinger *et al*. 2002
	R: CAA GTC ATG GGT GCA GAA TG				Redesigned from Huebinger *et al*. 2002
11HDZ562	F: AAA GAG TTA GAT GCA ACT GAG TGA C	60	109-131	9	Huebinger *et al*. 2002
	R: TTC CTC AGG GCT CAG CAT AG				Redesigned from Huebinger *et al*. 2002
11HDZ1004	F: CTC ATG TCT CTT GCA CTG GC	56	141-161	5	Huebinger *et al*. 2002
	R: GTA ATG GCA TAT TTC ACT CTT TTT C				Huebinger *et al*. 2002
Gica13905	F: CAG ACA GAT GGG GAA ACT GAG	63	222-236	6	Novel primer[1]
	R: TTT GGC TAA ATT TTT CAT ACA CAC A				Novel primer[1]

a) GIRAFFE (CONTINUED).

Locus	Primer Sequence	T_A (˚C)	Product size (bp)	Number of alleles	Reference
Gica13619	F: CAG GTT TTC ATT GTA TTG CTC TG	60	253-261	4	Novel primer[1]
	R: ATG CAG AAT GGG GGT TAC AG				Novel primer[1]
Gica9976	F: GGG AGG AGA CTG GAT TGT CA	63	256-276	11	Novel primer[1]
	R: AGT GGC TCT CCA AAG CAC AT				Novel primer[1]
Gica16120	F: AAA GTA ATT TGG GCA AAT GTG G	60	134-140	3	Novel primer[1]
	R: TTT GGC CAG TCT TCA GAT CA				Novel primer[1]
Gica16160	F: TGC AGA GCA ATT GCA AAC AT	60	129-171	18	Novel primer[1]
	R: GTG GGC AAC TGT TCA TAG GG				Novel primer[1]
Gica10894	F: TGT TGT CAC TTA CCC GTT TTC C	63	240-256	7	Novel primer[1]
	R: AGA GTC TGG GAT GCA TTT GG				Novel primer[1]
Gica7401†	F: TGA TTG GCC CTG ATT AGC TG	60	120-126	2	Novel primer[1]
	R: GAC AAG AAT GTG GGC ACT CC				Novel primer[1]
Gica9905	F: ATG ATA TTC AGC TGG GCC TCT	60	289-327	16	Novel primer[1]
	R: CCT GAT GGA CAC CAG GTT G				Novel primer[1]
Gica14170	F: GTG AGG TGC CAT CAC CTT CT	63	261-269	5	Novel primer[1]
	R: CAC TGG AGG CAA GTC AAC AA				Novel primer[1]

†Locus Gica7401 showed high frequencies of null alleles and was discarded from population analyses.
[1]: Novel primers were designed for this study using Illumina sequencing of a Maasai giraffe tissue sample.

b) IMPALA

Locus	Primer Sequence	T_A (˚C)	Product size (bp)	Number of alleles	Reference
32HDZ122	F: CCT GTA ACC ACT TTC TTT TCA CC	57	124-142	8	Huebinger *et al*. 2006
	R: GAC CCA ATG ACC CAA AAC C				
32HDZ688	F: CCG AGG AGG AGA AAA AGG TC	55	197-219	12	Huebinger *et al*. 2006
	R: TGT TGT GTA ATC ATC AGT CCC C				
32HDZ707	F: CAT TCC CTG GCC ACT GTC	57	156-164	5	Huebinger *et al*. 2006
	R: CAA GGG GAT AGT GAT GGA AAG				
MAF209	F: TCA TGC ACT TAA GTA TGT AGG ATG CTG	52	125-135	6	Buchanan & Crawford 1992[Ψ]
	R: GAT CAC AAA AAG TTG GAT ACA ACC GTG G				
MCM38	F: TGG TGA ATG GTG CTC TCA TAC CAG	54	113-135	11	Hulme *et al*. 1994[Ψ]
	R: CAG CCA GCA GCC TCT AAA GGA C				
MCM58	F: CTG GGT CTG TAT AAG CAC GTC TCC	57	167-197	12	Hulme *et al*. 1994[Ψ]
	R: CAG AAC AAT AAA CGC TAA ACC AGA GC				
SR-CRSP9	F: AGA GGA TCT GGA AAT GGA ATC	57	116-136	7	Bhebhe *et al*. 1994[Ψ]
	R: GCA CTC TTT TCA GCC CTA ATG				
TGLA122	F: CCC TCC TCC AGG TAA ATC AGC	55	150-164	8	Georges & Massey 1992[Ψ]
	R: AAT CAC ATG GCA AAT AAG TAC ATA C				
32HDZ07	F: TCC CCG TAA AGA GCA GTT G	58	213-227	7	Huebinger *et al*. 2006
	R: AGG AGA CAG GCA AAG AAA GG				
32HDZ9	F: CCA TCC ACT ATC AGC ACC TCC	60	213-231	10	Huebinger *et al*. 2006
	R: CTC CCT CTT CCT TCC CCC				
SR-CRSP8	F: TGC GGT CTG GTT CTG ATT TCA C	52	211-223	7	Bhebhe *et al*. 1994[Ψ]
	R: CCT GCA TGA GAA AGT CGA TGC TTA G				

Ψ: Primers were used in Lorenzen et al. 2006 study of impala

c) ELAND

Locus	Primer Sequence	T_A (˚C)	Product size (bp)	Number of alleles	Reference
OarFCB193	F: TTC ATC TCA GAC TGG GAT TCA GAA AGG C R: GCT TGG AAA TAA CCC TCC TGC ATC CC	60	93-135	16	Buchanan and Crawford 1993
MMP9	F: CTT GCC TTC TCA TGC TGG GAC T R: GTG AGG ATA GCA CTT GGT CTG GCT	58	178-198	4	Maddox 2001
OarAE129	F: AAT CCA GTG TGT GAA AGA CTA ATC CAG R: GTA GAT CAA GAT ATA GAA TAT TTT TCA ACA CC	50	145-159	8	Penty *et al*. 1993
TGLA227	F: CGA ATT CCA AAT CTG TTA ATT TGC T R: ACA GAC AGA AAC TCA ATG AAA GCA	50	89-107	10	Georges & Massey 1992
ETH225	F: GAT CAC CTT GCC ACT ATT TCC T R: ACA TGA CAG CCA GCT GCT ACT	58	133-159	12	Steffen *et al*. 1993
BM804	F: CCA GCA TCA ACT GTC AGA GC R: GGC AGA TTC TTT GCC TTC TG	60	139-167	13	Bishop *et al*. 1994
OarFCB304	F: CCC TAG GAG CTT TCA ATA AAG AAT CGG R: CGC TGC TGT CAA CTG GGT CAG GG	58	151-171	10	Buchanan and Crawford 1993
BM757	F: TGG AAA CAA TGT AAA CCT GGG R: TTG AGC CAC CAA GGA ACC	58	169-217	17	Bishop *et al*. 1994
TCRBV62‡	F: TGA GTC CTC AGC AAG CAG GT R: ACT GGG ACA CTA CTC CAG CTC TT	60	140-170	6	Crawford *et al*. 1995
MAF209‡	F: TCA TGC ACT TAA GTA TGT AGG ATG CTG R: GAT CAC AAA AAG TTG GAT ACA ACC GTG G	59	123-131	5	Buchanan and Crawford 1992

‡Loci TCRBV62 and MAF209 both deviated from Hardy Weinberg equilibrium proportions in multiple populations and were discarded from population analyses.

Table 5. Hypotheses of landscape effects on gene flow of (a) giraffe, (b) impala, and (c) eland in southern and central Tanzania. For each landscape I generated one or more resistance surfaces with different weighting schemes.

a) GIRAFFE

Landscape Hypothesis	Prediction	Model Name	Weighting Scheme
Distance	Gene flow in giraffe is correlated with geographic distance	IBD	All cells (except lakes) given a resistance value of one
Slope[1]	Gene flow in giraffe is correlated with slope	Slope^(0.2, 0.5, 1, 1.2, 1.5, 1.75, 2)	Slope (in degrees) to a power of 0.2, 0.5, 1, 1.2, 1.5, 1.75, 2
Minimum distance to water[1]	Gene flow in giraffe is not correlated with distance to water	Water^(0.5, 1, 1.5, 2)	Minimum distance to water (in km) to a power of 0.5, 1, 1.5, 2
Woodland density	Moderate woodland densities offer forage, shrublands and grassland do not; woodland is correlated with gene flow	Woods1	Moderate woodland densities (41-60%) weighted as a cost of 1, low (<40%) and high (>60%) density woodlands assigned a cost of 2. All shrublands assigned a cost of 5; grassland assigned a cost of 10.
	Low density woodlands and shrublands offer forage, higher density woodlands have increased predation risk; low density woodland is correlated with gene flow	Woods2	Low density (<40%) woodland assigned a cost of 1; low density shrubland (<40%) assigned a cost of 2. Moderate density (41-60%) woodland and shrubland assigned a value of 5; high density (>60%) woodland and shrubland assigned a cost of 10.

[1]Although I made specific hypotheses about the effect of slope and distance to water on each species, I tested all three species with the entire range of weighting schemes for these landscape features to optimize my models.

b) IMPALA

Landscape Hypothesis	Prediction	Model name	Weighting Scheme
Distance	Gene flow in impala is correlated with geographic distance	IBD	All cells (except lakes) given a resistance value of one
Slope[1]	Gene flow in impala is correlated with slope	Slope^(0.2, 0.5, 1, 1.2, 1.5, 1.75, 2)	Slope (in degrees) to a power of 0.2, 0.5, 1, 1.2, 1.5, 1.75, 2
Minimum distance to water[1]	Gene flow in impala is correlated with distance to water	Water^(0.5, 1, 1.5, 2)	Minimum distance to water (in km) to a power of 0.5, 1, 1.5, 2
Grassland Density	Grassland offers increased forage and decreased predation risk	Grass1	High density grassland areas (>60%) given a cost of 1, moderate grassland (40-60%) given a cost of 5, low density grassland (<40%) given a cost of 10. Areas of predominantly woodland assigned a cost of 20.
	Edge areas preferred for foraging and shelter	Grass2	Low and moderate (<60%) density grassland given a cost of 1; areas with high density grassland (>60%) assigned a cost of 5. Areas of predominantly woodlands given a cost of 10.

c) ELAND

Landscape Hypothesis	Prediction	Model name	Weighting Scheme
Distance	Gene flow in eland is correlated with geographic distance	IBD	All cells (except lakes) given a resistance value of one
Slope[1]	Gene flow in eland is not correlated with slope	Slope^(0.2, 0.5, 1, 1.2, 1.5, 1.75, 2)	Slope (in degrees) to a power of 0.2, 0.5, 1, 1.2, 1.5, 1.75, 2
Minimum distance to water[1]	Gene flow in eland is not correlated with distance to water	Water^(0.5, 1, 1.5, 2)	Minimum distance to water (in km) to a power of 0.5, 1, 1.5, 2
Woodland density	High density woodlands preferred for forage and predator avoidance; shrublands avoided	Woods3	High density woodland areas (>60%) given a cost of 1, moderate woodland (41-60%) given a cost of 2, low density woodland (<40%) given a cost of 3. All shrubland densities assigned a cost of 10. Areas of predominantly grassland assigned a cost of 20.
	Woodland-savannah preferred; grassland and thick shrubland avoided	Woods4	Low and moderate (< 60%) woodland assigned a cost of 1; high density woodland (>60%) assigned a cost of 3. Low density shrubland (<40%) assigned a cost of 5. Moderate and high density (>41%) shrubland assigned a value of 5; high density (>60%) shrubland assigned a cost of 10. Areas of predominantly grassland assigned a cost of 20.

Table 6. Pairwise population differentiation between reserves in Tanzania for (a) giraffe, (b) impala, and (c) eland. Nei's unbiased genetic distances (D_{hat}) are shown above the diagonal; F_{ST} values are below. Katavi NP and Rukwa GR samples were combined for eland analyses.

a) GIRAFFE

	KNP	MNP	RNP	RkGR	RgGR	SeGR	SwGR	TNP
KNP	-	0.22	0.09	0.06	0.07	0.27	0.14	0.19
MNP	0.13	-	0.10	0.15	0.15	0.13	0.11	0.07
RNP	0.07	0.06	-	0.05	0.03	0.22	0.07	0.10
RkGR	0.04	0.09	0.03	-	<0.01	0.19	0.09	0.15
RgGR	0.06	0.08	0.02	<0.01	-	0.19	0.08	0.16
SeGR	0.17	0.09	0.13	0.12	0.12	-	0.17	0.15
SwGR	0.09	0.06	0.04	0.05	0.05	0.12	-	0.06
TNP	0.12	0.04	0.06	0.09	0.08	0.10	0.03	-

b) IMPALA

	KNP	MNP	RNP	RkGR	SeGR	SwGR	TNP
KNP	-	0.30	0.08	0.05	0.38	0.24	0.13
MNP	0.11	-	0.30	0.48	0.07	0.37	0.22
RNP	0.03	0.12	-	0.07	0.35	0.16	0.11
RkGR	0.01	0.15	0.03	-	0.42	0.24	0.22
SeGR	0.13	0.04	0.13	0.13	-	0.43	0.39
SwGR	0.08	0.14	0.06	0.07	0.15	-	0.14
TNP	0.06	0.10	0.05	0.08	0.14	0.06	-

c) ELAND

	KNP/RkGR	MNP	RNP	TNP
KNP/RkGR	-	0.14	0.01	0.11
MNP	0.04	-	0.10	0.27
RNP	<0.01	0.03	-	0.07
TNP	0.03	0.07	0.02	-

Table 7. Estimates of effective population size (*Ne*) of giraffe, impala, and eland in reserves in Tanzania, as estimated by ONeSAMP. Mean estimates are shown, with 95% confidence levels for population size in parentheses. Maximum effective population sizes were set with priors of n = 1000 for giraffe, n = 5000 for impala and n =1500 for eland.

Reserve	Giraffe	Impala	Eland
KNP	25 (19-50)	27 (17-81)	22 (16-66)*
MNP	36 (23-98)	22 (14-104)	7 (5-12)
RNP	44 (30-119)	25 (16-72)	31 (11-82)
RkGR	33 (24-56)	16 (10-37)	*
RgGR	10 (8-16)	n/a	n/a
SeGR	21 (15-35)	63 (32-211)	n/a
SwGR	14 (11-25)	7 (5-15)	n/a
TNP	16 (13-25)	130 (54-394)	135 (64-402)

*KNP and RkGR were assessed jointly for eland.

Table 8. Results from Mantel tests of IBD (isolation by distance) and IBR (isolation by resistance) tests for (a) giraffe, (b) impala, and (c) eland. Pearson's correlation coefficient (*r*) and significance values (*P*) from simple Mantel tests are shown. Where models addressing landscape effects performed better than the null (IBD) model, partial Mantel test results are also shown to parse out the effect of geographic distance.

a) GIRAFFE

			Simple Mantel		Partial Mantel (controlling for geographic distance)	
Landscape Hypothesis	Model name	Test	Correlation *(r)*	*P*	Correlation *(r)*	*P*
Distance	IBD	Genetic distance (D_{hat})*geographic distance	0.72	<0.0001		
Slope	Slope^1	Genetic distance (D_{hat})*IBR_{Slope^1}	0.80	<0.0001	0.52	0.005
Minimum distance to water	Water^0.5	Genetic distance (D_{hat})*$IBR_{Water^0.5}$	0.69	0		
Woodland density	Woods1	Genetic distance (D_{hat})*IBR_{Woods1}	0.68	<0.0001		
	Woods2	Genetic distance (D_{hat})*IBR_{Woods2}	0.74	<0.0001	0.24	0.22

b) IMPALA

			Simple Mantel		Partial Mantel (controlling for geographic distance)	
Landscape Hypothesis	Model name	Test	Correlation *(r)*	*P*	Correlation *(r)*	*P*
Distance	IBD	Genetic distance (D_{hat})*geographic distance	0.48	0.026		
Slope	Slope^2	Genetic distance (D_{hat})*IBR_{Slope^2}	0.91	<0.0001	0.88	<0.0001
Minimum distance to water	Water^1.5	Genetic distance (D_{hat})*$IBR_{Water^1.5}$	0.48	0.027	0.19	0.41
Grassland Density	Grass1	Genetic distance (D_{hat})*IBR_{Grass1}	0.49	0.024	0.14	0.54
	Grass2	Genetic distance (D_{hat})*IBR_{Grass2}	0.43	0.046		

c) ELAND

			Simple Mantel		Partial Mantel (controlling for geographic distance)	
Landscape Hypothesis	Model name	Test	Correlation *(r)*	*P*	Correlation *(r)*	*P*
Distance	IBD	Genetic distance (D_{hat})*geographic distance	0.47	0.37		
Slope	Slope^1	Genetic distance (D_{hat})*IBR_{Slope^1}	0.78	0.029	0.72	0.06
Minimum distance to water	Water^2	Genetic distance (D_{hat})*IBR_{Water^2}	0.92	0.007	0.92	0.007
Woodland density	Woods3	Genetic distance (D_{hat})*IBR_{Woods3}	0.70	0.085	0.60	0.25
	Woods4	Genetic distance (D_{hat})*IBR_{Woods4}	0.55	0.26	0.34	0.57

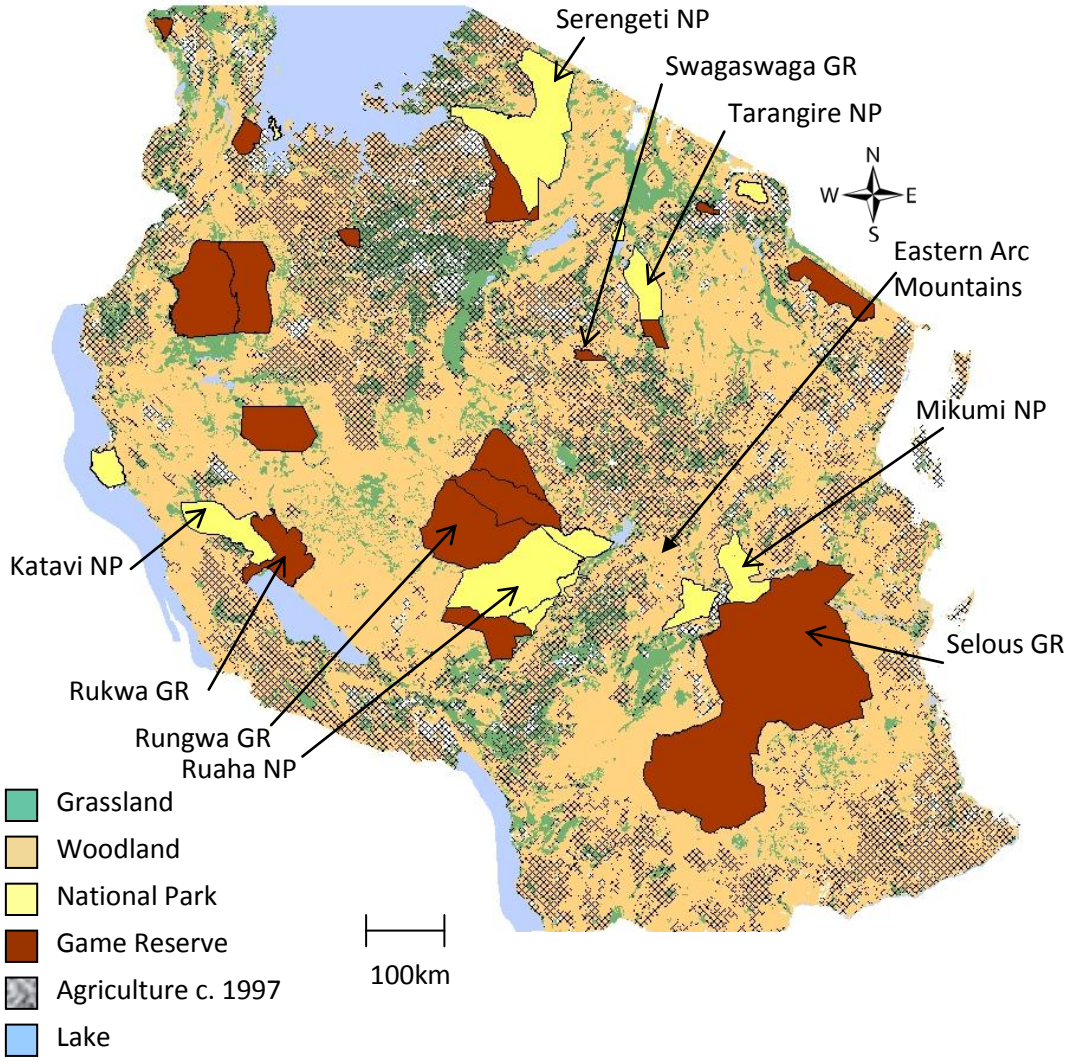


Figure 1. Map of Tanzania showing national parks and game reserves, major lakes, woodland, grassland, and agriculture (c. 1997). Reserves that were sampled in this population genetic study of giraffe, impala, and eland are labelled, as are the Eastern Arc Mountains.

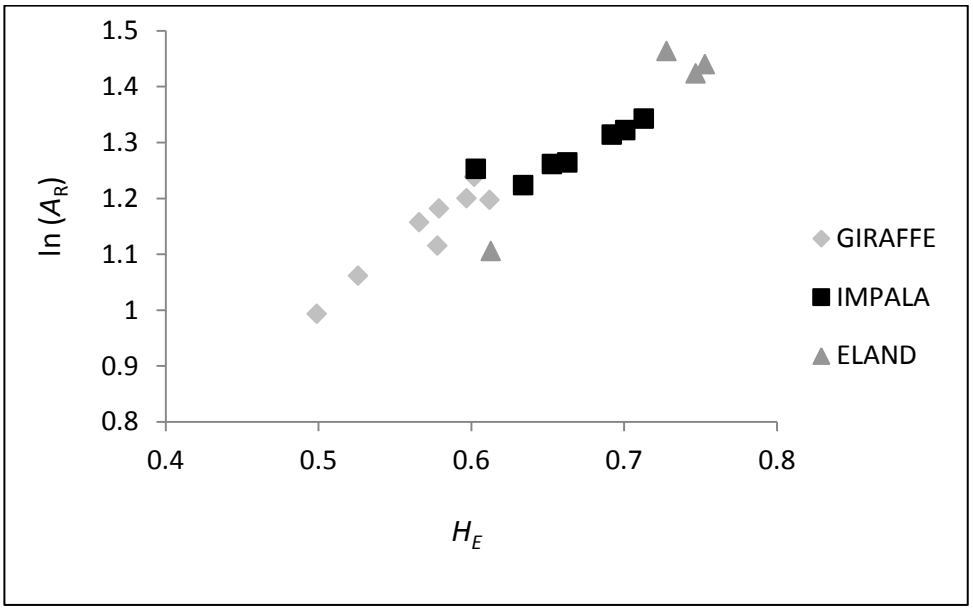
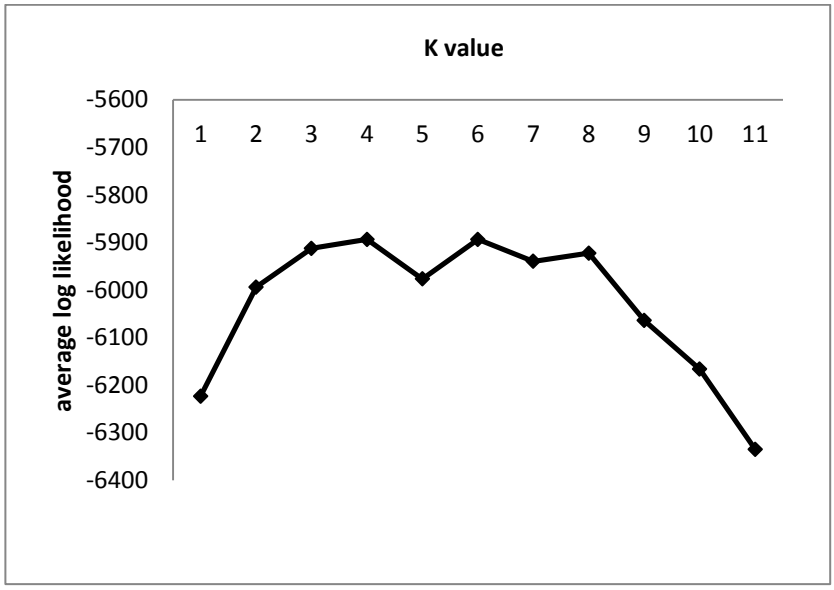


Figure 2. Mean allelic richness (A_R) versus mean expected heterozygosity (H_E) for populations of giraffe, impala and eland, averaged over all loci. Allelic richness was rarefied to eight alleles for all species, and log-transformed.

a) GIRAFFE



b) IMPALA

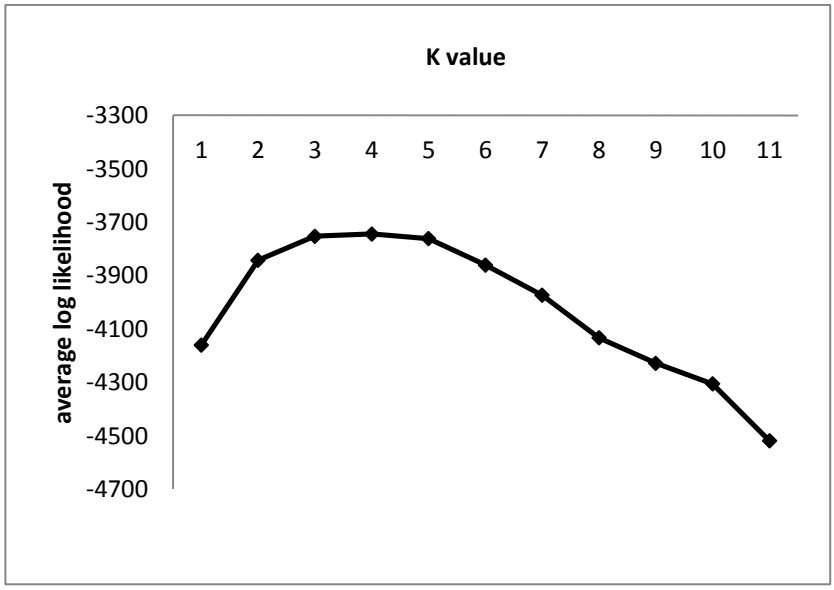


Figure 3. Estimation of the number of genetic clusters (*K*) generated by STRUCTURE using the mean log likelihood averaged over ten runs for each value of *K* for (a) giraffe, (b) impala, and (c) eland. The most likely number of genetic clusters in the data are chosen as the value with the largest average log likelihood.

c) ELAND

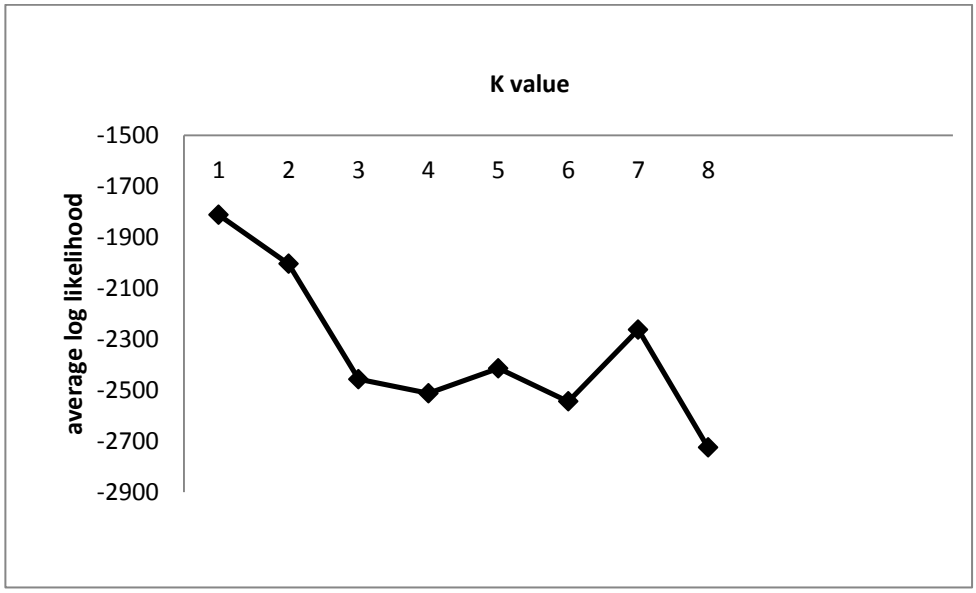


Figure 3. Continued.

a) GIRAFFE

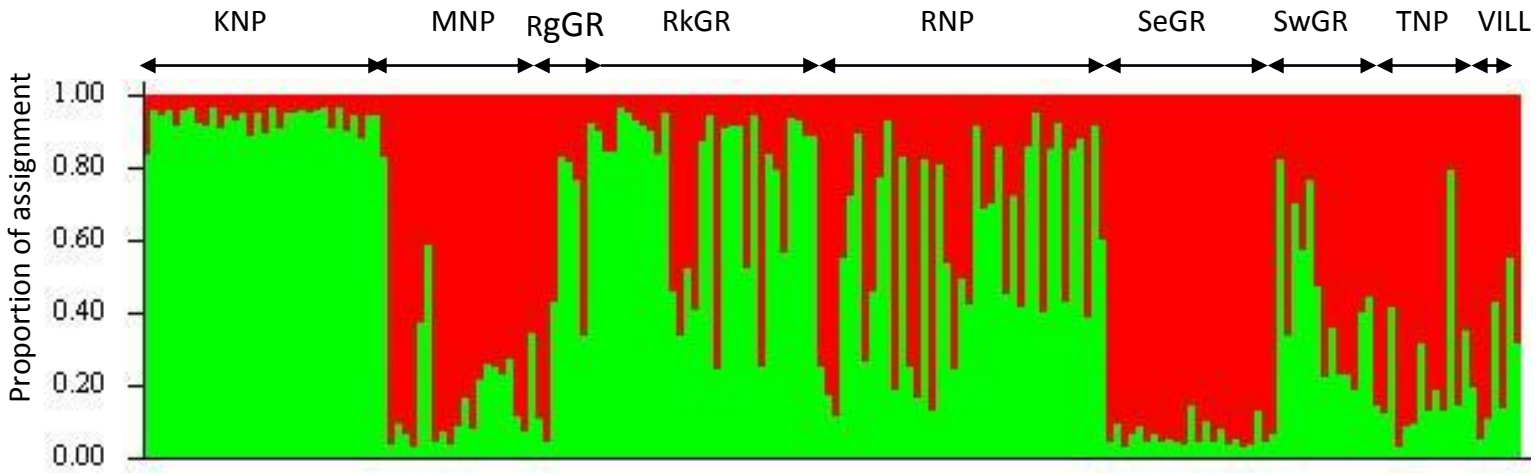


Figure 4. Bar plot of individual sample assignment to genetic clusters (*K* = 2) inferred by STRUCTURE for (a) giraffe (n = 186) and (b) impala (n = 133). Genetic clusters are shown in different colours, and each vertical bar represents one individual.

b) IMPALA

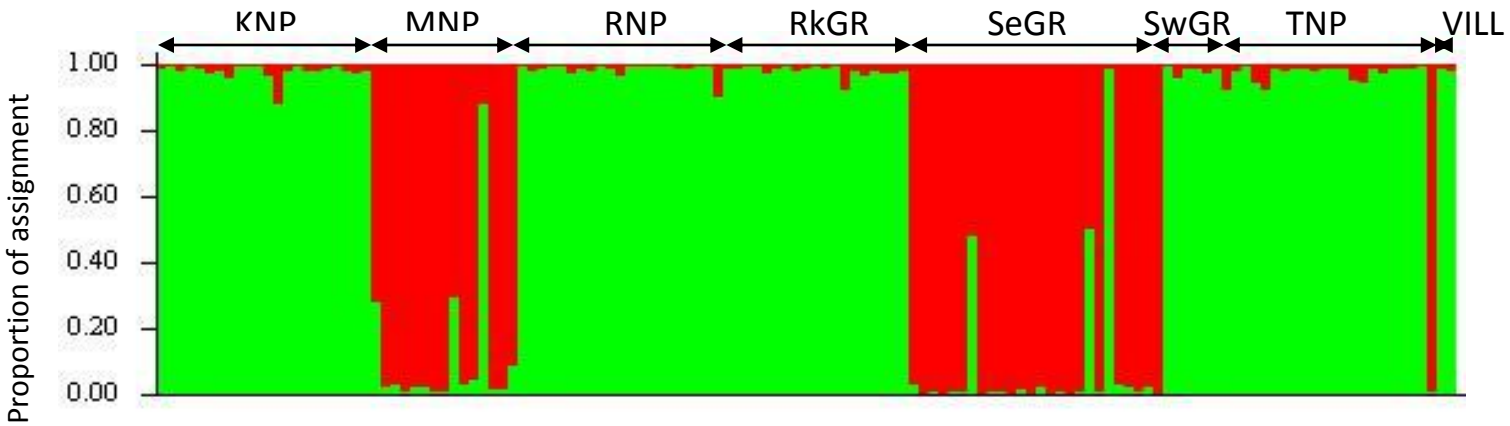


Figure 4. Continued.

a) GIRAFFE

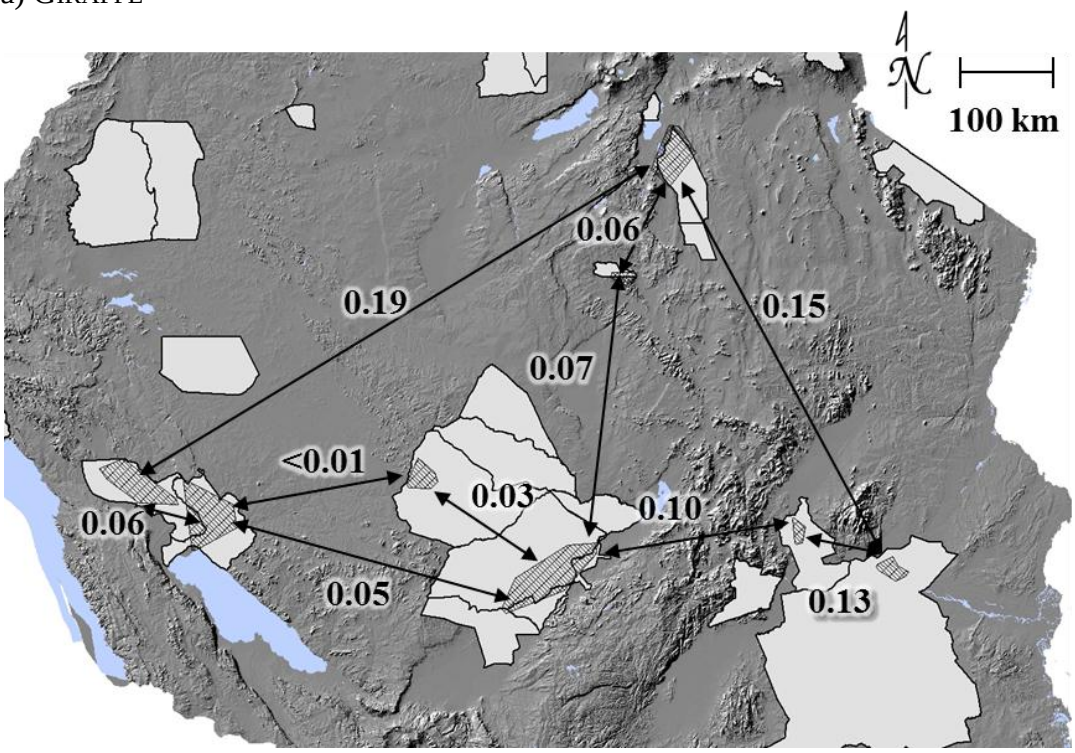
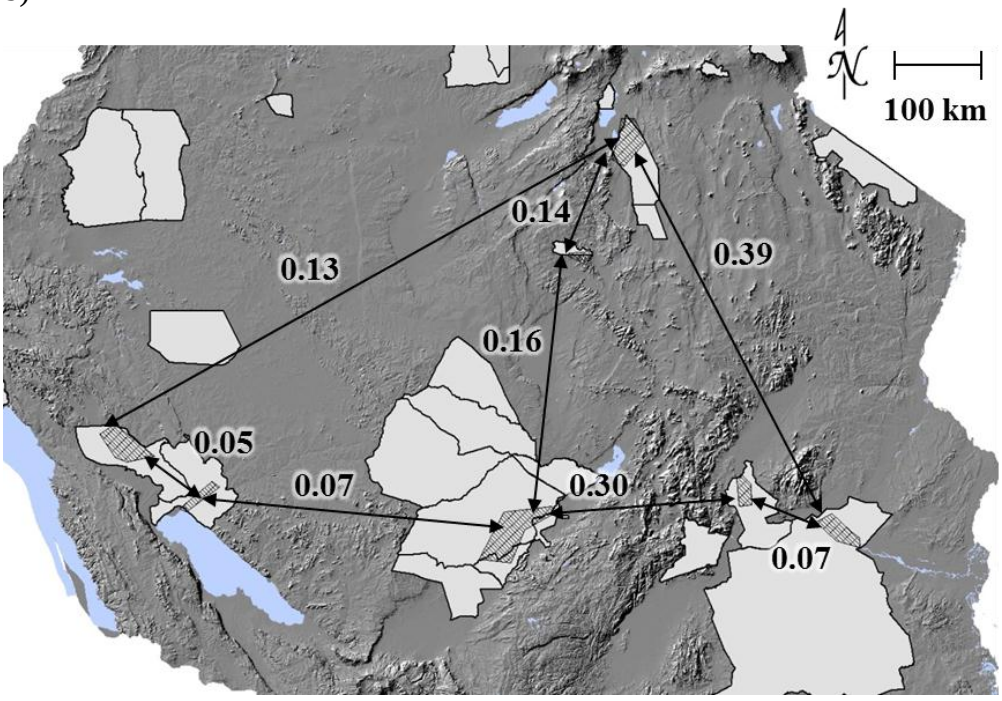


Figure 5. Sampling polygons and interpopulation Nei's unbiased genetic distance (D_{hat}) for (a) giraffe, (b) impala, and (c) eland. Reserves (both national parks and game reserves) are shown in light grey, and sampling polygons within each reserve are shown as hatched areas. Although all population-pairwise distance values were calculated (Table 6), only comparisons of neighbouring reserves are shown.

b) IMPALA



c) ELAND

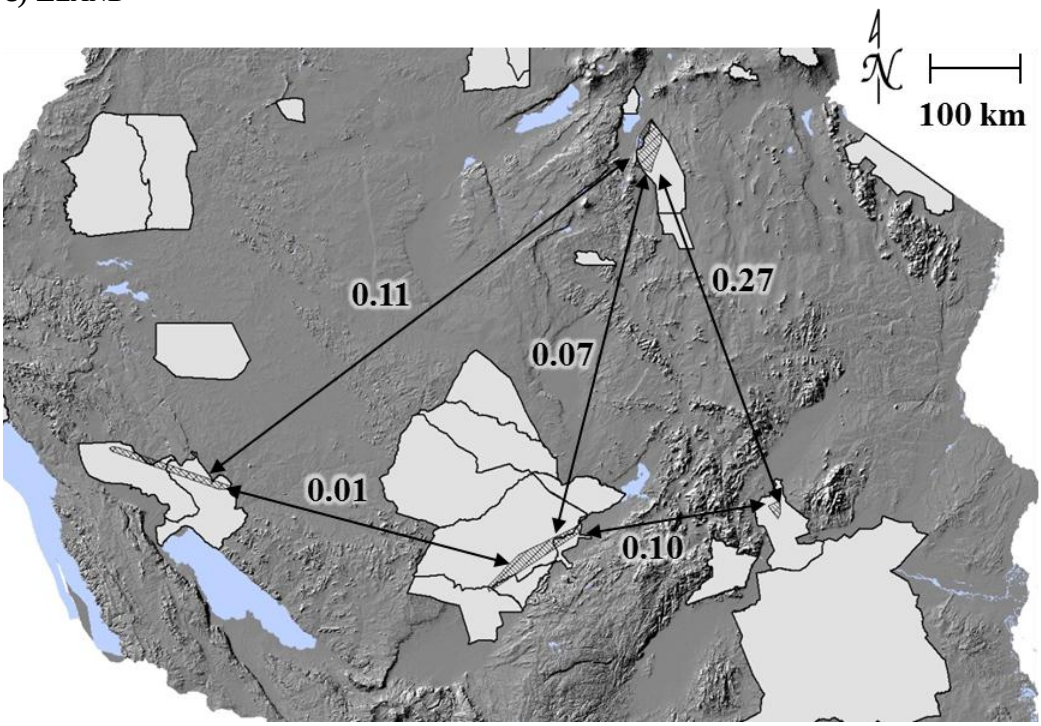
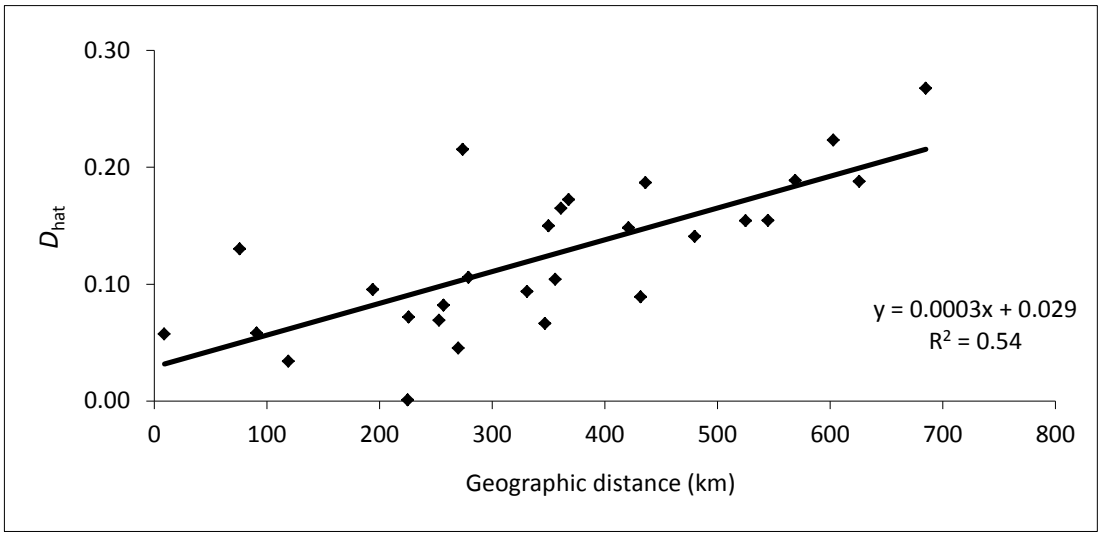


Figure 5. Continued.

a) GIRAFFE



b) IMPALA

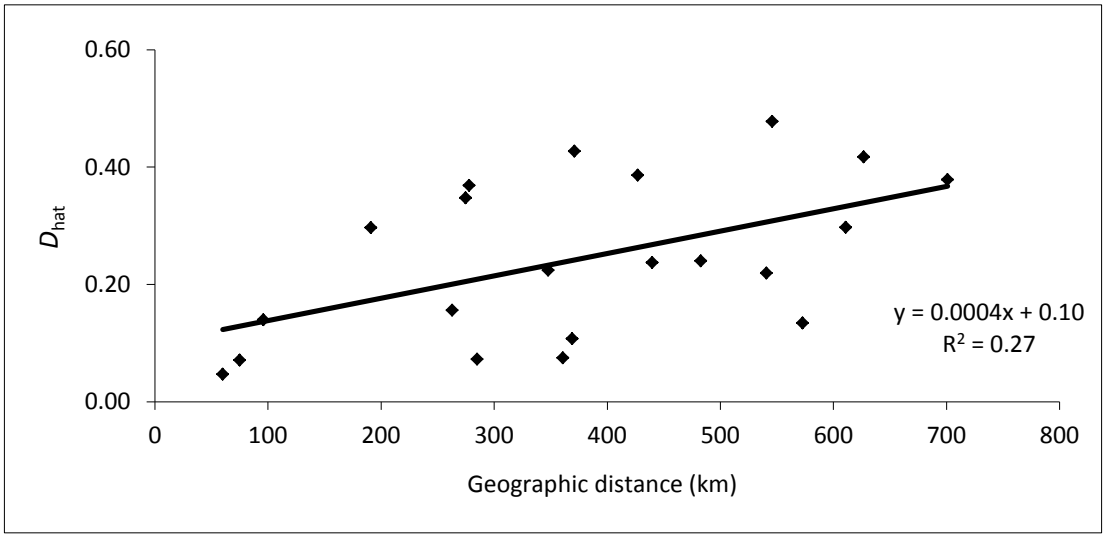


Figure 6. Nei's unbiased genetic distance, (D_{hat}), versus geographic distance (km) for (a) giraffe, (b) impala, and (c) eland. Geographic distance was measured as the minimum distance between edges of sampling polygons for each pair of populations sampled for each species.

c) ELAND

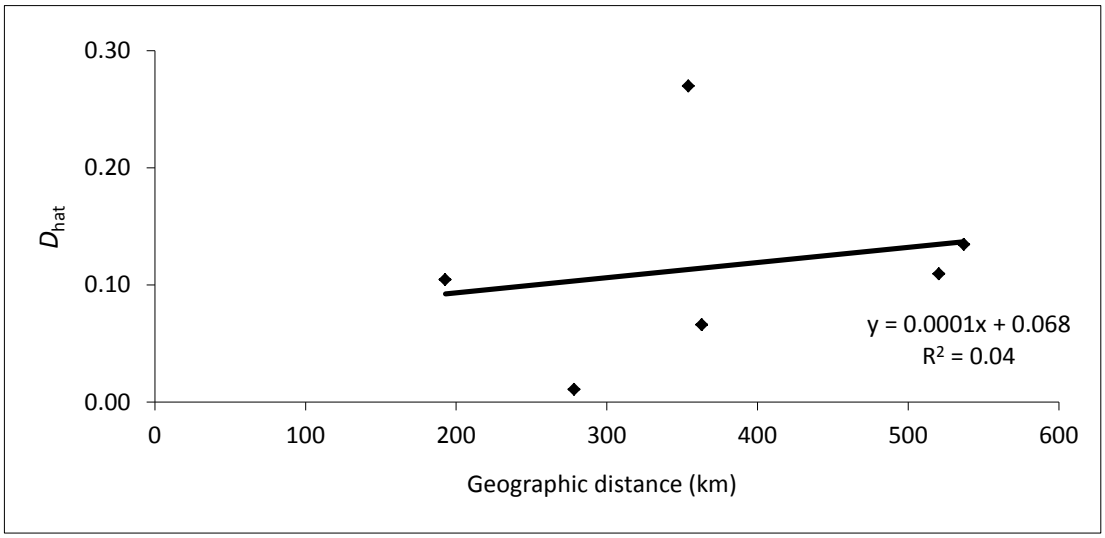


Figure 6. Continued.

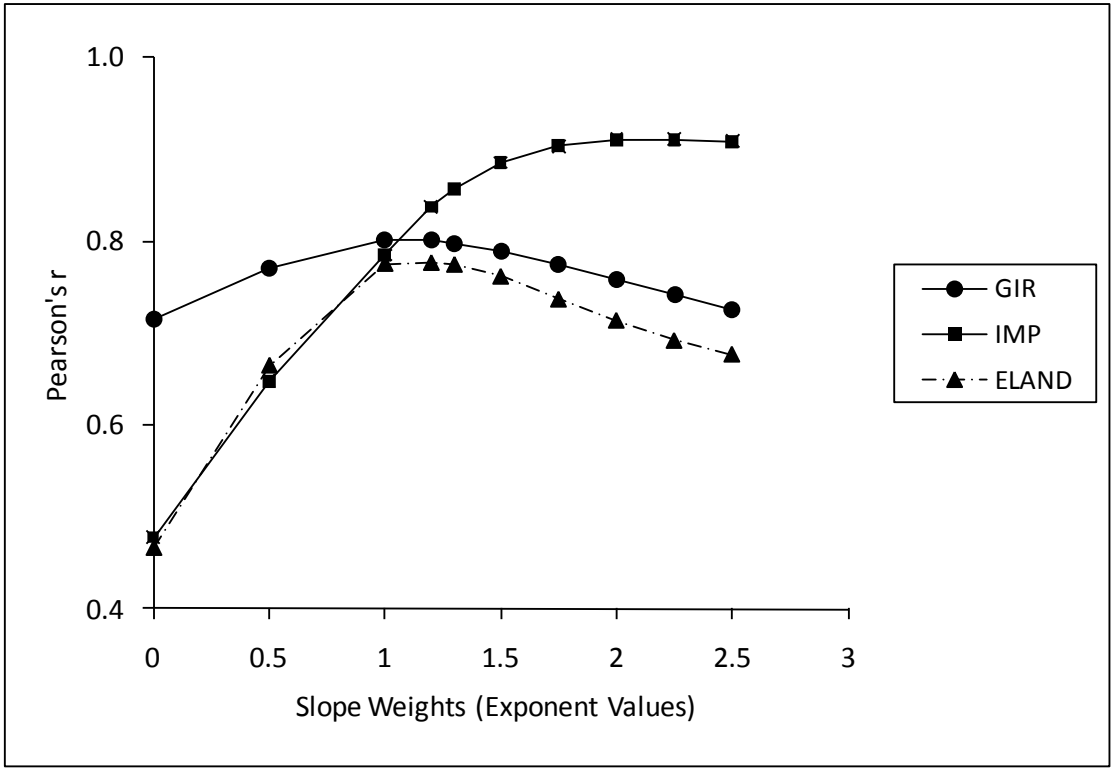
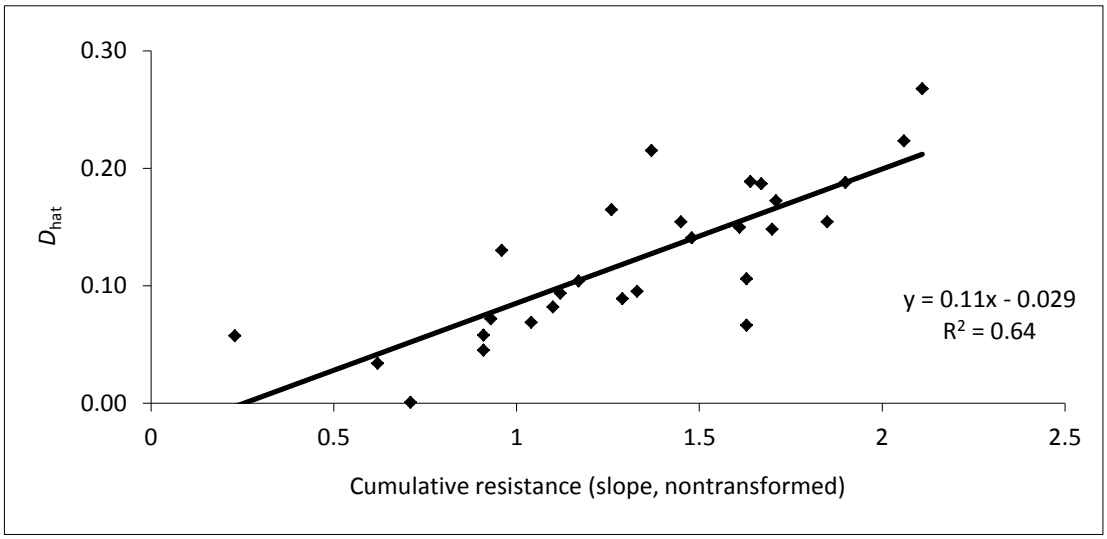


Figure 7. Pearson's correlation coefficient (r) from Mantel tests of the correlation of Nei's unbiased genetic distances (D_{hat}) with cumulative resistance values obtained from slope resistance models. The optimal slope weighting scheme for each species was chosen as the weight that produced resistance values most highly correlated with genetic distance. Slope to a power of zero represents the null model of isolation by distance (no effect of slope). In all three cases models incorporating slope performed better than the null models. In giraffe and eland, nontransformed slope (in degrees) had the highest correlation with D_{hat}. In impala, the squared value of slope was optimal.

a) GIRAFFE



b) IMPALA

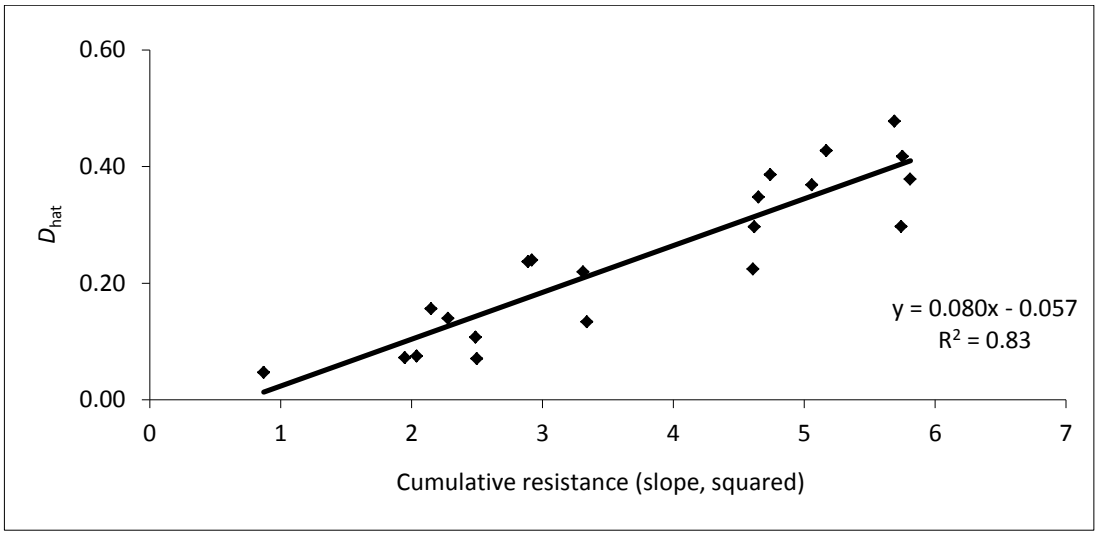
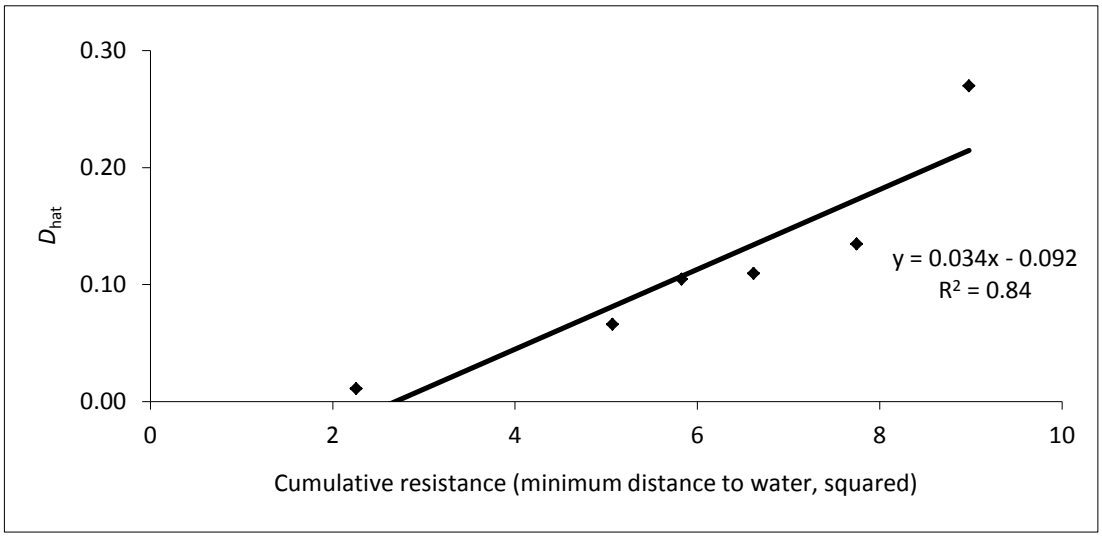


Figure 8. Genetic distance (D_{hat}) versus cumulative resistance values for best landscape models for (a) giraffe (slope in degrees, nontransformed), (b) impala (slope in degrees, squared), (c) eland (minimum distance to water in km, squared) and (d) eland (slope in degrees, nontransformed). For each species, the optimal model was the one that produced cumulative resistance values most highly correlated with genetic distances (D_{hat}). In eland, the correlation between resistance values from the squared value of the minimum distance to water and D_{hat} was highest. However, cumulative resistances from nontransformed slope were also highly correlated with D_{hat} after controlling for geographic distance and distance to water, and thus they are also presented for comparison.

c) ELAND



d) ELAND

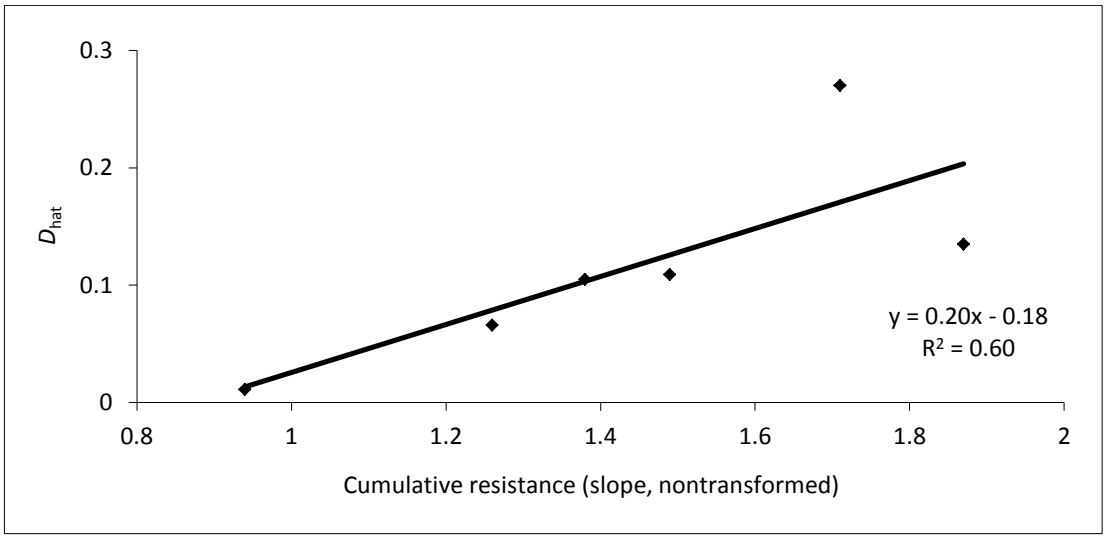
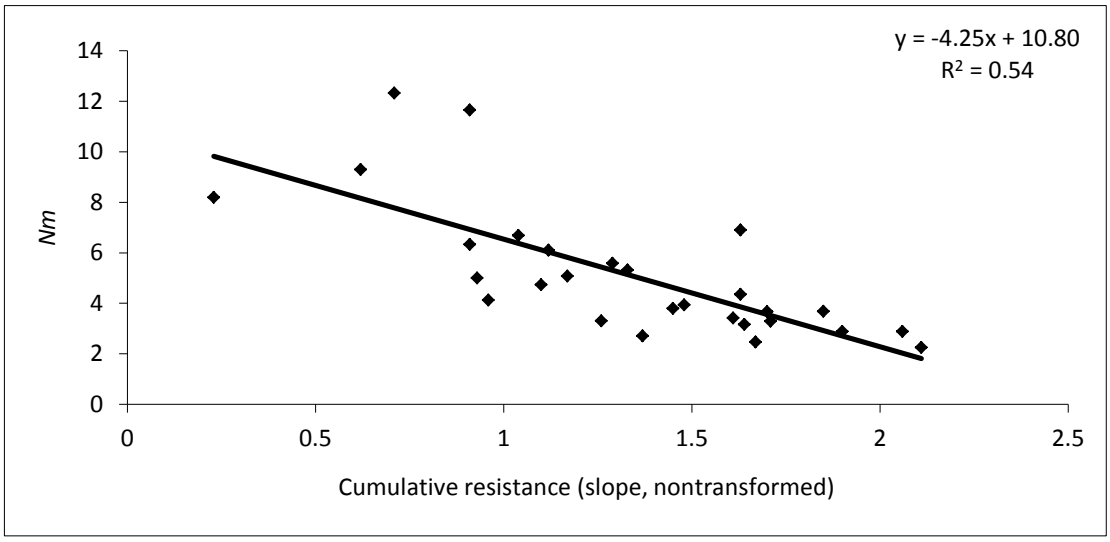


Figure 8. Continued.

a) GIRAFFE



b) IMPALA

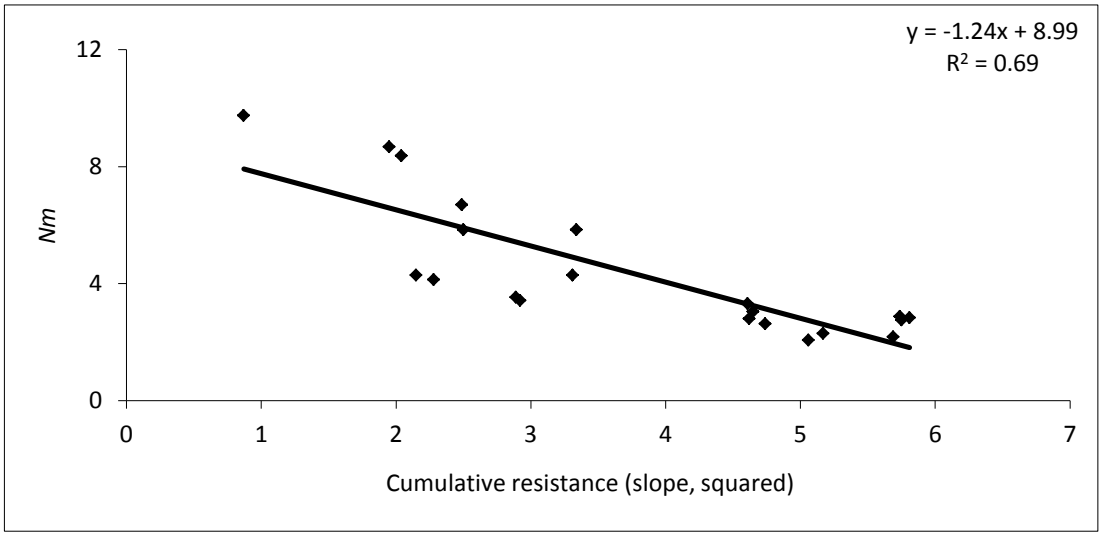
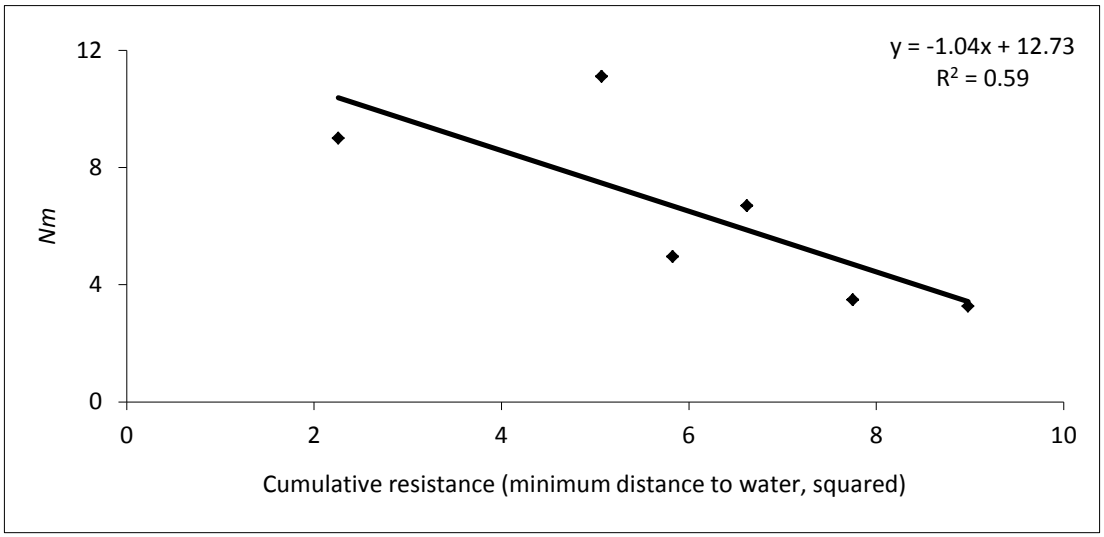


Figure 9. Number of migrants per generation (*Nm*) versus cumulative resistance values from the optimal landscape model for each ungulate species: a) giraffe (slope, nontransformed), b) impala (slope, squared), c) eland (minimum distance to water, squared) and d) eland (slope, nontransformed). *Nm* estimates were calculated from interpopulation F_{ST} values (Wright 1921).

c) ELAND



(d) ELAND

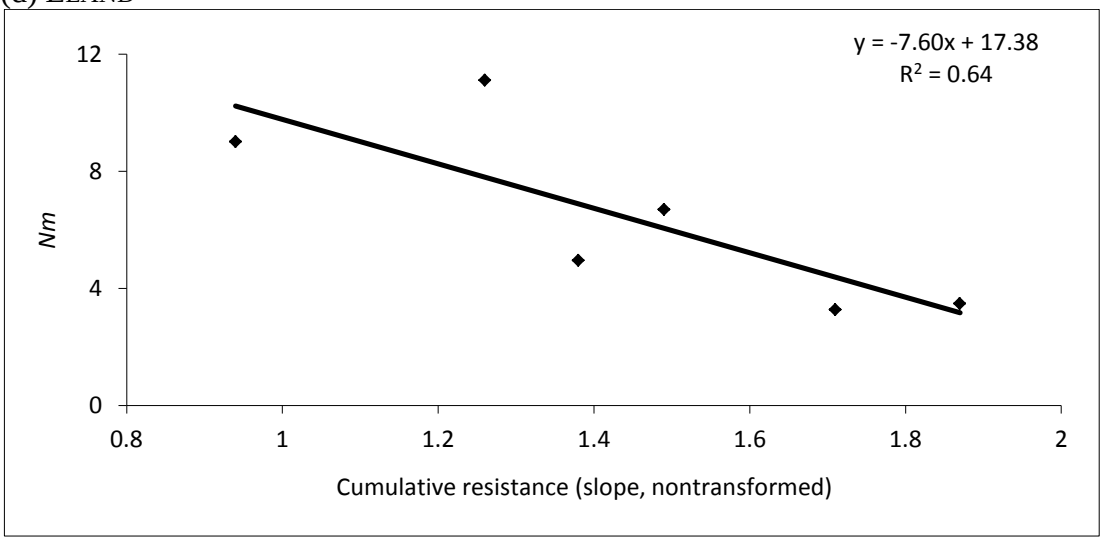


Figure 9. Continued.

a) GIRAFFE

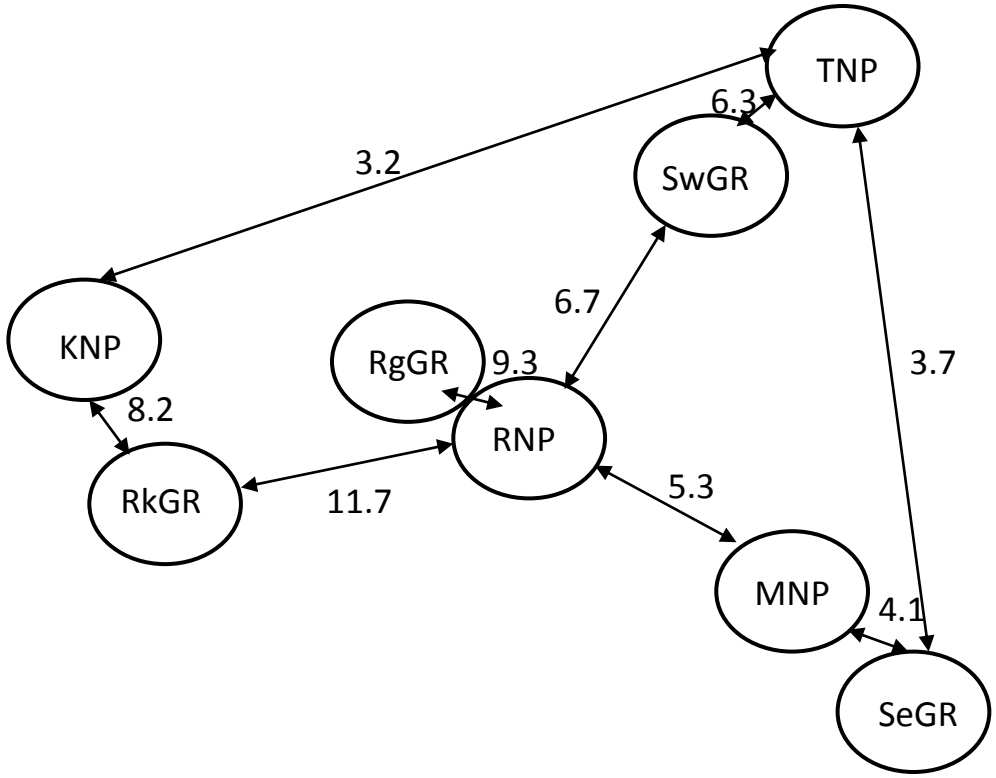


Figure 10. Schematic of the estimated number of migrants (*Nm)* between selected reserves for (a) giraffe, (b) impala, and (c) eland. *Nm* values were inferred from population F_{ST} values (Wright 1921). Although all population-pairwise values of *Nm* were calculated, only comparisons of neighbouring populations are shown.

b) IMPALA

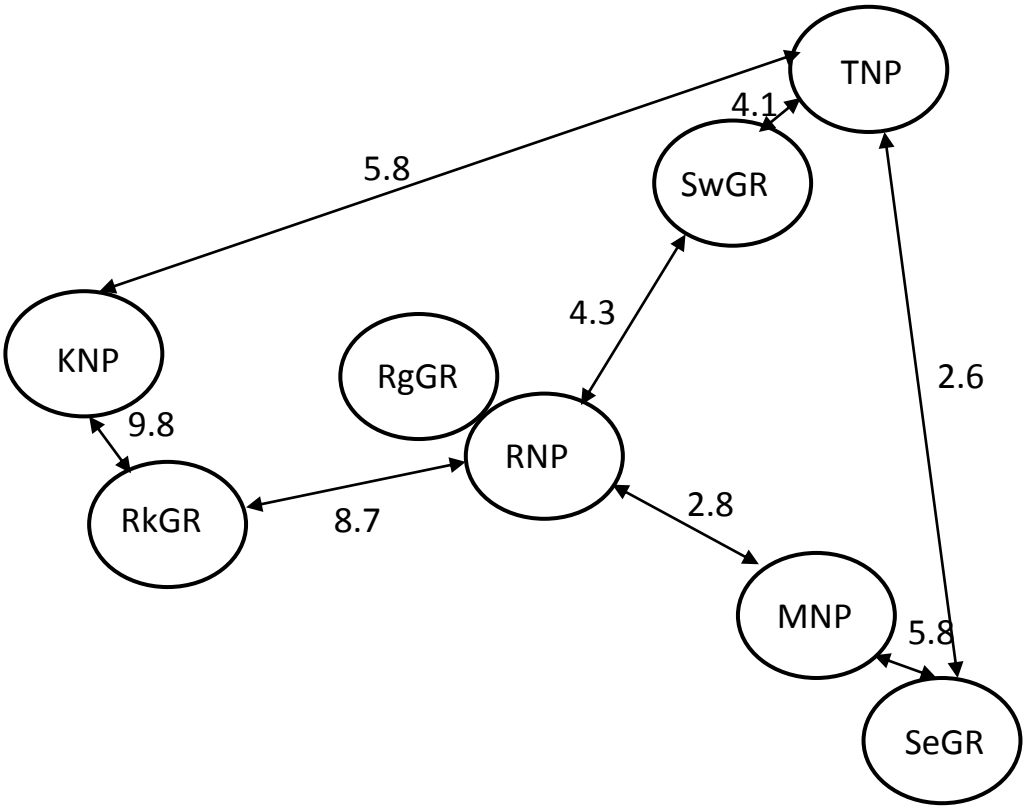


Figure 10. Continued.

c) ELAND

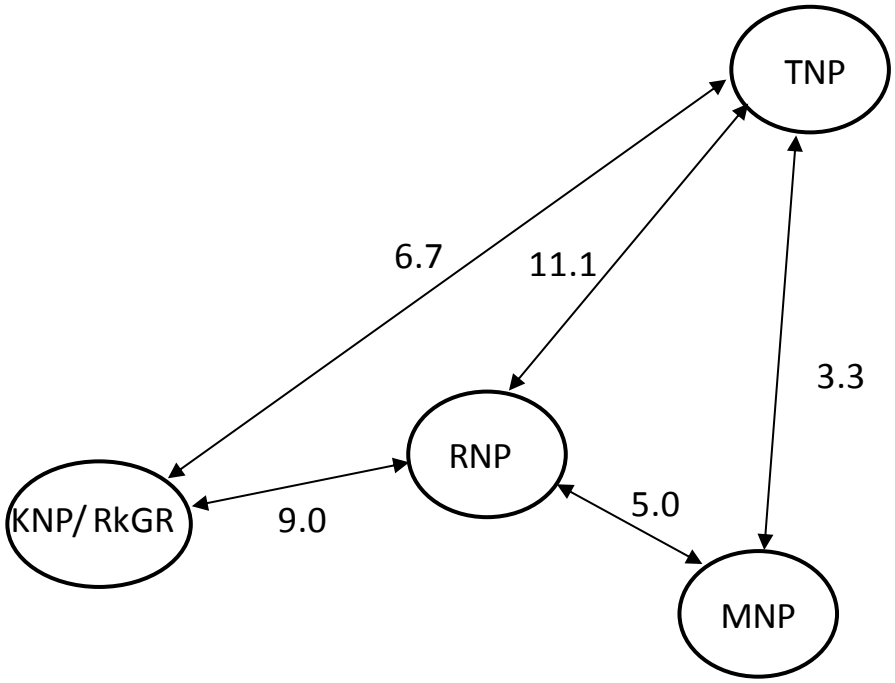


Figure 10. Continued.

a) GIRAFFE

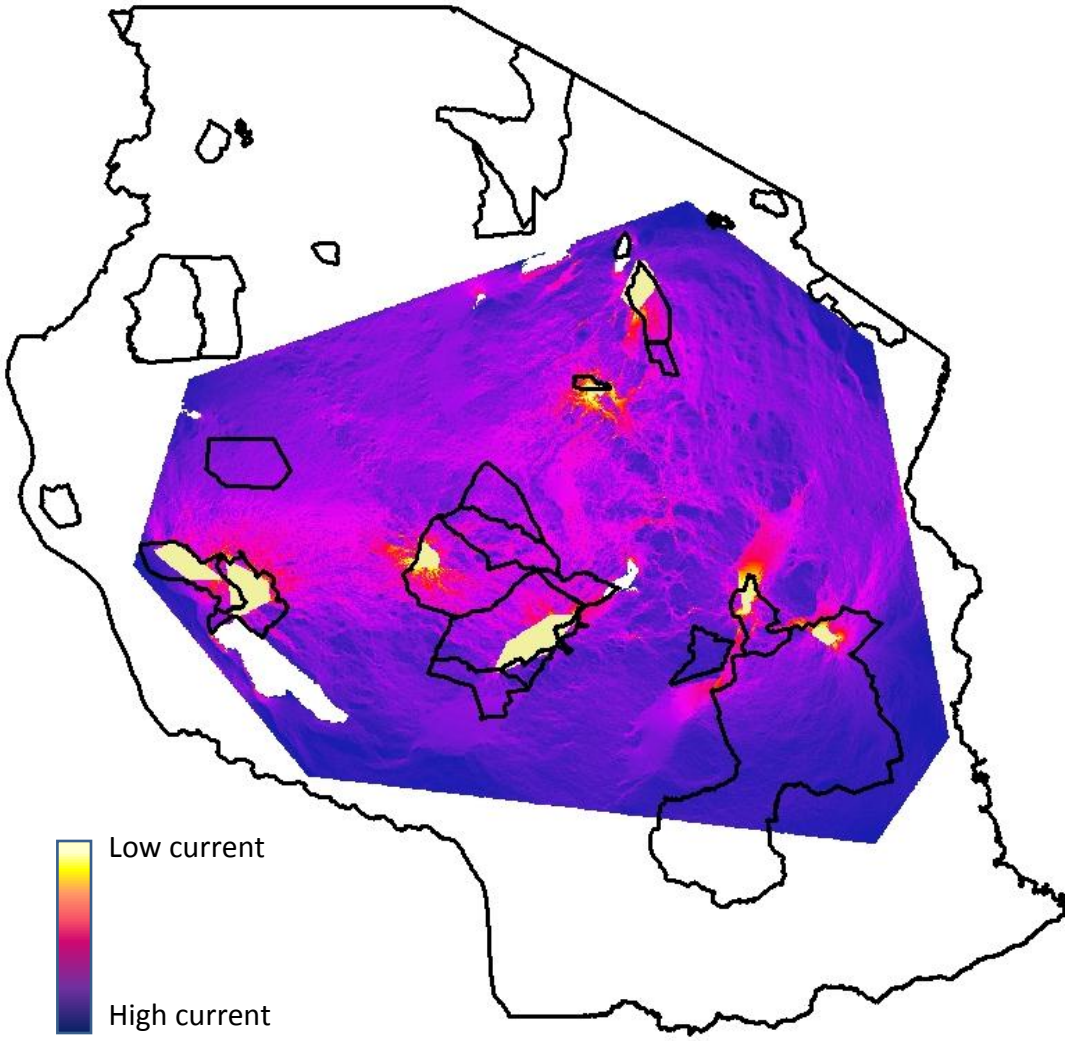


Figure 11. CIRCUITSCAPE current map of best fitting models of landscape resistance to gene flow for three ungulate species in Tanzania. Best fitting model was chosen as the one with cumulative resistance estimates that were most strongly correlated with genetic distances (D_{hat}). Areas with high current are conducive to movement and gene flow. The giraffe model (a) used resistance estimates based on nontransformed values of slope (in degrees). Resistance values in impala (b) were based on the squared value of slope. The eland resistance surface (c) was based on a squared transformation of minimum distance to water (km). For comparison, the surface of current based on nontransformed slope is also shown for eland (d).

b) IMPALA

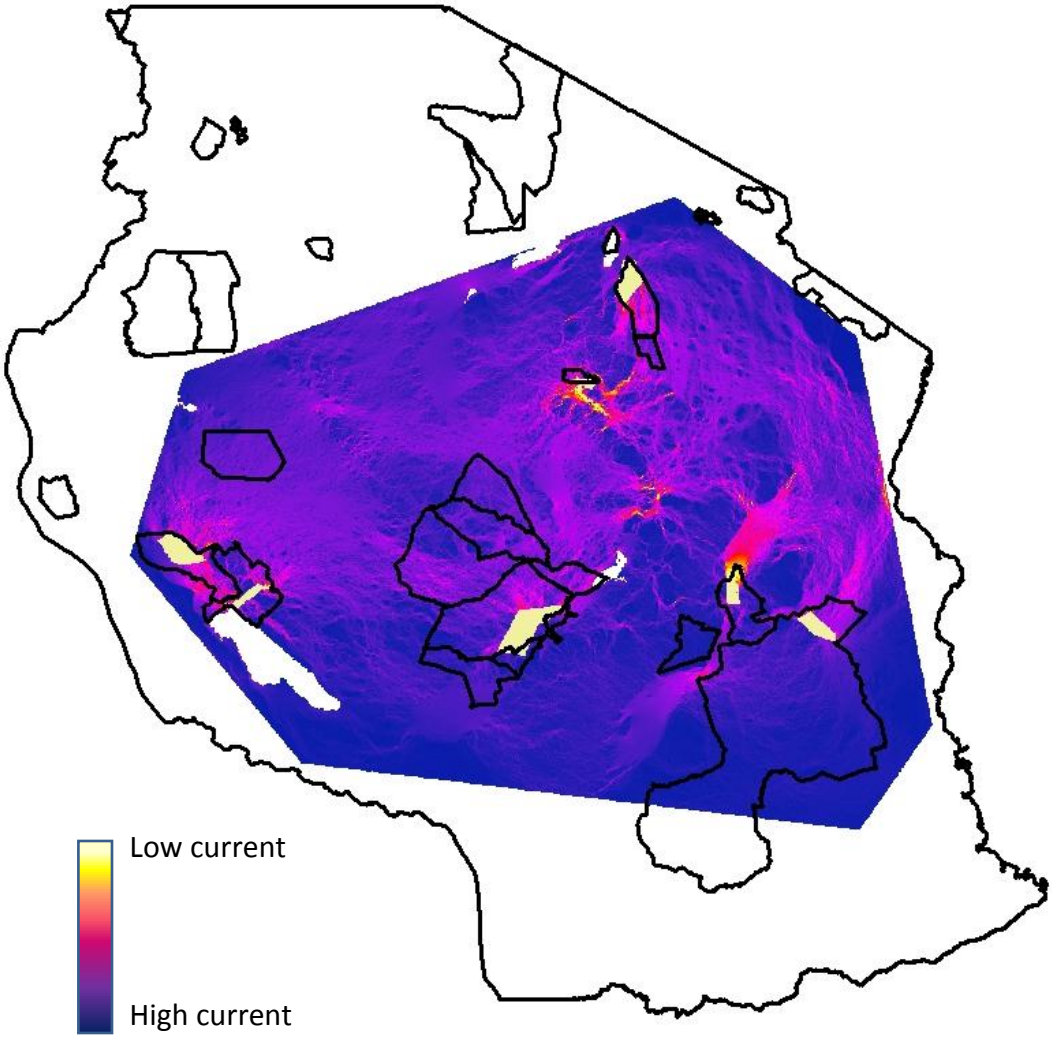


Figure 11. Continued.

c) ELAND

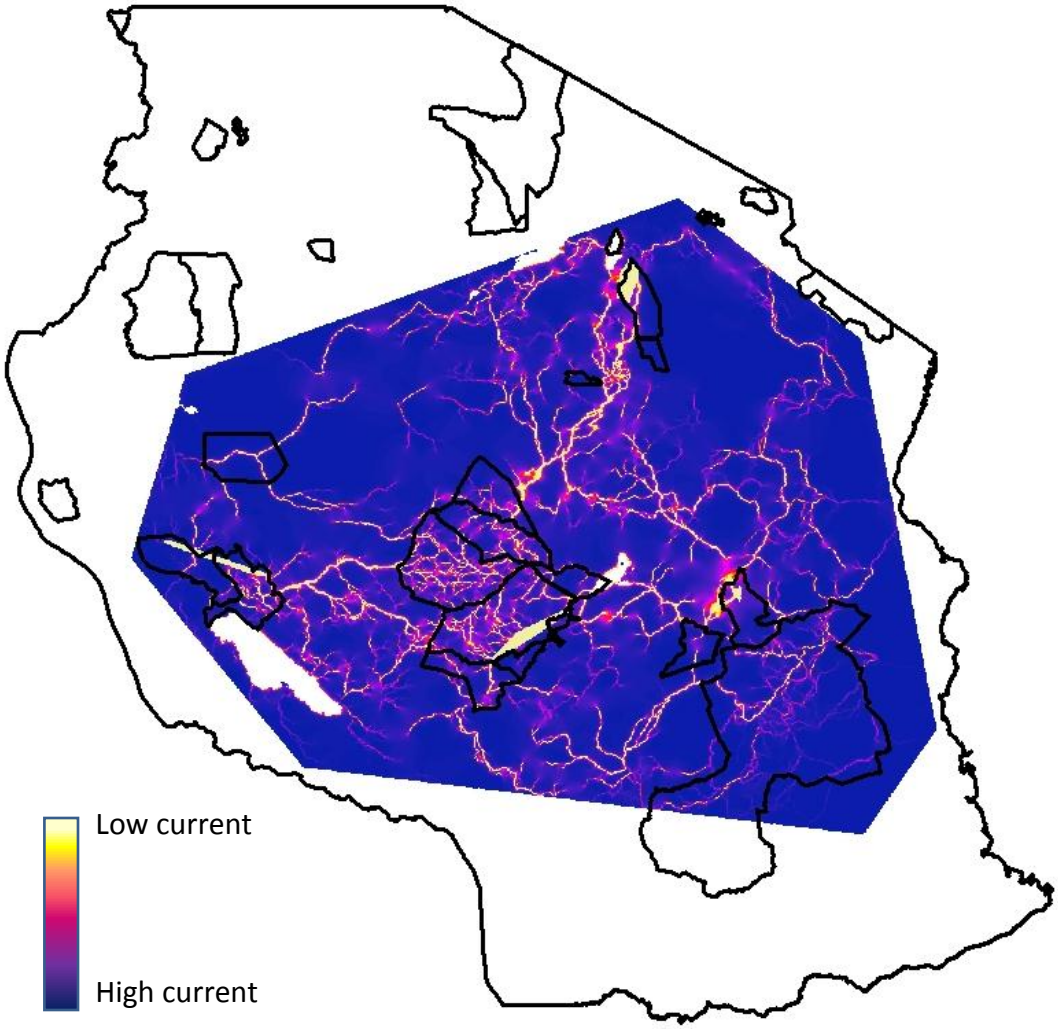


Figure 11. Continued.

d) ELAND

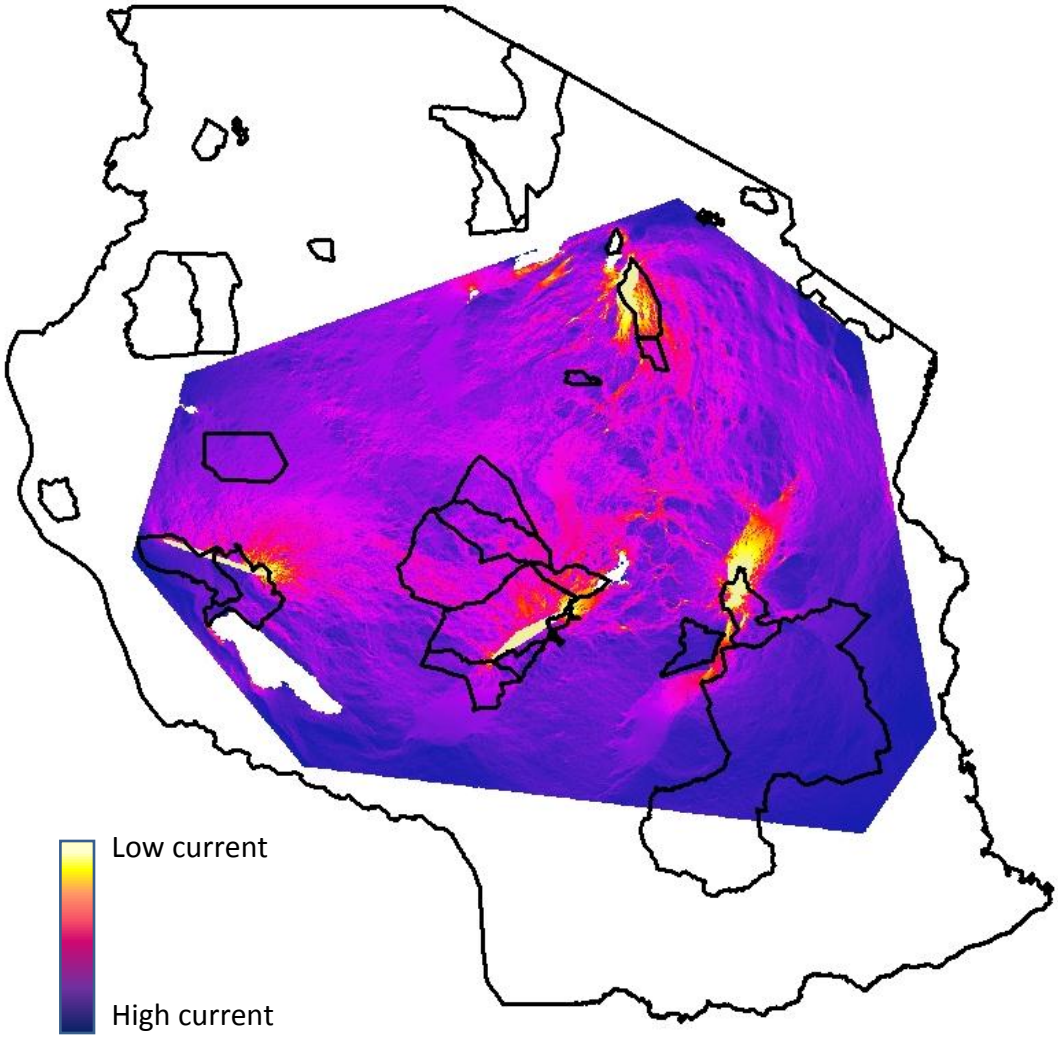


Figure 11. Continued.

APPENDICES

APPENDIX 1: Modified version of the AquaGenomic Stool and Soil Protocol (Multitarget Pharmaceuticals, Salt Lake City, Utah) used to extract DNA from ungulate faecal samples.

Sample processing (scraping) protocol:

In faecal samples the highest quality DNA is on the outer surface of the pellet; the inside portion of the pellet contains too little of the herbivore's DNA to amplify well. Material on the interior of the pellet may also contain PCR inhibitors.

1) Using the corner of a UV-sterilized square-edged razorblade, scrape 30 mg (0.03 g) of the outer layer of the pellet onto a piece of sterilized computer paper.

Pellets must be clean and dry; any sand, ash, dirt and large flakes from the pellets will decrease DNA quality and potentially clog pipette tips. Collect only fine scrapings from the exterior of the pellet.

2) Collect the dust in a 1.5 mL tube; store at –80 ºC until ready to extract the DNA. Discard any scraped pellets and store the remaining sample pellets at ambient temperature in a dry area.

Sample extraction protocol:

To decrease the possibility of contamination between samples use filter tips for all steps except B3 (optional to use nonfilter tips for the ethanol wash).

A) Lyse cells and extract DNA:

1) Bring samples to ambient temperature.

2) Add 450 μL of Aquagenomics solution and 0.2 mL of 1.0 mm zirconium beads to each 0.03 g sample of pellet dust.

3) Vortex at medium speed for 15 minutes to lyse the cells.

4) Add 25 μL of proteinase K (Qiagen), then incubate at 60 ºC for 90 minutes to lyse mitochondria and recover mitochondrial DNA. Vortex periodically throughout incubation to thoroughly mix the sample.

5) Incubate at 95 ºC for ten minutes to deactivate the proteinase K.

B) Remove cellular debris and pellet DNA:

1) Vortex for one minute, then centrifuge at 14,000 xg for four minutes to pellet cellular detritus.

2) Pipette the supernatant into a new 1.5 μL tube and estimate the total solution volume. Add one-half of that volume of 100% ethanol and one-half of the volume of AquaPrecipi solution. Vortex briefly to mix, then place the sample in the -20 ºC freezer for thirty minutes to precipitate the DNA from solution.

3) Centrifuge for four minutes at 14,000 xg to pellet the DNA. Discard solution, then rinse tubes 3-4 times using 350 μL of 70% ethanol; squirt ethanol from a large caliber pipette tip to cleanse the entire interior surface (including the lid) of the tube.

4) Place the open tube upside down on a Kimwipe to dry the DNA pellet.

C) Rehydrating DNA:

1) Add 100 μL of 1x TE buffer to the DNA pellet; gently dislodge and break pellet with pipette tip to mix. Leave all day (or overnight) at 4 ºC to solubilize the DNA.

2) Centrifuge for two minutes to pellet unsolubilized DNA; pipette supernatant (DNA) into a new 1.5mL tube.

3) Repeat the last two steps to obtain an additional 100 μL aliquot of DNA.

APPENDIX 2: Correlations between explanatory variables in optimal resistance models for giraffe, impala, and eland.

			Simple Mantel test		Partial Mantel test (controlling for geographic distance)	
Species	Variable 1	Variable 2	Correlation (*r*)	*P*	Correlation (*r*)	*P*
Giraffe	slope - nontransformed	geographic distance	0.900	<0.0001		
Impala	slope - squared	geographic distance	0.539	0.013		
Eland	distance to water - squared	geographic distance	0.664	0.175		
	slope - nontransformed	geographic distance	0.735	0.146		
	distance to water - squared	slope - nontransformed	0.944	0.007	0.899	0.029