

**Phylogeography, population genetics and
conservation of the okapi (*Okapia johnstoni*)**

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy

by

David William Graham Stanton, BSc. (Hons)

Illustrations by Margaret Stanton

April 2014

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed (candidate) Date

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed (candidate) Date

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references.

Signed (candidate) Date

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed (candidate) Date

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed (candidate) Date

STATEMENT 5

Work involving human participants and/or samples was performed with the approval of the appropriate Research Ethics Committee.

Signed (candidate) Date

To Re, Mum, Dad and Sarah
I could never have done any of this without you

Acknowledgements

First and foremost, I would like to thank Mike. I will always be grateful for the opportunity that you gave me – it has been an absolute privilege. In the end I just about escaped being nibbled to death by an okapi...

Thank you to my other supervisors, Jinliang (one of the smartest *and* nicest people I've ever met), John Ewen (thank you for the tireless patience with my endless drafts, re-drafts, re-re-drafts...), Noelle (remember when we were in Goma when they were firing rockets at the city?!), and John Hart (you still need to come to Cardiff to see the valley girls...).

I would like to thank everyone in the G10/C5.15 labs. Mafalda (estou farto desta merda), Niall (for always wanting to do something awesome to distract us from our theses), Klanky Isa (vol ye okopi koffi?), Pablo ("COFFEE'S READY!"), Rose (running away from your genotype?), Hannah Montana, Milena (never forget those early mornings and muddy banks...), Geoff (gotta love those otters), Mario (can you streamline my script?), Joanna (queres Café?), Silka (Milky Pilky), Peir, Xiangjiang, Azniza, Benoit, Rui, Tania...

Sixth floor cru et al! Tom (my man, you can still beat me at squash, even with a stroke), Sues (what can I say? Sheep eyes, pisscuit...), Steve (only person I know who would want to statistically analyse profanities in rap lyrics [*and* could actually do it], I'll keep the crack den damp for you), Matt (still got back!), Don (great socks), Caitlin, Beky & Laura (pardon the pun... still bitter... fucking Cambridge), Big El (what's your park run time?), Little El, Jo Lello (I'll get round to doing that DIY for you at some point), James Vaf, Marion...

All the rest of the people from the OnE department – thank you for making my time here so enjoyable, I'm going to miss you all!

Thank you to Chrissy (incredible dinners), Alan, Rhys, Ga, Mrytle, luce.

The work presented here was funded by a NERC CASE studentship with the ZSL. Additional funding was provided by Gilman International Conservation.

Table of Contents

Declaration	iii
Acknowledgements	v
Summary	viii
List of publications	ix
CHAPTER ONE – General Introduction	2
1.1 Background	2
1.1.1 First taxonomic description	3
1.1.2 Molecular studies on okapi	4
1.1.3 Okapi distribution	4
1.1.4 Okapi conservation	4
1.1.5 Biogeography of the Congo Basin	12
1.1.6 Congolese fauna	13
1.2 Project aims	14
1.2.1 Global objectives	14
1.2.2 Genetics and conservation	14
1.2.3 Aims of this phd research project	16
CHAPTER TWO - Microsatellite loci for the okapi (<i>Okapia johnstoni</i>)	27
2.1 Abstract	27
2.2 Article	27
2.3 Acknowledgements	31
CHAPTER THREE - Non-invasive genetic identification confirms okapi (<i>Okapia johnstoni</i>) presence southwest of the Congo River	33
3.1 Abstract	33
3.2 Article	34
3.3 Acknowledgements	43
CHAPTER FOUR - Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered okapi (<i>Okapia johnstoni</i>)	45
4.1 Abstract	45
4.2 Introduction	46
4.3 Methods	48
4.3.1 Study area and sampling	48
4.3.2 Molecular methodology	48
4.3.3 Nuclear DNA primers	49
4.3.4 Sequence analysis	50
4.3.5 Partitioning of genetic diversity	51
4.3.6 Population and sequence divergences	52
4.4 Results	53
4.4.1 Primer design	53
4.4.2 Sequence analysis	55
4.4.3 Partitioning of genetic diversity	57
4.4.4 Population and sequence divergences	60
4.5 Discussion	64
4.5.1 Okapi genetic diversity and evolutionary history	64

4.5.2 Evolutionary biogeography of the Congo Basin	65
4.5.3 Partitioning of present-day genetic diversity	68
4.6 Acknowledgements	69
4.7 Supplementary	70
CHAPTER FIVE - Enhancing knowledge of the ecology of a highly elusive species, the okapi, using non-invasive genetic techniques	78
5.1 Abstract	78
5.2 Introduction	79
5.3 Methods	80
5.3.1 Study species and site	80
5.3.2 DNA extraction and amplification	81
5.3.3 Data validation	84
5.3.4 Spatial autocorrelation	84
5.3.5 Patterns of relatedness	85
5.3.6 Multiple dung piles	85
5.3.7 Mating system	86
5.3.8 Duplicate genotypes	86
5.3.9 Spatial genetic structuring	86
5.3.10 Sex-biased dispersal	87
5.4 Results	88
5.4.1 Spatial autocorrelation	88
5.4.2 Patterns of relatedness	88
5.4.3 Multiple dung piles	89
5.4.4 Mating system	89
5.4.5 Duplicate genotypes	89
5.4.6 Spatial genetic structure	89
5.4.7 Sex-biased dispersal	98
5.5 Discussion	100
5.5.1 Okapi sociality	100
5.5.2 Okapi mating systems	102
5.5.3 Sex-biased dispersal	104
5.6 Acknowledgements	105
CHAPTER SIX - An integrated genetic management assessment for the okapi (<i>Okapia johnstoni</i>)	107
6.1 Abstract	107
6.2 Introduction	108
6.3 Methods	111
6.3.1 Samples	111
6.3.2 Lab methodology	112
6.3.3 Data validation	112
6.3.4 Genetic diversity	113
6.3.5 Genetic structure	114
6.4 Results	115
6.4.1 Samples	115
6.4.2 Genotype validation	118
6.4.3 Genetic diversity	118
6.4.4 Genetic structure	120
6.5 Discussion	125

6.6 Acknowledgements	129
6.7 Supplementary	131
CHAPTER 7 – General discussion	139
7.1 Aims	139
7.2 Completion of aims	139
7.3 Conservation recommendations	141
7.3.1 Wild population	141
7.3.2 Captive population	142
7.4 Role of genetics in the future of okapi conservation	143
7.5 Conservation reality on the ground in the DRC	147
APPENDIX ONE - MTfate: A program for modelling the fate of mitochondrial DNA lineages in ex-situ populations (concept and preliminary results)	151
A1.1 Abstract	151
A1.2 Introduction	151
A1.3 Methods	152
A1.3.1 Model	152
A1.3.2 Testing	153
A1.4 Results	153
A1.5 Discussion	155
A1.6 Acknowledgements	158
A1.8 Supplementary	158
BIBLIOGRAPHY	160

Summary

The okapi (*Okapia johnstoni*) is an endangered, evolutionarily distinct giraffid, endemic to the Democratic Republic of Congo (DRC). The okapi is a flagship species for the DRC, a country that contains some of the greatest biodiversity in the world. The okapi is currently under major threat from habitat fragmentation, human encroachment and poaching, yet to date, very little is known about the species in the wild, and no genetic study in the wild or captivity has ever been carried out. This thesis aims to use genetics to aid conservation efforts of okapi, a species that, due to its elusive nature, is highly challenging to study using alternative methods. We describe 13 polymorphic microsatellite loci for the okapi that will be useful for future studies of population genetic diversity and genetic structure. We further knowledge of okapi distribution by confirming that they occur on the southwest side of the Congo River, in localized distributions west of the Lomami River, and develop a simple molecular diagnostic tool for identifying okapi dung. We provide new ecological information about this species, showing that they appear to be mostly solitary, demonstrate male-biased dispersal, and are genetically polygamous or promiscuous, and are also likely to be socially polygamous or promiscuous. We show that there are similar levels of nuclear genetic variation in the wild, founder and captive okapi populations, however, mitochondrial genetic diversity within captive okapi is considerably reduced compared to the wild. In addition, both nuclear and mitochondrial alleles present in captivity poorly represent the allelic diversity present in the wild. This PhD thesis constitutes the first genetic study of wild and captive okapi populations, and provides important conservation information for this emblematic species.

List of publications

Publications associated with this thesis

- Stanton, D. W. G.**, Penfold, L. M., Zhan, X., & Bruford, M. W. (2010). Microsatellite loci for the okapi (*Okapia johnstoni*). *Conservation Genetics Resources*, 2(S1), 337–339. doi:10.1007/s12686-010-9235-0
- Stanton, D. W. G.**, Ewen, J. G., Galbusera, P., Wang, J., Bruford, M. W., et al. Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered Okapi (*Okapia johnstoni*). (Submitted: *PLOS ONE*).
- Stanton, D. W. G.**, Ewen, J. G., Galbusera, P., Wang, J., Bruford, M. W., et al. An integrated genetic management assessment for the okapi (*Okapia johnstoni*) (Submitted: *Conservation Genetics*).
- Stanton, D. W. G.**, Ewen, J. G., Hart, J., Wang, J., Bruford, M. W., et al. Non-invasive genetic identification confirms okapi presence southwest of the Congo River. (Submitted: *Oryx*).
- Stanton, D. W. G.**, Ewen, J. G., Hart, J., Wang, J., Bruford, M. W., et al. Furthering knowledge of the ecology of a highly elusive species, the okapi (*Okapia johnstoni*), using non-invasive genetic techniques. (Submitted: *Animal Conservation*).

Publications not directly associated with this thesis

- Stanton, D. W. G.**, Hobbs, G. I., Chadwick, E. A., Slater, F. M., & Bruford, M. W. (2008). Mitochondrial genetic diversity and structure of the European otter (*Lutra lutra*) in Britain. *Conservation Genetics*, 10(3), 733–737. doi:10.1007/s10592-008-9633-y
- Stanton, D. W. G.**, Hobbs, G. I., Mccafferty, D. J., Chadwick, E. A., Philbey, A. W., Sachheri, I. J., Slater, F. M., Bruford, M. W. (In Press) Contrasting genetic structure of the Eurasian otter (*Lutra lutra*) across a latitudinal divide. (*Journal of Mammalogy*).
- Morrissey, C. A., **Stanton, D. W. G.**, Tyler, C. R., Pereira, M. G., Newton, J., Durance, I., Ormerod, S. J (In Press). Developmental impairment in Eurasian dipper nestlings exposed to urban stream pollutants. *Environmental Toxicology and Chemistry*.
- Morrissey, C. A., **Stanton, D. W. G.**, Pereira, M. G., Newton, J., Durance, I., Ormerod, S. J (2013) *Environmental Science and Technology*. 47(15) 8931-8939. doi: 10.1021/es402124z.
- McCann, N, Orozco ter Wengel, P, **Stanton, D. W. G.** Modelling genetics within ecosystems. *Nature Correspondence*. 495(47). doi:10.1038/495047d
- Crowther, T, **Stanton, D. W. G.**, Thomas, S, A'Bear, A, Hiscox, J, Jones, T, Voříškovác, J, Baldrian, P, Boddy, L (2012). *Ecology*. Metagenomic Sequencing Uncovers Mechanism Underlying Top-Down Control in Woodland Soil.
- Salgado-Lynn, M., **Stanton, D. W. G.**, Sakong, R., Cable, J., Goossens, B., & Bruford, M. W. (2010). Microsatellite markers for the proboscis monkey (*Nasalis larvatus*). *Conservation Genetics Resources*, 2(0), 159–163. doi:10.1007/s12686-010-9295-1
- Ma, Z., Song, H., Zhong, J., **Stanton, D.**, Wei, Y., & Sun, Y. (2012). Short Communication: molecular characterisation of the wild yak (*Bos grunniens mutus*) melanocortin receptor-4 (MC4R) gene and a comparative analysis with that of other Bovinae species. *Journal of Applied Animal Research*, 40(2), 81–85. doi:10.1080/09712119.2012.662775

1

CHAPTER ONE – General introduction



"This is rather as if you imagine a puddle waking up one morning and thinking, 'This is an interesting world I find myself in — an interesting hole I find myself in — fits me rather neatly, doesn't it? In fact it fits me staggeringly well, must have been made to have me in it!' This is such a powerful idea that as the sun rises in the sky and the air heats up and as, gradually, the puddle gets smaller and smaller, frantically hanging on to the notion that everything's going to be alright, because this world was meant to have him in it, was built to have him in it; so the moment he disappears catches him rather by surprise. I think this may be something we need to be on the watch out for."

Douglas Adams, The Salmon of Doubt

CHAPTER ONE – General Introduction

1.1 Background

The okapi (*Okapia johnstoni*) is an even-toed ungulate endemic to the Democratic Republic of Congo (DRC), classified as Endangered by the IUCN in 2013 (IUCN, 2013). The species is extremely elusive, and was not photographed in the wild until 2008 (Nixon & Lusenge, 2008). It belongs to the family Giraffidae and its closest extant relative is the giraffe (*Giraffa camelopardalis*). The okapi is a flagship species for the DRC, a relatively poorly known country containing some of the greatest biodiversity in the world, and appears on the emblem of the national conservation agency, the Institut Congolais pour la Conservation de la Nature (ICCN).

The majority of what is known about okapi biology is from the captive population (Bodmer & Rabb, 1992). Although breeding okapi in captivity is relatively easy, rearing calves has been found to be more difficult. Until the 1950's approximately 50% of okapi calves died within the first month (Gijzen & Smet, 1974). Okapi usually give birth to a single offspring, although there has been a single recorded case of twinning (Pearson et al., 1978). Gestation periods vary between 414-493 days (Gijzen & Smet, 1974). Okapi weigh 14-30 kg at birth and follow the mother for 1-2 days before settling in one place. For the first two months, okapi spend approximately 80% of the time at the nest, probably to ensure rapid growth. Females can defend infants by striking with their forelegs (Bodmer and Rabb, 1985 in (Bodmer & Rabb, 1992)).

There have only been two ecological studies carried out on wild okapi. Hart and Hart (1989) described how okapi in the Ituri forest, DRC are highly selective feeders, with browsing focussed on tree fall gaps. Data based on the movements of eight radio-collared individuals, home range for females was estimated at between 1.9 km² to 5.1 km² (n = 5, mean 3.2 ± 1.36). Home ranges of males were estimated at 1.6, 9.2 and 10.5 km². Two of the eight collared okapi were predated by leopards during the study period, between March 1986 – January 1988. Bodmer and Gubista (1988) used observations of okapi prints in a 1.69 km² section of the Ituri forest, DRC. This study followed four individuals, and inferred that these prints were created by two adults,

one juvenile and one calf. These individuals were shown to exhibit a loose group arrangement.

1.1.1 First taxonomic description

The first account of an okapi was by Henry Morton Stanley during his exploration

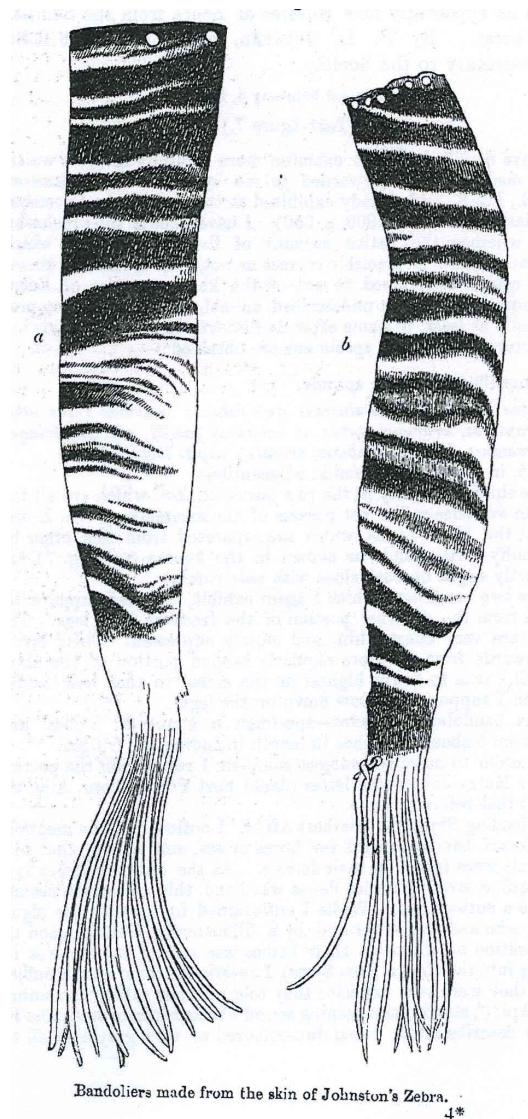


Figure 1.1 Okapi skins sent by Johnston to ZSL for identification (Sclater, 1901). Caption reads: “Bandoliers made from the skin of Johnston's Zebra”

through the Ituri forest in 1890. “The Wambutti knew a donkey and called it ‘Atti’. They say that they sometimes catch them in pits. What they can find to eat is a wonder. They eat leaves.” (Stanley, 1890). Sir Harry Johnston, British commissioner of Uganda, followed these reports up by visiting the Mbuti pygmies in the Ituri forest. The pygmies described an animal in their forests resembling a horse or ass that they called “ O’api (‘ equates to a gasping sound like an aspirate or Arabic K).” (Johnston, 1900). Johnston managed to procure two sections of skin, which he sent to The Zoological Society of London (ZSL; Figure 1.1; caption reads “Bandoliers made from the skin of Johnston's Zebra”), who soon identified that the skins did not belong to any known species of zebra or horse. The species was given the provisional name of “EQUUS (?) JOHNSTONI, sp. nov.” (Johnston, 1900). Johnston received a full skin and two skulls from the Swedish officer, Karl Eriksson, and soon realised that the okapi was more closely

related to the giraffe. However, due the okapi’s distinctiveness from the Giraffa species, the species name was revised to *Okapia johnstoni* (Lankester, 1901).

There is a distinct lack of paleontological information about *Okapia spp.* The only fossil predating the Pleistocene is *Okapia stillei* (Dietrich (1942) in Van der Made and Morales (2012)), which has since been reclassified as *Giraffa* (Harris et al., 2010). This is likely a consequence of the okapi's affinity for closed-canopy forest (Harris et al., 2010), where high humidity and decomposition rates leads to a scarcity of fossils.

1.1.2 Molecular studies on okapi

To our knowledge, no genetic study of wild okapi has ever been carried out, however DNA samples from captive individuals have been used to approximate the divergence time of okapi and giraffe (Fernandez & Vrba 2005; Hassanin et al. 2012). Fernandez and Vrba (2005) calculated a “best estimate” date of the divergence of okapi and giraffe as 17.8 mya. This was based on a local molecular clock phylogenetic tree, using a meta-analysis of 140 studies, yielding 660 point estimates for 127 nodes throughout the tree. This included 14 nodes with only fossil estimates, 40 with only molecular estimates, and 73 with estimates derived from both. Hassanin et al. (2012) estimated a divergence between okapi and giraffe of between 15.2 and 17.2 mya. The phylogeny used a relaxed molecular clock with six time constraints, based on fossil calibrations of six nodes.

1.1.3 Okapi distribution

The exact distribution of okapi is unknown. Most range descriptions of okapi claim they are only present on the North-East side of the Congo River (IUCN (2008), Kingdon (1997) and Stuart and Stuart (1997); Figures 1.2A, 1.2B and 1.2C respectively). However, anecdotal reports (J. Hart and S. Nixon Pers. Comm.), and putative okapi faecal samples (Quinn et al., In Prep), suggest okapi may also be present on the South-West side. This allows a “best guess” of the current range of okapi to be estimated (Quinn et al., In Prep; Figure 1.3).

1.1.4 Okapi conservation

The conservation of species can be extremely costly. It is therefore important that conservationists justify why species should be conserved, and why a particular

species requires conservation. We are currently in the midst of the sixth global extinction (Leakey & Lewin, 2000). Unlike previous mass extinction events however, this event is primarily due to human activities (IUCN, 2013). The okapi was listed as Near Threatened (NT) in the 2008 IUCN Red List (IUCN, 2008), based on the fact that there is still a considerable population in the Okapi Faunal Reserve. However, recent post-war surveys report a 44% decline between 1996 and 2006 (Hart, et al. 2008). If this decline is typical throughout the range, the global okapi population could be much lower than the 10,000 – 35,000 individuals estimated by IUCN at this time, and likely to be declining rapidly. The NT classification was recently reassessed during an okapi conservation workshop (Okapi Conservation Workshop, 2013), based on three estimates of change in okapi population size, and regional, anecdotal reports of threats to okapi.

The first estimate of change in population size (Hart et al., 2008) used a transect-recce methodology implemented using DISTANCE (Thomas et al., 2010). It constituted two surveys, finished in 1995 and 2007, and estimated okapi numbers of 4428 (min 2947, max 6655) and 2507 (min 1622, max 3871) respectively. The second estimate (Vosper et al., 2012) again used DISTANCE with recce-transects, and two surveys finished in 2007 and 2011. This survey estimated okapi numbers of 3363 (min 1752, max 6457) and 5255 (min 2775, max 9884) respectively, using the same dung degradation rate as Hart et al. (2008; 75 days). The Vosper et al. (2012) survey concluded that the increase in okapi abundance over the period of the two surveys was significant ($W = 11545$, $p = 0.0456$). These two surveys are summarised in Figure 1.4. The third estimate used law enforcement monitoring patrol data (Okapi Faunal Reserve [RFO] and ICCN unpublished data; Figure 1.5), which recorded encounter rate between 2008 and 2012. Patrols covered distances of between 10,125 and 25,467 km annually. These showed a 47.3% decrease in okapi encounter rate and a 64.6% decrease in direct observations over this period. Confidence intervals cannot be calculated using this methodology however. One output of the 2013 okapi workshop was a recommendation that the conservation status of this species be reclassified to Endangered (E) based on a decline estimated to have exceeded

50% over three generations (beginning in 1995). This recommendation was made based on only the first and the third estimate. The second estimated an okapi population increase in the RFO of 1892 (or 56.3%) in a four-year period. An increase of this magnitude was considered infeasible in this time period, and the overlap in the confidence intervals of the two surveys (overlap of 3682) was considered too high. Due to these reasons, and also that the estimates for this survey were in contradiction to all local reports that hunting and human disturbance in the area was on the increase, this second survey was not taken into account in the status review.

Okapi are also subject to multiple threats. They seem able to coexist with small-scale, low-level human occupation but not in areas of active settlement or disturbance (Hart et al., 2008). Approximately one-third of the okapi's known range is thought to be at risk of major human incursions during the first quarter of this century. The primary threat to the okapi is therefore habitat loss due to logging and human settlement, and the resultant habitat fragmentation (IUCN, 2013). Okapi also appear to be hunted fairly heavily as bush-meat and skins are frequently seen in eastern DRC (S Nixon Pers. Comm.; Hart et al., 2008; Nixon & Lusenge, 2008). The mining industry in the DRC is also likely to be an important anthropogenic threat. The DRC contains a vast amount of natural resources including half of Africa's rainforests (86 million hectares) as well as an

Figure 1.2. Okapi range according to IUCN (2008; A), Kingdon (1997; B), and Stuart and Stuart (1997; C) .

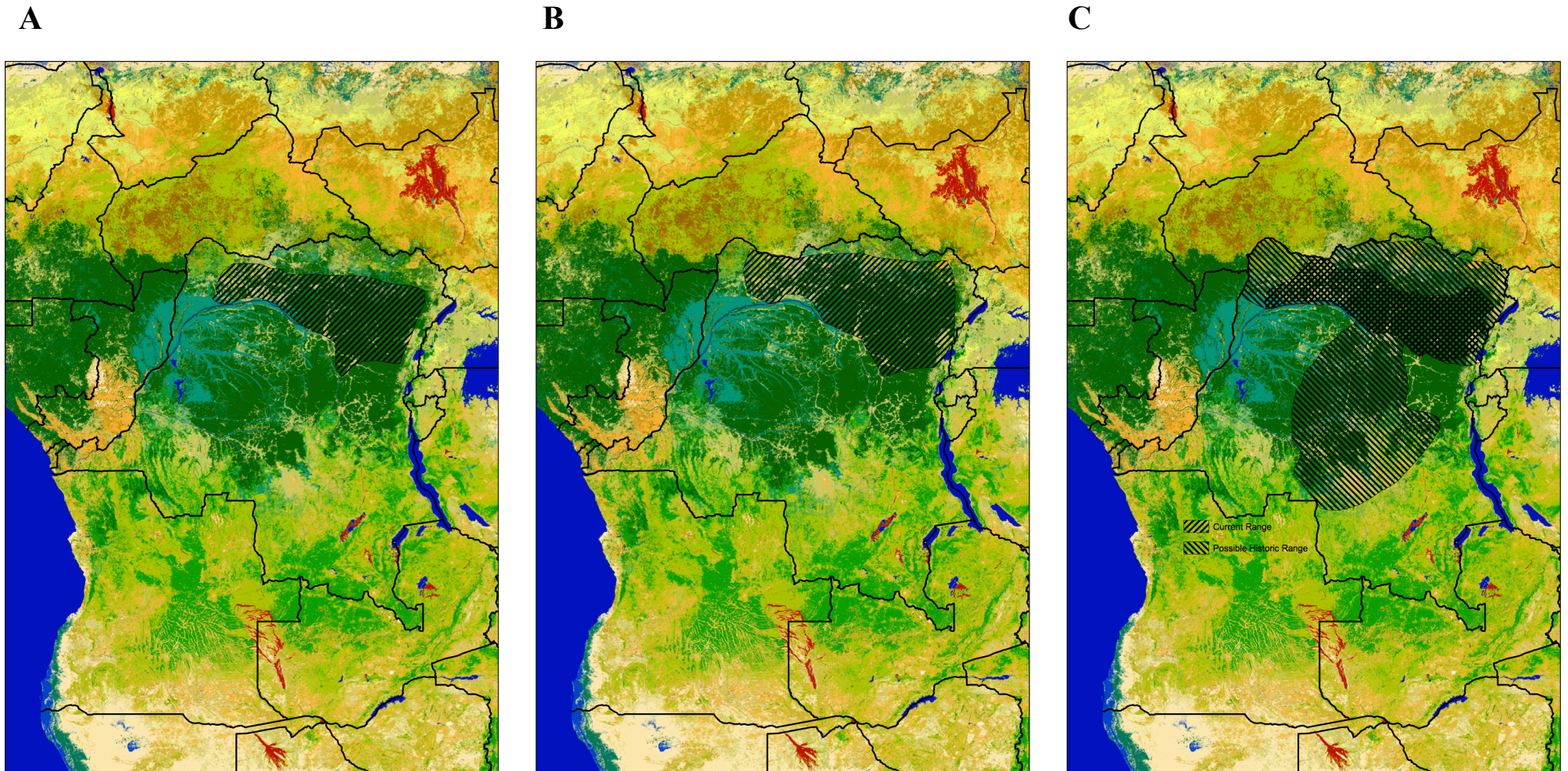
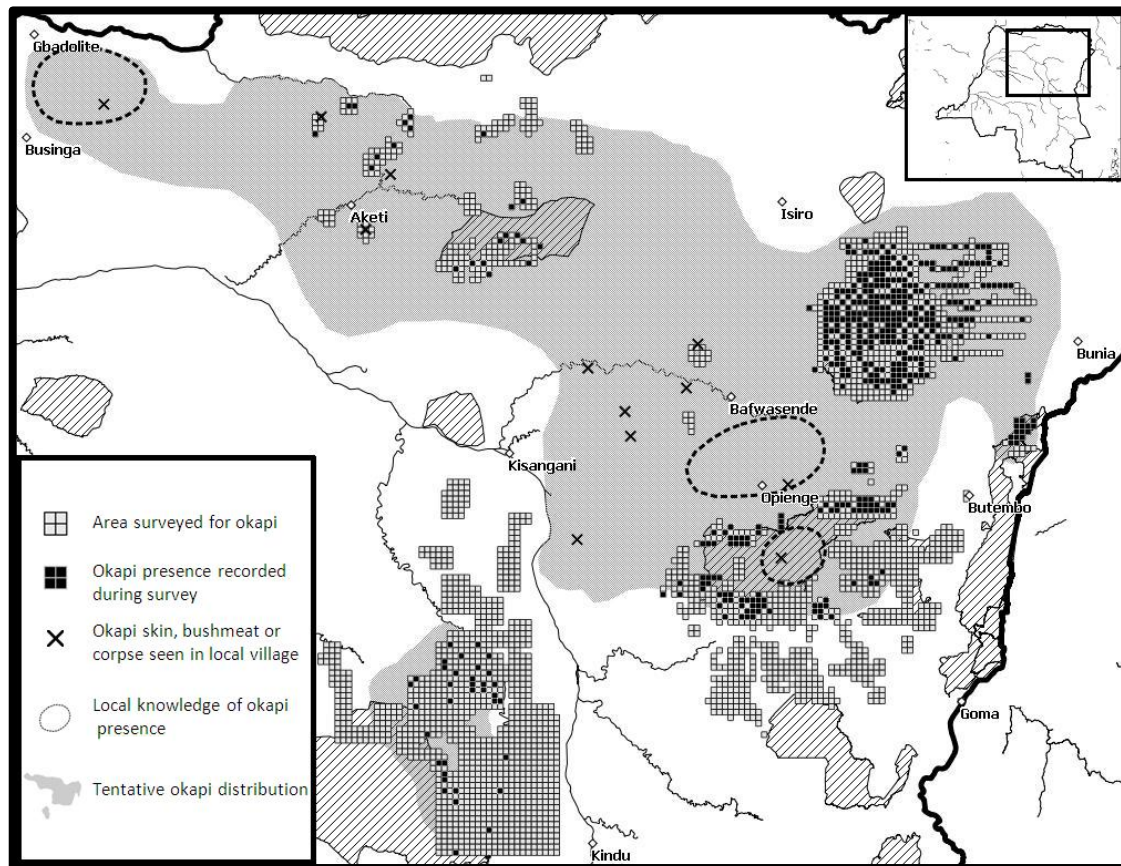


Figure 1.3. Map from okapi conservation workshop (Okapi Conservation Workshop, 2013; Quinn et al., (In Prep)), showing skin, and putative okapi dung samples.



enormous mineral wealth, especially coltan, but also including copper, cobalt, diamonds, gold, zinc and oil (IES, 2008; Malele Mbala, 2010; United Nations, 2011). The presence of these minerals in a particular area can lead to an influx of people, and result in a significant increase in deforestation and bush-meat hunting (Hayes & Burge, 2003). Also, the change in land use that accompanies mining practice is putting an increased pressure on the forest biodiversity in the DRC (Potapov et al., 2012). Mineral wealth is also thought to help finance rebel movements throughout the Eastern DRC, which would be expected to increase the levels of subsistence hunting throughout the forests, as it is in these forests that the various rebel groups predominantly live (S. Nixon Pers. Comm.).

Figure 1.4 Combination of the okapi survey results of the Hart et al. (2008) and Vosper et al. (2012) studies

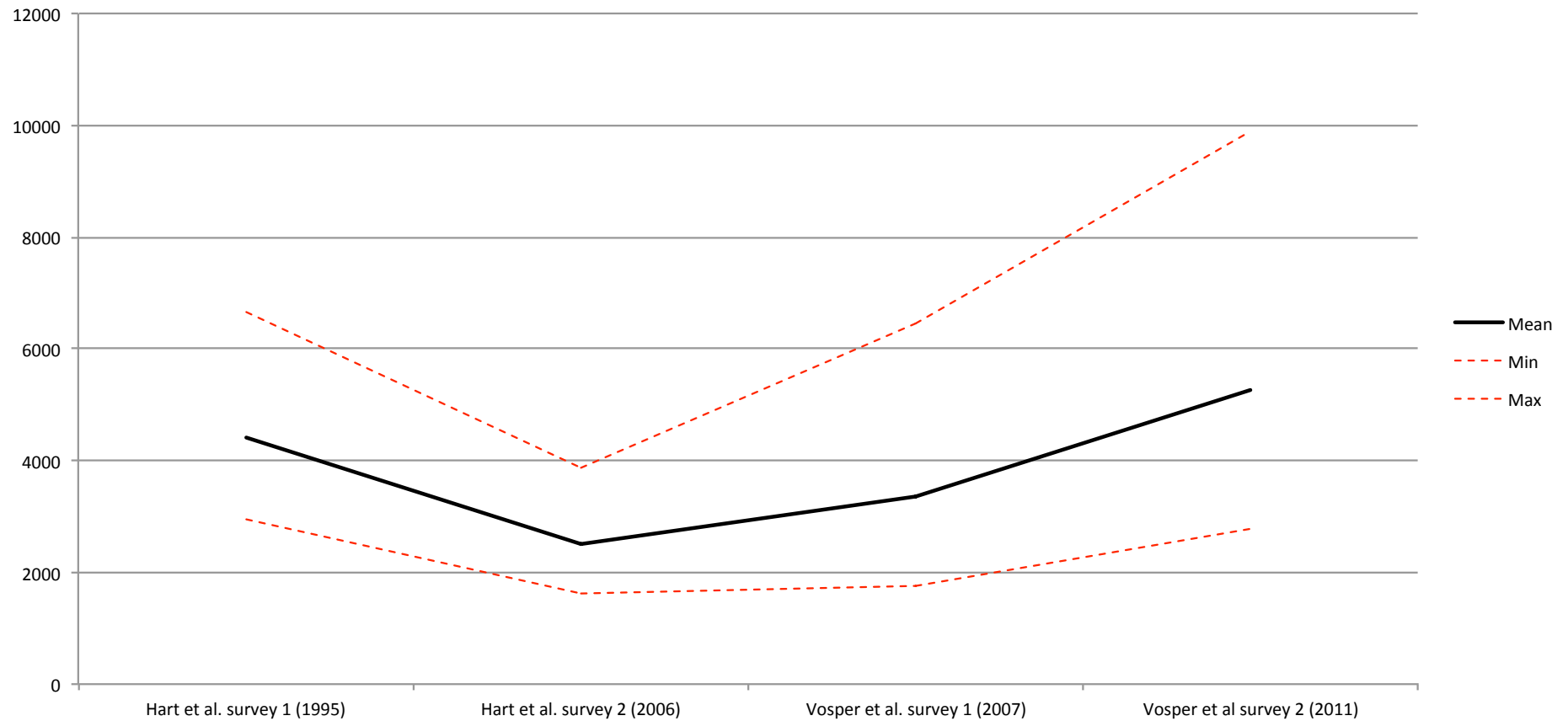
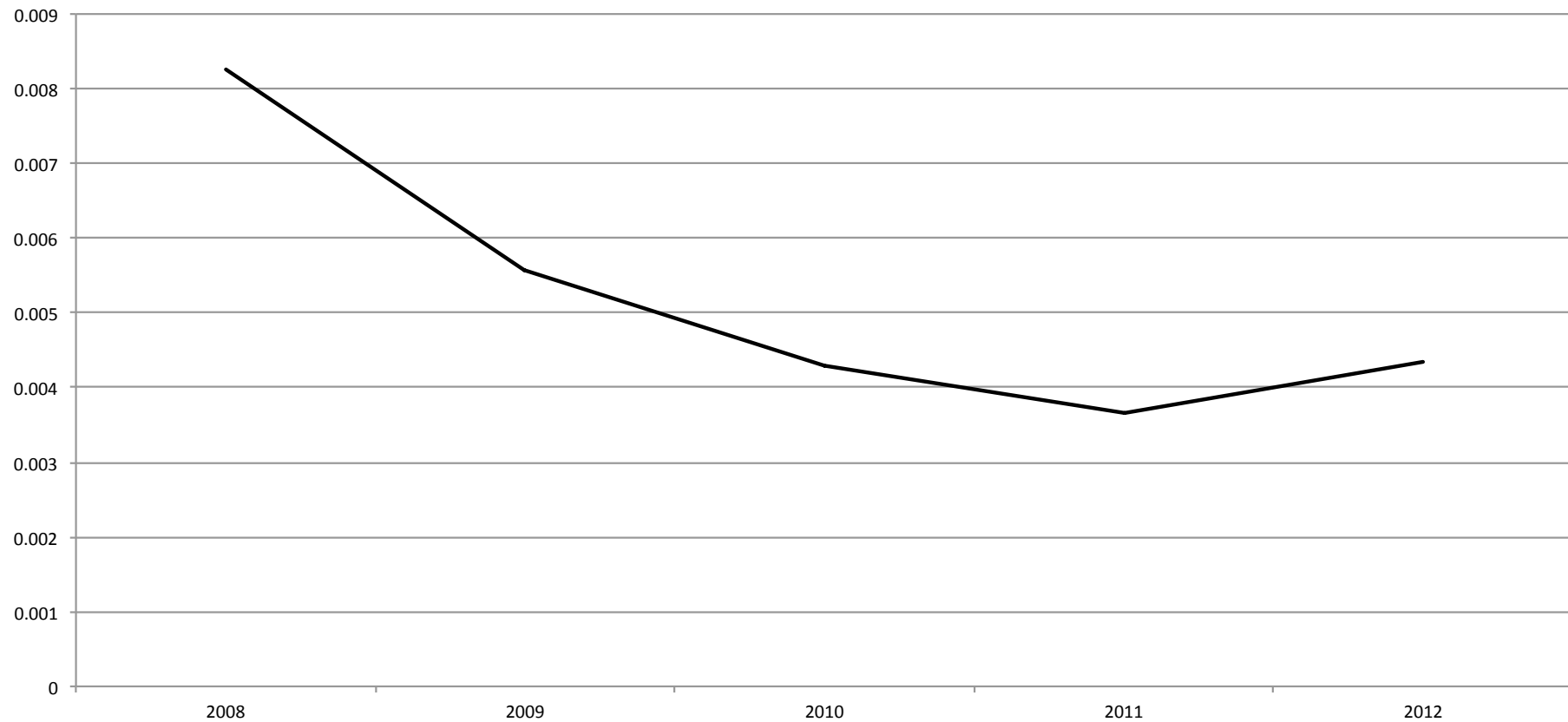


Figure 1.5 Encounter rates of okapi sign based on ICCN (Institut Congolais pour le conservation de la Nature) surveys between 2008-2012



Okapi are therefore clearly in need of conservation efforts, however, in light of limited funding for conservation, the question remains if conservation efforts should be directed toward okapi rather than towards other taxa and/or questions. In recent years there has been a push towards more all-encompassing conservation projects, taking an ecosystem-services, rather than an individual taxon-based approach (Armsworth et al., 2007). In a sense it is a tragedy that we have to justify conserving such a unique and incredible organism. The argument for less conservation emphasis on large, charismatic mammals is understandable however, considering eutherian mammals almost certainly constitute less than 1% of the world's extant species (Primack, 2000). Nonetheless, okapi are thought to be separated from giraffe, the only other member of the Giraffidae family, by ~16 million years of evolution (Hassanin et al., 2012) and comprise a highly distinct evolutionary lineage. Protecting evolutionarily distinct species is also recognised as an important aspect in conservation, with recent programmes such as the EDGE (Evolutionarily Distinct & Globally Endangered) of existence project, a ZSL-led project that focuses on globally endangered, evolutionarily distinct species (EDGE, 2014). Clearly there needs to be a balance, with some conservation projects acting on single taxa and others adopting a more generalist approach. In addition, the okapis striking appearance, and evolutionary distinctness, make them a potentially excellent flagship species for the DRC. They are already on the emblem of the country's conservation agency, the ICCN. The Epulu Conservation and Research Center (RFO, Orientale Province, DRC), housed 14 captive okapi through the entire period of "Africa's World War" (1998 – 2003), which mostly occurred on Congolese soil. The okapi survived this entire period and were a symbol of conservation in the area during this time. Tragically, MaiMai Simba rebels attacked the Epulu Conservation and Research Center in 2012, killing six people and all 14 okapi. These events highlight the challenging conservation situation in the DRC. Disturbingly, the events at Epulu were attributed to a response to a crack-down on poaching in the RFO, implying there is still a long way to go in the management of this faunal reserve.

1.1.5 Biogeography of the Congo Basin

There has been considerable debate regarding the biogeographical history of Central Africa, in particular the date and extent of forest fragmentation. The biogeographic history of Central Africa is usually interpreted based on hypothesised climatic changes throughout the Pleistocene, specifically increases in African climate variability (Potts, 2013) and aridity (DeMenocal, 2004) at approximately 2.8 – 2.5, 1.9 – 1.7 and 1.1 – 0.9 mya. There are competing theories of the extent to which these increases in climatic variability and aridity affected forest cover in Central Africa. The island refugia model (Hamilton, 1975; Livingstone, 1975) infers isolated patches of forest, broken up by grassland. The competing hypothesis (Cowling et al., 2008), is based on simulated paleovegetation of Central Africa and last glacial maximum simulations, indicating that although tropical broadleaf forest may not have been severely displaced by expanding grassland in central Africa, the structure of the forests may have been very different from today (with forests characterized by lower leaf area indices, lower tree heights and lower carbon content).

In the case of the island refugia model, we would expect the flora and fauna of Central Africa to show a structured pattern of mtDNA variation, caused by prolonged isolation in separate refugia. In the case of forest cover being reduced only minimally, we would expect a relatively unstructured phylogeographic pattern.

A structured pattern of mtDNA variation is evident in certain African species (e.g. gorillas, Anthony et al., 2007, and Jensen-Seaman and Kidd, 2001), whereas a relatively unstructured phylogeography can be observed in others (e.g. chimpanzees, Morin et al., 1994; mandrills, Telfer et al., 2003). This difference originates from different species having different ecological niches and limits, causing a different response to the same biogeographical factors. Therefore, more phylogenetic studies are required to refine the theories of central African biogeography. We believe okapi to be a useful model to test these theories as the species range covers a significant proportion of the central African rainforest, and the species has a unique ecology compared to any previously investigated.

1.1.6 Congolese fauna

The DRC contains some of the greatest biodiversity on the planet (Primack, 2000), and the okapi shares its range with a considerable number of other medium to large-bodied mammals. Table 1.1 (adapted from Hart et al. (2008)) lists the medium and large-bodied mammals known to inhabit the RFO. The faunal assemblages outside of the RFO are not as well catalogued. A recent survey of the Rubi-Tele hunting reserve (Orientale Province) identified the presence of nine primate species (one *Cercocebus* spp, two *Colobus* spp, five *Cercopithecus* spp and *Pan troglodytes*), elephants (*Loxodonta africana cyclotis*) and okapi (*Okapia johnstoni*). All species appeared to be at low density and there was considerable evidence of hunting. Faunal assemblages southwest of the Congo River is even more poorly catalogued, however a key difference in the fauna is the presence of bonobos (*Pan paniscus*) and absence of chimpanzees (*P. troglodytes*).

Table 1.1. Faunal assemblage in the RFO based on the wildlife surveys from Hart et al. (2008).

Taxon	Species
Okapi	<i>Okapia johnstoni</i>
Elephant	<i>Loxodonta africana cyclotis</i>
Small ungulates	<i>Cephalophus monticola</i> <i>Neotragus batesi</i> <i>Hyemoschus aquaticus</i>
Red duikers	<i>Cephalophus nigrifrons</i> <i>C. leucogaster</i> <i>C. weynsi</i> <i>C. dorsalis</i>
Yellow duikers	<i>Cephalophus sylvicultor</i>
Suids	<i>Potamochoerus porcus</i> , <i>Hylochoerus meinertzhageni</i>
Buffalo	<i>Syncerus cafer</i>
Primates	<i>Cercopithecines</i> (6 species) <i>Colobus</i> (3 species) <i>Cercocebus</i> (2 species)

1 General Introduction

	<i>Baboon (1 species)</i> <i>Pan troglodytes</i>
Leopard	<i>Panthera pardus</i>

1.2 Project aims

1.2.1 Global objectives

A necessary prerequisite to conservation efforts is a basic understanding of the species in question. Broadly speaking, this translates to answering the following questions: Evolutionary status (i.e. intra- and interspecific taxonomy, information about their ecology). Distribution (i.e. what is their current geographic range). Demographics (i.e. age/sex ratio, census sizes and effective population sizes). Without this information, conservation efforts could be unnecessary, and/or totally ineffective. Captive breeding programs are also often an important part of conservation efforts, providing a source of individuals for study *ex-situ*, and a potential source of reintroductions (Rahbek, 1993; Woodworth et al., 2002). Genetics can be used to investigate all of these questions (Frankham et al., 2002).

1.2.2 Genetics and conservation

The International Union for the Conservation of Nature (IUCN) recognises the importance of genetics in conservation, as genetic diversity is required for populations to respond to environmental change (Frankham et al., 2002; IUCN, 2013).

Following the first description of genetics and the mode of inheritance by Mendel, the first understanding of how evolution is defined by changes in allele frequencies, through the processes of mutation, selection and genetic drift, influence evolution was brought about by the work of Fisher (1930) and Wright (1931). The first person to recognise the importance of genetics in conservation was Sir Otto Frankel (Frankel & Soulé, 1981). Richard Frankham, among others, built on this foundation to give us an understanding of key aspects of conservation genetics, such as how inbreeding effects extinction (Frankham, 1995; Frankham & Ralls, 1998), the relationship between genetic variation and population size in wild populations (Frankham, 2002), selection in captive populations (Frankham et al., 1986) and the effect of genetic bottlenecks on

populations (Frankham et al., 1999; Li & Durbin, 2011). Today, these parameters can be inferred for a whole population, based on a single genome (Li & Durbin, 2011).

As described above, much of the information needed for the effective conservation of okapi is lacking. Genetics provides a potential means of elucidating this information, however, the okapi is a particularly challenging animal to study. They occur at low density across their range, and appear to only be present in dense forest, away from human presence (Bodmer & Rabb, 1992; J Hart Pers. Comm.). It is situations like this that non-invasive genetics may also become a useful option (Taberlet et al., 1996; Taberlet & Luikart, 1999; Taberlet et al., 1999). Non-invasive genetic methods are increasingly being used in ecology and conservation (e.g. Taberlet et al., 1997; Goossens et al., 2005; Zhan et al., 2007), and are likely to be highly appropriate here.

Since the advent of next-generation sequencing, genetic diversity can be assessed across the whole genome (e.g. Mardis, 2008; Li et al., 2010). Due to time and monetary restraints however, whole-genome sequencing is still not accessible to most population studies on non-model organisms. Genome-wide variation is often approximated using one, or a number of different genetic markers, for example microsatellites (Bruford & Wayne, 1993), mitochondrial DNA sequences (Avise et al., 1987), or Single Nucleotide Polymorphisms (SNPs (e.g. Aitken et al., 2004; Seddon et al., 2005)).

Genetic diversity is usually assessed either by heterozygosity or allelic diversity. Heterozygosity measured by calculating the proportion of heterozygous individuals at any given locus, or the proportion of heterozygous loci for any given individual. Allelic diversity is simply the number of alleles, averaged across all loci (Freeland, 2005). Unfortunately, a basic estimate of genetic diversity based on a small number of genetic markers is not necessarily a good measure of the ‘health’ of a population. Genetic structure (subpopulations that do not freely mate with each other (Pritchard et al., 2000)) can lead to differential patterns of genetic diversity in different parts of a species’ range. Genetic structure is therefore important to identify for conservation

management. Also, nuclear markers are usually chosen in areas of the genome that are non-coding (Aitken et al., 2004), and therefore not under selection, as this should give a description of population genetic demography that is unbiased by selection pressure (Frankham et al., 2002). However, these supposedly neutral markers can be out of linkage disequilibrium (i.e. non-randomly associated) with genes under selection, and/or other neutral markers (Freeland, 2005), and linkage disequilibrium and selection therefore needs to be tested for molecular markers.

1.2.3 Aims of this phd research project

This thesis aims to use genetic tools to help elaborate:

- 1) The geographic range of okapi and their evolutionary diversity across that range;
- 2) The sociogenetic structure and dispersal characteristics of okapi populations using the most densely populated okapi reserve remaining and
- 3) Evaluate the captive population of okapi in comparison to their wild ancestors.

Each chapter in this thesis aims to investigate an aspect of one or more of these aims. These questions form the basis of the chapters of this thesis. The following sections introduce each of the chapters of this thesis, and explain the particular question that is being addressed in detail. These sections are intended to give added background information of justification for the study, and theory behind methodology carried out. Although the focus of this PhD project is clearly on okapi, many of the questions being investigated transcend taxonomic boundaries. Many of the investigations carried out therefore have implications for other geographically proximate species. Also, some of the theoretical analyses implemented provide a methodological framework that can be used much more widely.

1.2.3.1 CHAPTER TWO - Microsatellite loci for the okapi (*Okapia johnstoni*)

As discussed above, okapi are a highly elusive animal, making study by direct observation virtually impossible. Non-invasive genetic methods therefore provide a potentially useful means of studying this organism. Mitochondrial DNA

has been used for studying non-model organisms for many years (Awise et al., 1987; Awise, 1998), due to it containing highly conserved regions, high copy number and, barring some exceptions, lack of recombination, well-studied, easy to extract (Ladoukakis & Zouros, 2001; Burzynski, 2003) and maternally inherited (Gyllensten et al., 1991; Schwartz & Vissing, 2002; Kvist, 2003). However molecular ecological studies now usually utilise nuclear markers also, with microsatellites becoming the marker of choice over the past ~20 years (Bruford & Wayne, 1993; Beheregaray, 2008). It is important to use a combination of different molecular markers for ecological studies where possible, as results based on a single, non-combining locus (i.e. mitochondrial DNA) may give a biased and/or limited perspective of a species' evolutionary history, as it is essentially a single data point, and is therefore not be representative of the complete genome (e.g. Roca et al., 2007). Developing *de novo* microsatellite markers is a more complex process than mitochondrial markers as interspecific PCR amplification rates are usually lower for microsatellites than for mitochondrial DNA (Freeland, 2005). This chapter therefore describes the development of a set of 13 polymorphic microsatellite loci for use in non-invasively collected samples. Microsatellite markers can, and have been used to describe processes such as dispersal (Goudet et al., 2002; Goossens et al., 2005; Zhan et al., 2007), sociality (Möller, 2012; Ribeiro et al., 2012) and mating systems (Blyton et al., 2012). Our markers were therefore intended for use in investigating these processes in okapi, and this is discussed in Chapter 5. Microsatellites are also frequently used for investigating the genetic structure of populations (Pritchard et al., 2000; Brown et al., 2007), and this is one of the objectives investigated in Chapter 6 for wild, founder and captive okapi populations.

1.2.3.2 CHAPTER THREE - Non-invasive genetic identification confirms okapi (*Okapia johnstoni*) presence southwest of the Congo River

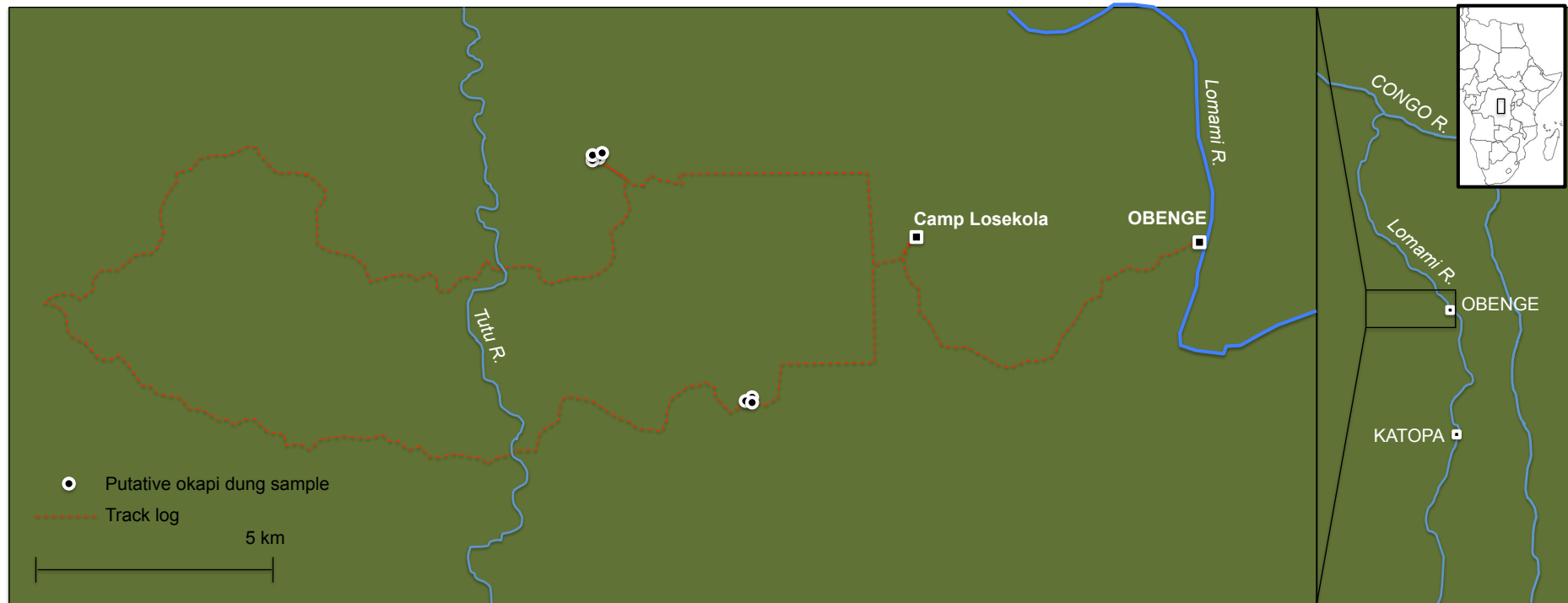
Some of the most basic questions that ecologists can ask for any given organism are: "How many are there?" and "Where do they live?" Without this information, conservation efforts are likely to be ineffective and/or futile. Distribution maps of okapi are often contradictory, with some indicating okapi presence on only the

northeast side of the Congo River (e.g. IUCN (2008), Figure 1.2A; Kingdon (1997), Figure 1.2B), but others indicating a potential, or historic distribution spanning the river (e.g. Stuart and Stuart (1997), Figure 1.2C). For the limited number of okapi population censuses that have been carried out (discussed above, Hart et al. (2008) and Vosper et al. (2012)), the most important source of okapi sign is faecal samples. Field researchers have considered it likely that okapi faeces is being confused with bongo (*Tragelaphus eurycerus*) faeces in the wild (J. Hart and S. Nixon Pers. Comm.). If this were the case, it would mean that okapi census estimates may be inaccurate, which could explain the high variation in okapi census estimates that has been observed in the surveys discussed above (Hart et al., 2008; Vosper et al., 2012). Therefore clear descriptions for some of even these most basic questions of distribution and abundance are missing for this species. This chapter aims to make a first attempt at answering these questions for okapi. This study used faecal samples collected from four quite geographically distant regions. This study also used samples collected from the RFO, as part of a separate large-mammal survey (Vosper et al., 2012), as well as other, opportunistically collected samples from a range of wildlife surveys. In addition to these, we carried out a more okapi-focussed survey in the TL2 region (Figures 1.6 and 1.7). The main aim of this survey was to identify okapi sign, and collect putative okapi faecal samples. The survey consisted of two expeditions, the first of which (Figure 1.6) was unsuccessful in collecting any putative okapi faecal samples, due to refusal to be allowed to enter the forests by individuals from neighbouring villages. However, these individuals claimed that okapi do not occur in local forests. This information is anecdotal, but is in accordance with surveys that have previously been carried out in the region (Figure 1.3). The second expedition (Figure 1.7) consisted of a reconnaissance circuit, via multiple forest clearings, where large mammals sign would likely be easier to identify. Where possible, putative okapi signs were followed in an attempt to local faecal samples. GPS coordinates were recorded for any faecal samples found, and these samples included with others that had previously been collected from this region.

Figure 1.6 First expedition in the TL2 (Tsuapa/Lualaba/Lomami) region to investigate okapi presence. Total distance covered was 80.4 km by foot and 13.4 km by bicycle, over a period of three days (August 2011).



Figure 1.7 Second expedition in the TL2 (Tsuapa/Lualaba/Lomami) region to investigate okapi presence. Total distance covered was 75.6 km by foot, over a period of six days (August 2011).



This chapter therefore uses a faecal sample dataset that covers several regions within the okapis range. The chapter investigates the proportion of samples that were collected that were correctly identified as okapi. The only species that is likely to be confused with okapi in the wild is bongo (J. Hart and S. Nixon Pers.Comm.). This chapter also investigates the possibility of developing a genetic test that uses just PCR amplification and gel electrophoresis to differentiate between these two species. This information is then used to determine if okapi do occur on the southwest side of the Congo River.

1.2.3.3 CHAPTER FOUR - Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered okapi (*Okapia johnstoni*)

Phylogeography is a field concerned with the principles and processes governing the geographic distributions of genealogical lineages (Avice et al., 1987; Avice, 1998). Phylogeography has the potential to identify genetically distinct populations, map dispersal of taxa through a region, identify speciation (Bermingham & Moritz, 1998) and to delineate evolutionary significant units (ESUs; Fraser & Bernatchez, 2001). These factors may be determined due to the fact that historical biogeographic factors, contemporary ecology and behaviour of organisms can play an important role in shaping the genetic architectures of extant species (Avice et al., 1987; Avice, 1998). This information can be of use simply to further scientific understanding of species, ecosystems and biogeography of a region (e.g. Moodley & Bruford, 2007; Lorenzen et al., 2010; Mboumba et al., 2011; Nicolas et al., 2011). It is also used however, in a more applied sense when deciding management policy, and is essential for highlighting priority areas, identifying barriers to movement, and planning translocations (Pennock & Dimmick, 1997).

Phylogeography has never been implemented before in okapi, but it has been previously investigated in giraffe (Hassanin et al., 2007). This study showed that northern and southern giraffes were genetically distinct. It also showed that *G. c. peralta* contains only Niger giraffes, whereas *G. c. antiquorum* appears to include populations living in Cameroon, Central African Republic, Chad and

southwestern Sudan. Brown et al. (2007) also demonstrated extensive population genetic structure between giraffe subspecies, using both mitochondrial DNA and nuclear microsatellites. As part of their analyses, this study used the program MDIV (Nielsen & Wakeley, 2001) to estimate divergence times between the giraffe clades.

This chapter investigates phylogeographic patterns in okapi. Specifically, this study aims to describe the genetic structure and evolutionary history of okapi, compared to other African ungulates including the giraffe, and to use this information to shed light on the biogeographic history of Congo Basin fauna in general. As mentioned above, the okapi range may span the Congo River. The Congo has been implicated in creating and maintaining ~1million years of genetic isolation, leading to the speciation between chimps and bonobos (Won & Hey, 2005; Caswell et al., 2008; Hey, 2010; Kawamoto et al., 2013). If okapi are shown to be present on the SW side of the Congo River, it will be important to identify if this geographic distinctness is also associated with genetic, or even taxonomic distinctness.

This study uses faecal samples, skin samples taken from trophies in Congolese villages, and museum specimens. Museum specimens can provide an invaluable source of samples for molecular ecology studies. These samples may have originated from locations that are currently inaccessible, or populations that are now extinct (Stanton et al., In Press), and may lead to a more complete description of a species' demographic history than based on extant populations alone (Mondol et al., 2013).

However, molecular damage and contamination by exogenous DNA can lead to difficulties when using DNA from museum samples. Following the death of an organism, DNA molecules can be degraded by bacteria, fungi and/or insects. These processes fragment DNA molecules, and can eventually lead to total degradation. DNA can also be fragmented by oxidation of deoxyribose residues. Further to degradation of DNA, DNA modifications can lead to incorrect calling of bases (Pääbo et al., 2004). To guard against these potential issues, studies that utilise museum

samples should design primers that amplify sequences as short as possible, and repeats can be carried out to ensure validity of results.

In order to thoroughly describe the biogeography of the Congo Basin, this study implements an Approximate Bayesian Computation (ABC) methodology (Beaumont et al., 2002) to estimate population divergence times using the program POPABC (Lopes et al., 2009). The POPABC analysis shares elements of coalescent theory (Wakeley, 2008) with the MDIV analysis implemented in the Brown et al. (2007) study, however ABC methods potentially allow models with additional levels of complexity at the expense of some approximation using summary statistics. Chapter 4 also investigates mitochondrial DNA sequence divergences times within okapi. Chapter 4 also uses a comparative phylogeographic methodology to estimate both *absolute* and *relative* (to other African ungulate taxa) divergence dates.

1.2.3.4 CHAPTER FIVE - Advancing knowledge of the ecology of a highly elusive species, the okapi (*Okapia johnstoni*), using non-invasive genetic techniques

As discussed above, conservation efforts are unlikely to be successful if even the most basic information is lacking on a given species. Only one long-term ecological study of wild okapi has ever been carried out (Hart & Hart, 1989). In Chapter 5 we use non-invasive genetic methods to try to determine aspects of ecology in okapi. Specifically we investigate sociality, mating system and dispersal of a population of okapi in the Réserve de Faune à Okapis, DRC. Dispersal is a crucial factor in understanding a species' ecology, particularly in spatially structured populations, as it is one of the main drivers in species persistence (Bowler & Benton, 2005). Natal dispersal is the movement of an individual from its natal site to the place where it reproduces, or would have reproduced if it had survived and found a mate (Howard, 1960). A species' mating system can effect genetic variability, inbreeding, and adaptive potential, thereby influencing population growth rate and extinction risk (DiBattista et al., 2008). In Chapter 5 we consider four forms of mating system:

1. Long-term monogamy; defined as male and female individuals forming lifelong mating pairs.
2. Monogamy; defined as individuals mating with a single individual of the opposite sex in any breeding year, but with multiple partners over multiple breeding years.
3. Polygamy; defined as individuals mating with a restricted group of individuals of the opposite sex in successive reproduction attempts.
4. Promiscuity; defined as individuals usually mating with different individuals of the opposite sex in successive reproduction attempts.

Darwin (1871) was the first person to discuss these ecological processes in an evolutionary context, and ecologists have been attempting to create a theory that can explain, and predict the vast array of social structures, dispersal patterns and mating systems observed in the wild ever since. However, the complex interaction between each of these three processes suggests that a single theory being formed that could accurately explain or predict any of them for any given animal seems unlikely (Emlen & Oring, 1977; Greenwood, 1980; Clutton-Brock, 1989b; Clutton-Brock, 1989a; Lawson Handley & Perrin, 2007). Instead, studies have focussed on specific interactions. For example, dispersal has been shown to be influenced by geographic, environmental and anthropogenic factors (Wiens, 2001; Funk et al., 2005; Baguette & Van Dyck, 2007), and social structure and mating systems have each been shown to be affected by genetic structure (Sugg et al., 1996; Storz, 1999; Ross, 2001; Möller, 2012; Ribeiro et al., 2012).

Generalities do appear to exist however, for example, ungulates that utilize dense forested habitats tend towards forming a smaller social unit. This is because the coordination of a social group is difficult in a forest, especially if the animal is large (Eisenberg & Lockhart, 1972). Predator threat may also influence social structure. Since in a savannah habitat, an animal is likely to be seen before it is heard, forming social groups may be an effective defence strategy. A forest habitat however, affords both predator and prey with cover to avoid detection. In this situation, an animal at risk of predation is more likely to adopt a hiding

strategy (Eisenberg & Lockhart, 1972) and be predominantly solitary to reduce social interaction and therefore detection probability (Geist, 1974).

1.2.3.5 CHAPTER SIX - An integrated genetic management assessment for the okapi (*Okapia johnstoni*)

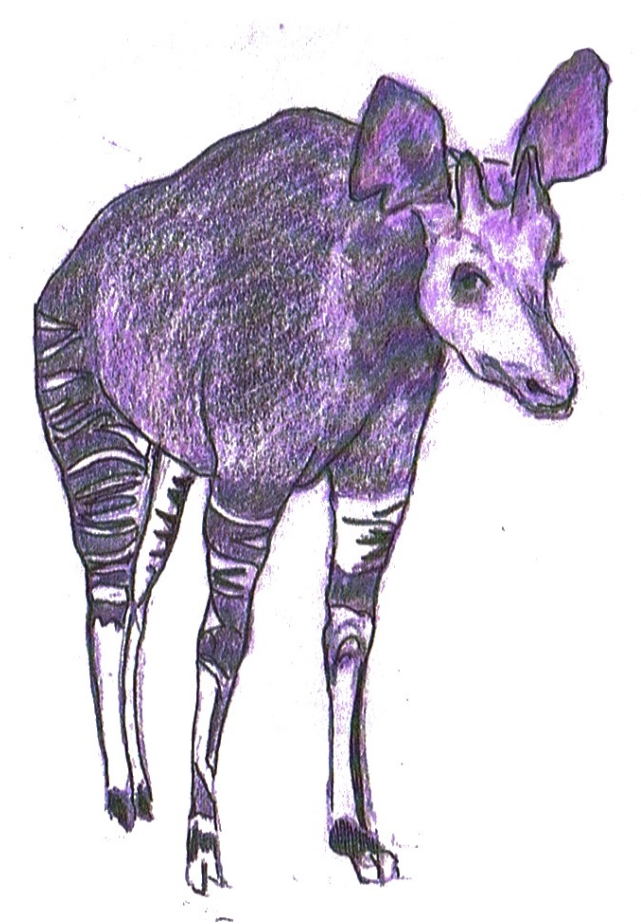
Captive populations can be an invaluable resource in conservation efforts (Rahbek, 1993; Woodworth et al., 2002). Many unsuccessful attempts were made to extract a living okapi from the wild in the years following the first scientific description of okapi in 1901, and after okapi were successfully transported out of Africa, it took many more years before okapi were successfully bred in captivity. Ebola, born in 1957 in Paris Zoo was the first okapi to be born and successfully bred in captivity. The okapi studbook currently lists 174 okapi in captivity (The Okapi Studbook, 2013), which descended from 34 wild-caught individuals that founded the captive population (Leus & Hofman, 2012).

Captive okapi are predominantly located in America (55%) and Europe (40%), with 4% of the global population in Asia and 1% in South Africa. There are carefully managed translocations across the ex-situ okapi populations, in order to minimise mean kinship between individuals (Leus & Hofman, 2012).

Chapter 6 extends the application of conservation genetics to the *ex-situ* okapi population. This chapter compares genetic diversity among the wild, founder and captive populations, and investigates if the captive okapi population, and the founders of that population, were/are genetically representative of the wild. Results are linked back to the descriptions of genetic diversity from Chapter 4. These results are particularly relevant considering the tragic events at the Epulu Conservation and Research Center, discussed above, and okapis Endangered conservation status.

2

CHAPTER TWO – Microsatellite loci development



"Professionalism is... and that is what I want."

David Brent

CHAPTER TWO - Microsatellite loci for the okapi (*Okapia johnstoni*)

David W. G. Stanton^{1*}, Linda M. Penfold², Xiangjiang Zhan¹, Michael W. Bruford¹

Published in: Conservation Genetics Resources (Accepted 16th April 2010)

¹ School of Biosciences, Cardiff University, Cathays Park, Cardiff, Wales, UK CF10 3AX

² White Oak Conservat Ctr, Yulee, FL 32097 USA

* Lead author contribution: All labwork, primer design and data analysis. Preparation of first manuscript draft, and modifications of manuscript based on co-author comments.

2.1 Abstract

We describe 13 polymorphic microsatellite loci for the okapi (*Okapia johnstoni*). These markers were tested with 20 samples collected from the Okapi Faunal Reserve and the Rubi-Tele hunting reserve, Democratic Republic of Congo, and exhibited a mean of 6.1 alleles per locus and a mean expected heterozygosity of 0.759. All but one locus was in Hardy–Weinberg equilibrium, and no evidence for linkage disequilibrium was detected between any loci. These loci will be useful for the future study of population genetic diversity and genetic structure in this elusive and emblematic species.

2.2 Article

The okapi (*Okapia johnstoni*) is an elusive forest ungulate both endemic to and iconic for the Democratic Republic of Congo (DRC), not described until 1901 or photographed in the wild until 2008 (Nixon & Lusenge, 2008). Okapi are at risk from poaching, and habitat fragmentation due to slash-and-burn deforestation as well as other anthropogenic disturbances such as roads and settlements. They are listed as near-threatened by the 2008 IUCN Red List (IUCN, 2008), based on the fact that the core population in La Réserve de Faune à Okapis (RFO) remains stable. However, recent post-war surveys report a 44% decline between 1996 and 2006 (C. Hicks Pers. Comm.) and large gaps in knowledge exist for this species. Primary steps towards its conservation require description of demographic features such as population structure, dispersal and gene-flow. Given the elusiveness of this species, such analysis requires a non-invasive approach. As part of an ongoing conservation effort, we report the isolation of microsatellite markers for this species.

Total DNA was extracted from four blood samples of okapi from White Oak Conservation Centre, Florida using a Qiagen Blood and Tissue Extraction Kit following manufacturer's instructions. Of these, one was chosen to construct a genomic library following the protocol of Glenn and Schable (2005).

Approximately 100 ng/μl of DNA was digested overnight with *HaeI* and *SspI* (New England BioLabs) and the digested DNA was ligated to Super SNX24 linkers. We electrophoresed the two ligated products in a 2% agarose gel and excised fragments between 300 and 700 bp and re-extracted the DNA using a peqGOLD Gel Extraction Kit (PeqLab). We then captured the fragments using a set of biotinylated microsatellite-containing oligo-nucleotides (mix 2 in Glenn and Schable 2005; (AG)₁₂, (TG)₁₂, (AAC)₆, (AAG)₈, (AAT)₁₂, (ACT)₁₂, (ATC)₈). The biotinylated probe-DNA complex was enriched by binding to -coated magnetic beads (Dynabeads M-280, Invitrogen). The DNA-probe streptavidin mixture was split into three repeats and non-specific DNA was removed from the probe mixture by washing twice with 2x SSC (Saline-Sodium Citrate), 0.1% SDS at room temperature, twice with 1x SSC, 0.1% SDS at room temperature, and the three repeats were washed two final times with 1x SSC, 0.1% SDS, at 45°C, 50°C and 55°C. The enrichment products were then combined, and after recovery by PCR using the forward SuperSNX-24 primer, the enriched library was constructed using a TOPO TA Cloning Kit according to the manufacturer's instructions (Invitrogen). Positive colonies were amplified using universal M13 forward and reverse primers (M13F: 5'-GTAAAACGACGGCCAG-3'; M13R: 5'-CAGGAAACAGCTATGAC-3') and fragments between 500 and 1200 bp (to try to maximize flanking region of any microsatellites found) were selected for sequencing, which was carried out by Macrogen Inc., Korea.

Sequences were assembled and edited in Sequencher version 4.9 (Gene Codes Corporation Inc.) and visually checked for microsatellite repeats. Of the 120 colonies analysed, 102 were successfully sequenced and 27 (26%) had repeats with more than 10 repeat units, melting temperatures between 50 and 66°C, calculated in MSATCOMMANDER (Faircloth, 2008), and PCR product lengths between 100 and 200 bp.

2 Microsatellite loci development

Table 2.1. Primer and microsatellite motif sequences, PCR conditions and genetic diversity for 13 okapi microsatellite markers.

Locus	Primers (F, forward; R, Reverse)	Motif	5'-Label	Ta (°C)	A	Size range (bp)	H _o	H _E	GenBank ID
Oka-01	F: AAGAGAGACTGCACTGTGGACC R: GCTCTTGTGTCTGACATGTTCTC	(AC)19	HEX	56	7	147-179	0.95	0.814	HM056235
Oka-02	F: CCAGGAAGAATCCTCATAATTGA R: TCCTAACAGCACCAGGGTTT	(TG)19	FAM	56	8	122-144	0.579	0.817	HM056236
Oka-03	F: GAAGAGGGAAGGCTGAAACA R: AAGCCAAGCCTGGTTCCC	(GT)17	FAM	56	6	96-120	0.5	0.735	HM056237
Oka-04	F: TCCAGTGTCTCTGGGCTTCC R: GCATCACACTCCTTGCTCAC	(CA)14	HEX	56	5	122-144	0.45	0.733	HM056238
Oka-05	F: CAGAGAAGACTGGTGGGCTACAG R: TGGTGGTAGTTCTCTGGGAGC	(AC)15	FAM	56	3	110-114	0.55	0.594	HM056239
Oka-06	F: CACACGTACATATGGACTC R: TCCTTCCCAGAGAGCTTGAC	(AC)14	FAM	56	6	155-169	0.95	0.801	HM056240
Oka-07	F: TGCCCCTCCCCCAACTCTCAA R: GCAGGGGACTCAACTTTGTTTCCT	(TG)27	HEX	56	8	183-209	0.6	0.86	HM056241
Oka-08	F: TGCCCCTCCCCCAACTCTCAA R: CCTGTGGGGCTCAGTGCT	(CA)15	NED	56	6	129-147	1	0.778	HM056242
Oka-09	F: AGGAGAGCAGGTAGTTCTCAGGA R: GAGCACAAGTGAAGTACCAAG	(TG)18	HEX	56	9	111-137	0.556	0.619	HM056243
Oka-10	F: GCAAGCAAGGTGGAATAC R: GGGAACGGATACATCTCAG	(CA)15	HEX	59	5	136-144	0.714	0.717	HM056244
Oka-11	F: GACAATGTAGGAGGGAGGCA R: ACTGGATGGCTGCAGTTTAG	(GT)30	FAM	59	11	161-189	1	0.902	HM056245
Oka-12	F: AGCAAGAAAGCATGACCTTACTCT R: ACTGACTAACATATGCGCATAAGC	(TG)13	FAM	59	5	97-113	0.789	0.713	HM056246
Oka-13	F: CTCGAGGGCATGTTTAGGAG R: AACCTCAGCATTTTCCGTCT	(TG)15	NED	59	7	125-145	0.65	0.747	HM056247

2 Microsatellite loci development

Table 2.1. Primer and microsatellite motif sequences, PCR conditions and genetic diversity for 13 okapi microsatellite markers.

Locus	Primers (F, forward; R, Reverse)	Motif	5'-Label	Ta (°C)	A	Size range (bp)	H _o	H _E	GenBank ID
Oka-01	F: AAGAGAGACTGCACTGTGGACC R: GCTCTTGTGTCTGACATGTTCTC	(AC)19	HEX	56	7	147-179	0.95	0.814	HM056235
Oka-02	F: CCAGGAAGAATCCTCATAATTGA R: TCCTAACAGCACCAGGGTTT	(TG)19	FAM	56	8	122-144	0.579	0.817	HM056236
Oka-03	F: GAAGAGGGAAGGCTGAAACA R: AAGCCAAGCCTGGTTCCC	(GT)17	FAM	56	6	96-120	0.5	0.735	HM056237
Oka-04	F: TCCAGTGTCTCTGGGCTTCC R: GCATCACACTCCTTGCTCAC	(CA)14	HEX	56	5	122-144	0.45	0.733	HM056238
Oka-05	F: CAGAGAAGACTGGTGGGCTACAG R: TGGTGGTAGTTCTCTGGGAGC	(AC)15	FAM	56	3	110-114	0.55	0.594	HM056239
Oka-06	F: CACACGTACATATGGACTC R: TCCTTCCCAGAGAGCTTGAC	(AC)14	FAM	56	6	155-169	0.95	0.801	HM056240
Oka-07	F: TGCCCCTCCCCCAACTCTCAA R: GCAGGGGACTCAACTTTGTTCCCT	(TG)27	HEX	56	8	183-209	0.6	0.86	HM056241
Oka-08	F: TGCCCCTCCCCCAACTCTCAA R: CCTGTGGGGCTCAGTGCT	(CA)15	NED	56	6	129-147	1	0.778	HM056242
Oka-09	F: AGGAGAGCAGGTAGTTCTCAGGA R: GAGCACAAGTGAAGTACCAAG	(TG)18	HEX	56	9	111-137	0.556	0.619	HM056243
Oka-10	F: GCAAGCAAGGTGGAATAC R: GGGAACGGATACATCTCAG	(CA)15	HEX	59	5	136-144	0.714	0.717	HM056244
Oka-11	F: GACAATGTAGGAGGGAGGCA R: ACTGGATGGCTGCAGTTTAG	(GT)30	FAM	59	11	161-189	1	0.902	HM056245
Oka-12	F: AGCAAGAAAGCATGACCTTACTCT R: ACTGACTAACATATGCGCATAAGC	(TG)13	FAM	59	5	97-113	0.789	0.713	HM056246
Oka-13	F: CTCGAGGGCATGTTTAGGAG R: AACCTCAGCATTTTCCGTCT	(TG)15	NED	59	7	125-145	0.65	0.747	HM056247

Twenty okapi DNA templates were extracted, four from the blood samples used to assemble the library; fourteen from faeces from wild-born and captive-born okapi collected at the Okapi Breeding and Research Station, Epulu, DRC; and two collected in the field in the RFO. PCR amplifications were performed in a GeneAmp® PCR System 9700 (Applied Biosystems) in a 6 µL reaction containing 1.5 µL of template DNA, 0.2 µM of each primer, 2.5 µL of 2x Multiplex PCR Master Mix (Qiagen, 3 mM MgCl₂, 3 x 0.85 ml) and 0.7 µg/µL BSA (New England Biolabs). The amplification conditions were as follows: 94 °C for 5 min, 45 cycles at 94 °C for 45 s, T_a for 30 s, 72 °C for 1.5 min and a final extension at 72 °C for 10 min. 14 primers, which produced a specific target band were labelled with one of the fluorescent dyes (5'-FAM, HEX; Eurofins MWG Operon and 5'-TAM; Applied Biosystems UK). PCR reactions were performed on all 20 DNA extracts as previously described but using the optimal annealing temperatures (Table 2.1). Amplification products were scanned together with the GeneScan™ **350 ROX™** Size Standard (Applied Biosystems), and fragment lengths were scored using Genemapper® version 4.0 (Applied Biosystems).

Thirteen polymorphic microsatellite loci provided reliable and consistent results and are detailed in Table 2.1. Genotyping success rate was 95.8% using a multiple-tubes approach (a consensus genotype is built by retaining alleles at each locus that occur above a set threshold; Navidi et al., 1992; Taberlet et al., 1996). The mean number of alleles was 6.8 (range 3–11). Observed and expected heterozygosities were calculated using ARLEQUIN version 3.5.1.2 (Excoffier & Lischer, 2010). The mean expected heterozygosity was 0.759 (range 0.594 – 0.902) and mean observed heterozygosity was 0.715 (range 0.450 – 1.000). Exact Hardy–Weinberg probabilities were assessed, and linkage disequilibrium was tested for, using GENEPOP version 4.0.10 (Raymond & Rousset, 1995; Rousset, 2008). Significance levels were adjusted using Bonferroni corrections for multiple testing ($P < 0.004$ in our dataset). All loci were in Hardy–Weinberg equilibrium except Oka-04 ($P < 0.001$), and no evidence was found for linkage disequilibrium between any pair of loci. The deviation from Hardy–Weinberg equilibrium for some loci may be due to sampling across populations that have some form of genetic structure. Caution is therefore advised when using the locus Oka-04 until further large-scale testing can be carried out on individuals from a single

population. These 13 microsatellite loci will be useful for the future study of population genetic diversity and genetic structure in the okapi.

2.3 Acknowledgements

The authors are grateful to Carolyn Sanguinetti and Brookfield Zoo for assistance with sample collection and processing and to Gilman International Conservation, for partial funding for this study. We would like to thank everyone in C5.15 for assistance in the lab. We would also like to thank The Damned for inspiring the project. Nibbled to death by an okapi, nibbled to death by an okapi...

3

CHAPTER THREE – Okapi distribution



“Lack of comfort means we are on the threshold of new insights.”

Lawrence M. Krauss

CHAPTER THREE - Non-invasive genetic identification confirms okapi (*Okapia johnstoni*) presence southwest of the Congo River

David W. G. Stanton^{1*}, John Hart², Ashley Vosper³, Noëlle F. Kümpel⁴, Jinliang Wang⁵, John G. Ewen⁵, Michael W. Bruford¹

Submitted to: Oryx (27th March 2014)

¹*School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom,*

²*Lukuru Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), 1235 Ave Poids Lourds, Kinshasa, DRC*

³*Wildlife Conservation Society (WCS), Great ape and human impact monitoring, Okapi Faunal Reserve, DRC*

⁴*Conservation Programmes, Zoological Society of London, London NW1 4RY, United Kingdom*

⁵*Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom*

* Lead author contribution: All labwork and data analysis. A subset of the sample collection. Preparation of first manuscript draft, and modifications of manuscript based on co-author comments.

3.1 Abstract

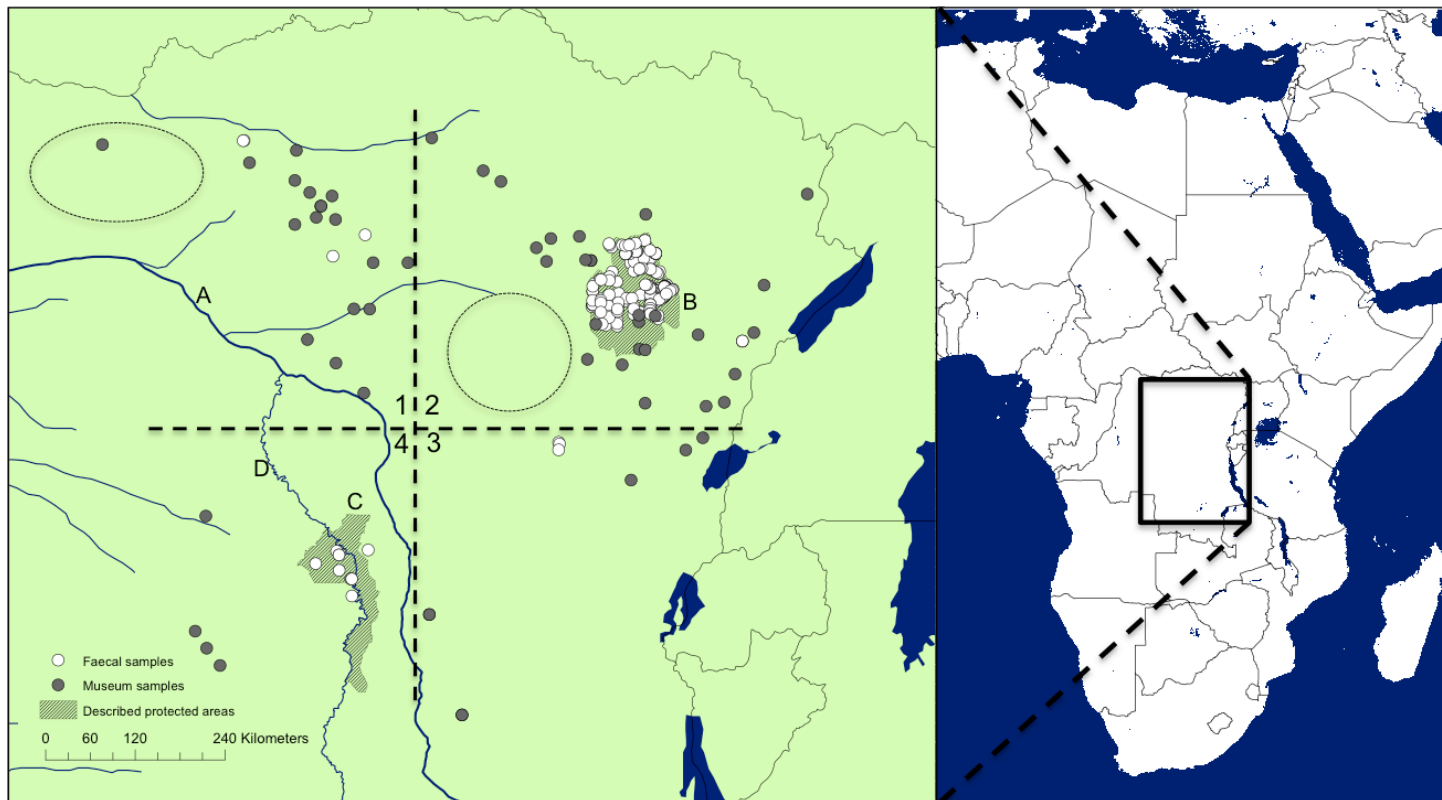
The okapi, a rainforest giraffid, endemic to the Democratic Republic of Congo, was recently reclassified as Endangered by the IUCN. Historic records and anecdotal reports suggest that okapi may have occurred in a population disjunct from the bulk of the range southwest of the Congo River. The current distribution and status of the okapi in this region is not well known. This study describes the use of non-invasive genetic identification in this species, assessing the success of species identification from putative okapi dung from the wild. Correct dung identification varied throughout the range, which may be related to varying okapi population densities and/or different sample collection strategies. Okapi were confirmed to occur on the southwest of the Congo River, in localized distributions west of the Lomami River. This study demonstrates that non-invasive genetic methods can provide information on distribution of cryptic, uncommon species that is difficult to get by other methods. Further investigation is now required to genetically characterise this species across its range, and to investigate the biogeographic processes that lead to the distribution that is observed in okapi and the other fauna in the region.

3.2 Article

Okapi is a monotypic species within the giraffidae, endemic to the DRC. Despite being recently reclassified as Endangered by IUCN much information is lacking regarding its current distribution and population sizes across the range (Mallon et al., 2013). Most of the okapi's range lies to the north and east of the Congo River (Kingdon, 1997; Stuart & Stuart, 1997; IUCN, 2008; Hart, 2013). However, there are historic records and anecdotal reports of okapi also occurring southwest of the river (Figure 3.1). It is important to determine the validity of these reports because rediscovery of a southern okapi population would imply a geographically and potentially evolutionarily distinct population and reinforce current efforts to gazette protected areas within this region.

Okapi are highly elusive, with the first successful camera trap image not produced until 2008 (Nixon & Lusenge, 2008), and are consequently very difficult to monitor in the wild. Dung counts have been previously used to determine okapi presence/absence (Beyers, 2008; Vosper et al., 2012). However, previous studies of other species have shown that accurately identifying animal dung visually can be difficult (Busby et al., 2009; Faria et al., 2011). Field researchers have considered it likely that okapi dung is being confused with bongo dung in the wild (J. Hart and S. Nixon pers. comm.). Genetic methods are increasingly being used for investigating ecological questions in situations where only non-invasively collected samples, such as dung samples, are available. Non-invasive genetics has recently been used to investigate species identification success in Central Africa (Faria et al., 2011), demographic history of orangutans (*Pongo pygmaeus*) in borneo (Goossens et al., 2005), dispersal of giant pandas (*Ailuropoda melanoleuca*) in China (Zhan et al., 2007), and individual identification of brown bears (*Ursus arctos*) in Europe. These non-invasive genetic methods therefore provide a useful way of testing the possibility of misidentification of okapi dung.

Figure 3.1. Map showing dung samples used in the present study (before molecular identification), and historic records of okapi presence (Royal Museum for Central Africa, Tervuren and the Centre de Recherche en Science Naturelles, Lwiro, DRC). Sampling regions are labelled 1-4 for clarity. The map features the Congo River (A), the okapi faunal reserve (B), the Lomami NP (C) and the Lomami River (D). Dashed ellipses indicate areas of local knowledge of okapi presence and/or evidence of okapi seen in a local village.



In this paper we present an update on okapi distribution south and west of the Congo River through use of genetic analysis of okapi dung. Our results show the potential utility of this non-invasive, cost effective method as a means to confirm species' identities where numbers are low and confusion with other species is possible. We used two mitochondrial DNA (mtDNA) primer pairs, OJ1 and OJ2 (OJ1-F [15162-15180]: ATGAATCGGAGGACAACCA, OJ1-R [15359-15380]: GGCCTCTTCTTTGAGTCTTAGG, 217 bp; OJ2-F [15359-15380]: CCTAAGACTCAAAGAAGAGGCC, OJ2-R [15525-15542]: TGCTGCGTTAAGGCTGTG, 184 bp) from Chapter 4. These primers were based on a single available okapi mitochondrial genome (Genbank accession number NC_020730.1). These primers were designed so that primer OJ1 was sited in the okapi Cytochrome b (Cyt b) and tRNA genes, whereas the reverse primer for OJ2 was sited in the control region (CR). Cyt b usually has a lower mutation rate than the CR (Ballard & Michael, 2004), meaning that primers designed within the Cyt b gene are more likely to successfully amplify across species that are taxonomically distant than the CR. These primer pairs (OJ1 and OJ2) may therefore be able to be used for species identification based on their relative amplification successes, without the need for DNA sequencing. The study utilised two hundred and forty-seven putative okapi dung samples, originating from throughout the okapi's known range. Seven samples were from sampling region one (see Figure 3.1 for definition), 209 from region two, seven from region three and 24 from region four. Faecal samples were collected either by, i) walking randomly placed transects through forest sites and collecting any faeces observed, or ii) identifying okapi prints and/or grazing and searching the surrounding area for faeces. Sampling methodology i) was used in areas of high okapi density (the okapi faunal reserve [RFO; Figure 3.1]), and sampling methodology ii) was used in areas of low okapi density (everywhere else in the range that faecal samples were found). Samples were collected by a number of different collaborators in the field, and since okapi population density is known to vary throughout the range, variation was expected in the species identification success between sampling regions. In addition, field workers were encouraged to collect all dung that *might* be okapi in areas where dung encounters rates were low, but were more stringent in areas of higher density. Therefore variations in species

identification success between sample collectors, does not necessarily reflect identification ability of the field workers. In sampling locations where it was not clear whether dung piles corresponded to one or more individuals, samples were DNA profiled (Stanton et al., 2010) to determine the number of different individuals that each sample comprised.

mtDNA PCR reagents and conditions are given in Tables 3.1 and 3.2 respectively. Products were visualised on a 3% agarose gel and sequenced by Macrogen Europe. Species ID was determined for successfully sequenced PCR products using the GenBank (BLAST) database. The formulae from Faria et al. (2011) were used to quantify and describe the accuracy of species identification based on the observed pattern of PCR bands on agarose gel before DNA sequencing has been carried out. Briefly, these formulae estimate the error of omission (dung from a given species is overlooked by this method), error of commission (dung is mistakenly attributed to a particular species when it was produced by a different species) and identification accuracy (rate of identification accuracy, both within species and overall). In the present study, the identification accuracy estimation is only reflecting the accuracy of this method for identifying okapi, and is therefore equivalent to one minus the error of commission.

Table 3.1. PCR reagents for the mtDNA PCR in the present study. PCR was carried out in a total volume of 25 µl. PCR for primers OJ1 and OJ2 were carried out in separate reactions.

PCR reagent	Final concentration/quantity	Supplier
Bovine serum albumin	4 µg	New England Biolabs (Ipswich, MA, USA)
PCR buffer	1X	Invitrogen (Merelbeke, Belgium)
MgCl ₂	2.5 mM	Invitrogen (Merelbeke, Belgium)
dNTPs	0.2 mM	Invitrogen (Merelbeke, Belgium)
(each) Primer	0.5 µM	Sigma (Gillingham, UK)
GoTaq	1U	Invitrogen (Merelbeke, Belgium)
DNA	2 µl	

Table 3.2. PCR conditions for the mtDNA PCR in the present study. PCR was carried out in a total volume of 25 µl.

94°C	3 mins	60x
94°C	30 secs	
58°C	35 secs	
72°C	45 secs	
72°C	5 mins	

Of the 247 dung samples tested 122 did not successfully amplify using either of the mtDNA primers (four of seven [57.1%], sampling region one; 101 of 209 [48.3%], sampling region two; four of seven [57.1%], sampling region three; 13 of 24 [54.2%], sampling region four). For the samples where at least one band was amplified, all three individuals from region one amplified both OJ1 and OJ2. From region two, three samples amplified OJ1 only, four amplified a band in OJ2 only and the remaining 94 amplified using both OJ1 and OJ2. All three working samples amplified OJ1 and OJ2 in region three and from region four, six of the samples amplified OJ1 only, the remaining six amplified both OJ1 and OJ2. All fragments from regions one, three and four were sequenced. From region two, 41 samples where both OJ1 and OJ2 amplified were selected for sequencing, as well as the seven that could only be amplified using one primer. All samples from regions one, two and three aligned with okapi when using Genbank (BLAST). In region four, all six samples where a band was amplified using OJ1 but not OJ2 aligned with bongo, whereas the five where a band was amplified using both OJ1 and OJ2 aligned with okapi. All samples that primer OJ2 could be amplified in corresponded to okapi samples, based on a Genbank (BLAST) of sequenced PCR products. The primer OJ2 was always unsuccessful in amplifying a fragment in bongo. Successful PCR amplification of primer OJ2 can therefore be used to differentiate between okapi and bongo without the need for DNA sequencing, with primer OJ1 acting as a positive control. Using the species identification formulae from Faria et al. (2011), the present study estimates an error of omission of 5.0%, an error of commission of 0% and an identification accuracy of 95%, of correctly identifying an okapi individual by the presence/absence of a band for primer OJ2. This is due to the primer pair OJ1 priming in the Cyt b and tRNA genes, whereas the reverse primer of the OJ2 primer pair primes in the CR, which is much less conserved between species (Table 3.3). Assuming this error of commission of 0% using the pre-sequencing method described above, of the 125 samples (containing usable DNA) used in the present study, 119 were okapi and 6 were bongo. All six bongo, misidentified as okapi, were from sampling region four, southwest of the Congo River. The six okapi samples found in region four comprised four individuals (based on the their microsatellite genotypes), which were all found on the western side of the Lomami River (Figure 3.1).

Table 3.3. DNA alignment of the priming site (shaded grey) and 29 bp of flanking sequence of the OJ2-reverse primer used in the present study. The alignment contains five okapi haplotypes from individuals used in the present study, and two bongo haplotypes downloaded from GenBank. Insertions/deletions are denoted by a hyphen, and bases identical to the consensus (shown in bold at the top) are denoted by a full-stop. This contig shows 11 and 12 polymorphisms, including five indels between the consensus okapi priming-site and each of the two bongo haplotype priming sites. The top row is the consensus genotype, and a “.” indicates that that position is identical to the consensus.

OJ_1	A	C	T	-	C	C	G	C	A	C	C	C	C	A	C	A	G	C	C	T	T	-	-	-	A	A	C	G	-	C	A	G	C	A	A	A	C	T	A	A	A	C	T	T	A	-	C	A	C	A	T
BONGO1	.	.	.	A	A	T	T	A	.	.	.	A	.	.	.	T	.	G	T	G	.	G	G	C	.	.	G	.	A	.	.	-	T	.	.	T	A	.	G	T	.	.	A	.	.	G	T
BONGO2	.	.	.	A	A	T	T	A	.	.	.	A	.	.	.	T	A	G	T	G	.	G	G	C	.	.	G	.	A	.	.	-	T	.	.	T	A	.	G	T	.	.	A	.	.	G	T
OJ_2	.	.	.	-	-	-	-	-	-
OJ_3	.	.	.	-	-	-	-	-	.	.	A	T	-	C
OJ_4	.	.	.	-	T	-	-	-	-	-	
OJ_5	.	.	.	-	-	-	-	-	-	.	G	.	.	.

The present study therefore shows definitively that okapi still do occur southwest of the Congo River, despite a lack of successful camera trap photos from this area. Okapi, appear to be localised and at low density (encounter rate in the Lomami area was less than 0.02 dung / km of transect and reconnaissance walked even in the zones where they are known to occur. This compares with encounter rates between 0.11 – 0.51 dung / km of transect walked in the RFO (Hart et al., 2008; Vosper et al., 2012). This study demonstrates the utility of a molecular approach for differentiating between okapi and bongo dung, the species with which okapi are most likely to be confused with in the field. In particular, we demonstrate the possibility of differentiating between okapi and bongo without the need for DNA sequencing. These primers therefore provide a simple diagnostic test to positively identify okapi from non-invasively collected samples, such as dung, using just PCR amplification and gel electrophoresis. It should be noted that this methodology is only being tested for positively identifying okapi samples. The error of omission of 5% implies that using this method the investigator may class some okapi as bongo (this is essentially allelic dropout, Taberlet et al 1999), however this is easily accounted for by sequencing the putative bongo samples. The error of commission of 0% implies a high accuracy for samples positively identified as okapi. It should be noted that it is unlikely that the dung of any species other than bongo, with an overlapping range to okapi, will be confused with either of these two species (J. Hart and S. Nixon pers. comm.).

Accurately determining distribution and abundance is a priority for species conservation (Jenkins et al., 2013) and is especially challenging in species that are cryptic, uncommon and locally distributed as is the case of the okapi in many parts of its range. Figure 3.1 shows regions within the putative okapi range where there is either local knowledge of okapi presence and/or evidence of okapi seen in a local village (Okapi Conservation Workshop, 2013). These areas should be priority areas for implementing the methods we discuss in the present study, in order to get a better idea of this species' current distribution.

While we confirm the continued occurrence of okapi in Congo's central basin, the species is clearly uncommon, and appears to no longer exist in some areas where the species was reported historically. Bongo, on the other hand, appear to be more common in the areas surveyed southwest of the Congo River than in the RFO (J. Hart pers. comm.). These differences in density are also a likely explanation for the higher occurrence of bongo dung samples from sampling region four. Due to the low abundance of putative okapi dung outside of the RFO, field teams were asked to collect anything that could be okapi dung, whereas for the RFO surveys, they were asked to differentiate between species. Taken together, these results imply that okapi and bongo dung *can* be correctly identified in the field by eye, however correct identification should not be assumed, as not all sampling regions had 100% identification success. Molecular methods should therefore be used wherever possible to confirm species IDs. These results show that the surveys carried out by Vosper et al. (2012) were not biased by species misidentification.

Of the samples found in the Lomami National Park, only the ones West of the Lomami River could be confirmed to be okapi. This does not imply that okapi are not present elsewhere southwest of the Congo River, however there is no current evidence that this is the case. The presence of okapi southwest of the Congo River raises the possibility of genetically differentiated okapi populations either side of the Congo (and Lomami) River, if, as seems likely, the river acts as a dispersal barrier for the species. In addition to containing this highly geographically distinct okapi population, a new Cercopithecine, the lesula (*Cercopithecus lomamienis*), was recently discovered in the Lomami National Park (Hart et al., 2012), highlighting the unique biodiversity of this future protected area. Interestingly the okapi, lesula and tholloni red colobus (*Procolobus tholloni*) are all species restricted to the west side of the Lomami River, and absent from the Lomami-Lualaba interfluve (Hart et al., 2012). Further studies are now required to investigate biogeographic processes that have influenced the fauna of this region, and to genetically characterize okapi throughout their range.

3.3 Acknowledgements

We gratefully acknowledge the assistance of the *Institut Congolais pour la Conservation de la Nature* (ICCN) for allowing the field collection and export of samples. We thank Gilman International Conservation, the Okapi Conservation Project, the Frankfurt Zoological Society, the Wildlife Conservation Society, the Lukuru Foundation/TL2 project, the Zoological Society of London (ZSL) and the considerable number of people who assisted in various ways with sample collection. In particular, we would like to thank John Fataki Bolingo, Bryna Griffin, Terese Hart, Chrysostome Kaghoma, Luaision, John Lukas, Kambale Magloire, Ephrem Mpaka, Stuart Nixon, Linda Penfold, Elise Queslin, Alex Quinn, Rosemary Ruf. This project was jointly funded by the UK Natural Environment Research Council (NERC) and ZSL (ZSL as NERC CASE industry partner and via an Erasmus Darwin Barlow Expedition grant). We also acknowledge the financial support provided by Gilman International Conservation (the Okapi Conservation Project), the US Fish and Wildlife Service, the UK's Darwin Initiative and the Mohamed bin Zayed Species Conservation Fund.

4

CHAPTER FOUR – Okapi phylogeography



"Scholars maintain that the translation was lost hundreds of years ago."

Ron Burgundy

CHAPTER FOUR - Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered okapi (*Okapia johnstoni*)

David W. G. Stanton^{1*}, John Hart², Peter Galbusera³, Philippe Helsen³, Jill Shephard^{3,4}, Noëlle F. Kümpel⁵, Jinliang Wang⁶, John G. Ewen⁶, Michael W. Bruford¹

Submitted to: PLOS ONE (7th March 2014)

¹*School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom*

²*Lukuru Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), 1235 Ave Poids Lourds, Kinshasa, DRC*

³*Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningen Astridplein 26, 2018 Antwerp, Belgium*

⁴*School of Veterinary and Life Sciences, Murdoch University, Perth 6150, Australia*

⁵*Conservation Programmes, Zoological Society of London, London NW1 4RY, United Kingdom*

⁶*Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom*

* Lead author contribution: All labwork and data analysis. A subset of the sample collection. Preparation of first manuscript draft, and modifications of manuscript based on co-author comments.

4.1 Abstract

The okapi is an endangered, evolutionarily distinctive even-toed ungulate classified within the giraffidae family that is endemic to the Democratic Republic of Congo. The okapi is currently under major anthropogenic threat, yet to date nothing is known about its genetic structure and evolutionary history. This information on genetic structure and evolutionary history is an important prerequisite for conservation management given the species' current plight. The distribution of the okapi, being confined to the Congo Basin and yet spanning the Congo River, also makes it an important species for testing general biogeographic hypotheses for Congo Basin fauna, a currently understudied area of research. Here we describe the evolutionary history and genetic structure of okapi, compared to other African ungulates including the giraffe, and use this information to shed light on the biogeographic history of Congo Basin fauna in general. Using nuclear and mitochondrial DNA sequence analysis of mainly non-invasively collected samples, we show that the okapi is both highly genetically distinct and highly genetically diverse, an unusual combination of genetic traits for an endangered species, and feature a complex evolutionary history. Genetic

data are consistent with repeated climatic cycles leading to multiple Plio-Pleistocene refugia in isolated forests in the Congo catchment but also imply historic gene flow across the Congo River.

4.2 Introduction

The okapi (*Okapia johnstoni*) is an evolutionarily distinct even-toed ungulate endemic to the Democratic Republic of Congo (DRC) that has recently been reclassified as 'endangered' by the IUCN (Mallon et al., 2013). The okapi also holds iconic status among the Congolese people, and is thus a potentially important conservation flagship and umbrella species for the region. However, the species is under major on-going threat from habitat fragmentation, human encroachment, regional armed conflict and poaching (IUCN, 2008). The okapi was recognised as a member of the Giraffidae family in 1901 (Lankester, 1901) and to date has only been the subject of one long-term *in situ* ecological study (Hart & Hart, 1989). No photograph of a live, free-ranging, wild okapi was believed to be in existence until the release of a camera-trap image in 2008 (Nixon & Lusenge, 2008). The enigmatic nature of this species is due to its elusive behaviour, affinity for dense rainforest, and the on-going political instability in the regions of the DRC where it occurs, severely limiting scientific study. One important component in conservation management of endangered species is an understanding of the genetic structure of species and populations. This includes an understanding of the causes of any observed genetic differentiation, such as major geographic and demographic barriers in the ancient and recent past (Frankham et al., 2002). Virtually nothing is known of the diversity or evolutionary history of *Okapia*, which has almost no fossil record, a likely consequence of the okapi's adaptation to closed-canopy forest where the conditions for fossilisation are poor (Harris et al., 2010). This study therefore utilised a comparative phylogeographic approach, including data from multiple African ungulates to contextualise the history and diversity of the species. We used this approach to provide relative divergence estimates using a single methodology, based on the application of common analyses across species (Bermingham & Avise, 1986; Taberlet et al., 1998).

The historic range of the okapi is thought to have included large sections of the central/eastern Congo Basin, although it is likely that they are currently confined to a small fraction of their former distribution (Stuart & Stuart, 1997). This relatively wide historic range potentially makes them an important model for investigating historical processes governing the biogeography of the fauna of this region, a subject that remains under considerable debate (Hamilton, 1975; Livingstone, 1975; Cowling et al., 2008). In the absence of fossils a phylogeographic approach needs to be taken, as has been done a number of times with widely distributed African species (Arctander et al., 1999; Nersting & Arctander, 2001; Muwanika et al., 2003; Clifford et al., 2004; Moodley & Bruford, 2007; Mboumba et al., 2011), but comprehensive investigations within the Congo Basin have been much less common (Eriksson et al., 2004; Kawamoto et al., 2013). The Congo appears to have a profound effect in partitioning faunal diversity. For example, the river is implicated in maintaining one million years of evolutionary divergence between chimpanzees and bonobos (Won & Hey, 2005; Caswell et al., 2008; Hey, 2010; Kawamoto et al., 2013), and is thought to be the most important feature for structuring species diversity of *Praomys* in the Congo Basin (Kennis et al., 2011). Many questions regarding central African biogeography cannot, however, be resolved currently due to a paucity of studies. In particular, there are very few studies investigating the role of the Congo River on *within* species genetic diversity (Kennis et al., 2011). Okapi are a potential model large mammal to help test competing biogeographic theories, and investigate the role of the Congo River on within species genetic diversity due to the okapis close association with closed-canopy rainforest and relatively wide historic distribution (compared to other studied taxa) across the Congo Basin, including both sides of the Congo River (Chapter 3).

Here we used mitochondrial and nuclear DNA sequences to provide the first molecular-informed description of the evolution of the okapi, to identify genetically distinct populations and plausible hypotheses for the evolutionary history of this species.

4.3 Methods

4.3.1 Study area and sampling

This study analysed 74 okapi samples, including feces ($n = 39$), museum specimens (sampled with permission from the museum of Central Africa, Tervuren, Belgium; preserved skin [$n = 11$] and bone [$n = 9$]; museum sample numbers: 12604, 8305, 14235, 14906, 12517-a, 14454, 11043, 750, 1193, 8011, 9726, 9727, 13991, 13242, 14236, 14234, 909, 13336, 15298, 15299) and clippings of dried skin ($n = 15$) from individuals that had previously been hunted (Figure 4.1). All samples from wild okapi were non-invasively collected. Permission for sampling was provided by the Institut Congolais pour la Conservation de la Nature (ICCN; permit numbers: 0996/ICCN/DG/ADG/MG/KBY/2011 and 090/ICCN/ADG/DG/KV/2012). Fecal samples were collected either by, a) walking randomly placed transects through forest sites and collecting any feces observed, or b) by identifying okapi sign and searching the surrounding area for feces. Sampling methodology a) was used in areas of high okapi density (the Okapi Faunal Reserve [RFO; Figure 4.1]), and sampling methodology b) was used in areas of low okapi density (everywhere else in the range that fecal samples were found). Skin samples were in the form of clippings taken from skins owned by individuals living in villages in, or near field sites. Museum samples (skin and bone) were sampled with permission from the Royal Museum of Central Africa, Tervuren. These samples were collected between September 1911 and May 1939 and their locations were obtained from information accompanying the samples, usually the name of a town/village, likely representing the closest habitation to where the individuals were hunted. Samples were grouped into one of four broad sampling 'regions' (see Figure 4.1) for later analysis and as a descriptive reference.

4.3.2 Molecular methodology

Five pairs of mitochondrial DNA primers were designed using an available okapi sequence (Genbank accession number: NC_020730.1), for PCR amplification in this study. These primers amplify a section of the mitochondrial genome that comprised 422 bp of the Cyt *b* gene, the tRNA-Pro and tRNA-Thr genes, and 473 bp of the CR, (OJ1-F [15162-15180, based on the above-mentioned okapi mt

genome]: ATGAATCGGAGGACAACCA, OJ1-R [15359-15380]:
 GGCCTCTTCTTTGAGTCTTAGG, 217 bp; OJ2-F [15359-15380]:
 CCTAAGACTCAAAGAAGAGGCC, OJ2-R [15525-15542]: TGCTGCGTTAAGGCTGTG,
 184 bp; OJ3-F [15495-15515]: CCCACAACAACCAACACAAAC, OJ3-R [15741-
 15761]: CGGGATACGCATGTTGACRAG, 247 bp; OJ4-F [15645-15665]:
 ATATGCCCCATGCATATAAGC, OJ4-R [15885-15905]:
 CCCTGAAGAAAGAACCAGATG, 263 bp; OJ5-F: CTACCATGAGGACAAATATCATT,
 OJ5-R: CATTTCAGGTTTGATATGAGG), with fragment sizes all below 300 bp to
 allow them to be used in low quality DNA samples. PCR was carried out with a
 total volume of 25 µl with 4 µg BSA (New England Biolabs, Ipswich, MA, USA), 1X
 PCR buffer (Invitrogen, Merelbeke, Belgium), 2.5 mM MgCl₂, 0.2 mM dNTPs, 0.5
 µM each primer, 1 unit of GoTaq (Invitrogen) and 2 µl of DNA. PCR conditions
 were as follows: 94°C for 3 mins; 60x cycles of 94°C for 30 secs, 58°C for 35 secs,
 72°C for 45 secs; and a final extension of 72°C for 5 mins. PCR products were
 visualised on a 3% agarose gel and sequenced by MacroGen Europe.

4.3.3 Nuclear DNA primers

To complement the mitochondrial dataset and provide a nuclear perspective on
 phylogeographic structure, we developed a set of nuclear EPIC (Exon-Priming
 Intron-Crossing) markers (Table S4.1). Forty-eight primer pairs (selected from
 the 'best 96' loci from Aitken et al. (2004) were tested on DNA extracted from
 blood samples of two captive individuals, preferentially choosing loci that were
 reported to amplify a single band in *Bos taurus*. PCR conditions followed Aitken
 et al. (2004). PCR products were visualised on a 3% agarose gel, and primers that
 produced a single band (n = 20) were tested in four dried skin samples from wild
 okapi (two from region one; one from region three; and one from region four).
 PCR mix and conditions were the same as above, except annealing times were
 increased to 1 min and the number of cycles on the second step of the touchdown
 PCR was increased to 40. PCR products for all four samples were sequenced
 (Eurofins MWG Operon, Ebersberg, Germany) in the forward and reverse
 direction. Primers were redesigned to amplify shorter fragments (~100 bp) for
 use with non-invasive samples in fragments that contained at least one SNP in
 the four samples that they were tested in (n = 14). All 14 primer pairs were then

tested in 6 non-invasively collected samples (feces and dried skins), and the 12 most consistent primer pairs were selected. These primer pairs were then used for the full set of 28 samples in this study. In all cases, SNPs were scored manually in Sequencher 4.9 (GeneCodes) by sequencing individual PCR amplicons in the forward and reverse direction. To characterise population genetic diversity for this set of 12 EPIC loci, we calculated summary statistics in Arlequin v3.5 (Excoffier & Lischer, 2010), using the full set of 28 samples.

4.3.4 Sequence analysis

Sequences were aligned in Sequencher 4.9 (GeneCodes) and four mitochondrial DNA contig alignments were created. These consisted of 372 bp of the Cyt *b* gene, the tRNA-Pro (66 bp) and tRNA-Thr (69 bp) genes and 328 bp of the CR, and a concatenation of the Cyt *b*, tRNA-Pro and tRNA-Thr genes. Contigs consisted of shorter fragments than the original PCR product amplified in order to minimize missing data. To visualise the sequence data, a network of the CR and concatenated sequences were drawn in TCS 1.21 (Clement et al., 2000). Pairwise and average nucleotide diversities were calculated in DAMBE v4.2.13 (Xia & Xie, 2001), as were amino acid translations for Cyt *b* sequences, and haplotype diversities were calculated in DNAsp v5 (Librado & Rozas, 2009). The presence of nuclear inserts of mitochondrial DNA (*Numt*, Bensasson et al. 2001) was assessed by, i) the presence of a single band on an agarose gel, ii) comparison with known mitochondrial DNA sequence on GenBank, iii) for Cyt *b* sequences: the lack of stop codons in the translated amino acid sequence and the lack of any markedly distinct amino acid substitutions.

Okapi mitochondrial CR nucleotide diversity (π) was compared to CR sequence diversity of a number of other African ungulates. The taxa compared were hartebeest (*Alcelaphus buselaphus* spp; six subspecies), bontebok (*Damaliscus pygargus*), giraffe (*Giraffa camelopardalis* spp; six subspecies), roan antelope (*Hippotragus equinus* spp; five subspecies), African buffalo (*Syncerus caffer*), common eland (*Taurotragus oryx*) and bushbuck (*Tragelaphus scriptus* spp; 21 subspecies), chosen based on the availability of CR sequences in Genbank (studies and GenBank IDs are given in Table S4.3). The sequences from all taxa

were aligned, including any flanking regions, and the start position of the CR for this complete contig was identified, based on the annotation from GenBank. This was to ensure that a homologous section was being compared between taxa. A ubiquitous section of the complete contig was then separated out into the taxon groups shown in Table 4.1 and re-aligned. This re-alignment consisted of between 268-275 bp for each taxon, including indels. Indels were included in all π calculations. π and SD for Position 31 – 306 of the CR was calculated for each of the eight taxon groups described in Table 4.1, in DAMBE v4.2.13 (Xia & Xie, 2001).

4.3.5 Partitioning of genetic diversity

To partition relative contributions of genetic diversity, AMOVA statistics were calculated using Arlequin v3.5 (Excoffier & Lischer, 2010) using the four sampling regions defined in Figure 4.1. These regions attempt to cover as much of the okapi's known range as possible and comprised the Rubi-Tele Hunting Reserve and surrounding areas (region 1), the RFO (Okapi Faunal Reserve) and surrounding areas (region 2), the Aruwimi/Lindi/Tshopo (ALT) Rivers and Maiko National Park (region 3) and the Tsuapa/Lomami/Lualaba (TL2) Rivers (region 4). AMOVA groupings were then tested in different combination using these four regions. A total of 69 CR and 53 concatenated sequences were used for the AMOVA analysis.

A total of 28 individuals from the four sampling regions (region 1, n = 5; region 2, n = 14; region 3, n = 4; region 4, n = 5) were used with the EPIC loci. Of the twelve EPIC loci investigated, four contained greater than one SNP. All SNPs within one sequence were presumed to be linked. Therefore, for analyses using only SNPs (i.e. not the intron sequences), one SNP with high polymorphism was chosen from each of those three intron sequences. AMOVA statistics and F-statistics were calculated on the SNP data using Arlequin v3.5 (Excoffier & Lischer, 2010) with the same “groups” and “populations” used for the mitochondrial DNA.

4.3.6 Population and sequence divergences

To investigate inter- and intra-specific okapi mitochondrial lineage divergences, a time-calibrated phylogeny was created for okapi and giraffe Cyt *b*, tRNA-Pro and tRNA-Thr concatenated sequences using BEAST v1.7.5 (Bayesian Evolutionary Analysis by Sampling Trees; (Drummond & Rambaut, 2007)), with red deer (*Cervus elaphus*) as an outgroup (due to Cervidae being the taxonomically most proximate sister taxon; (Hassanin et al., 2012)). An HKY +Gamma model was used for the concatenated fragment, selected using jModelTest v2.1.1 (Guindon & Gascuel, 2003; Posada, 2008). The okapi and giraffe tree was constructed using lognormal relaxed and strict clocks, and with yule speciation, coalescent constant, coalescent expansion growth and speciation birth death tree models. The $2 \times \ln$ Bayes factors for these trees were calculated using Tracer v1.5 with 1000 bootstrap replicates to determine the most appropriate model. A value of greater than ten was taken as strong evidence for supporting a model, following Kass and Raftery (1995). In all cases, the MCMC chain was set at 20,000,000 iterations, with three repeats combined to create the final tree. Tracer v1.5 (Rambaut & Drummond, 2007) was used to assess the MCMC output of all BEAST runs. Divergence times and corresponding standard deviations were taken from Hassanin et al. (2012). To contextualise the results of the Giraffidae phylogeny, a phylogeny was also constructed including okapi, giraffe, duiker and bushbuck jointly. Pig (*Sus scrofa*) and collared peccary (*Pecari tajacu*) were selected as outgroups, based on the mtDNA phylogeny of Hassanin et al. (2012) giving good support for these species occurring within Cetartiodactyla, but outside Ruminantia. The okapi, giraffe, duiker and bushbuck tree used the 274 bp of Cyt *b* sequence that was overlapping between the present study, and the sequences from Genbank (Table S4.4).

PopABC (Lopes et al., 2009) was used to model divergence times between present-day okapi populations, as well as to infer present-day and ancestral effective population sizes, using the EPIC and mitochondrial sequences. For the EPIC loci, haplotypes were reconstructed using Phase v2.1.1 (Stephens et al., 2001). Pairwise analyses were carried out between sampling regions 1v2, 2v4 and 1v4 (Figure 4.1). The number of iterations used was $1e^6$, with the rejection

step set at $1e^{-5}$. In order to determine prior ranges, preliminary runs were carried out, starting with very wide priors and altering them until all the posterior distributions were distinct from the priors (priors are given in Table S4.5).

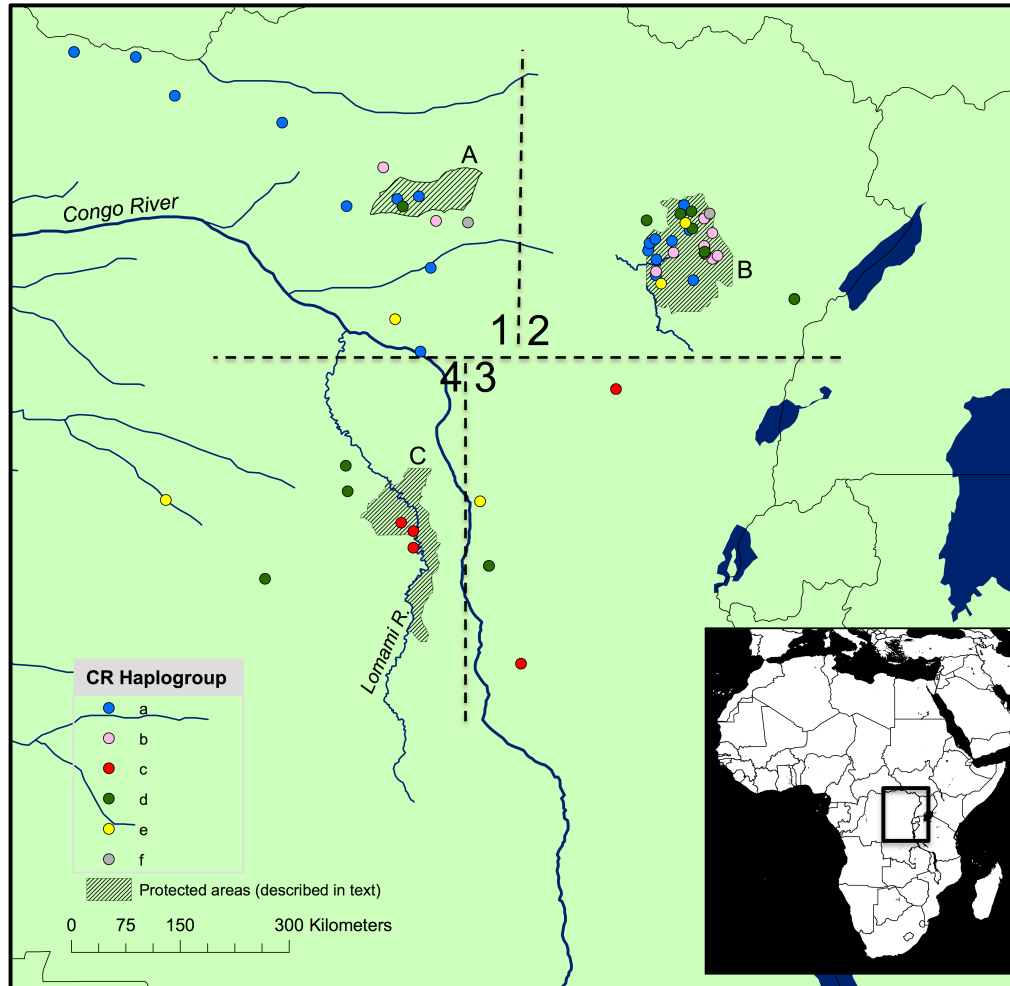
4.4 Results

4.4.1 Primer design

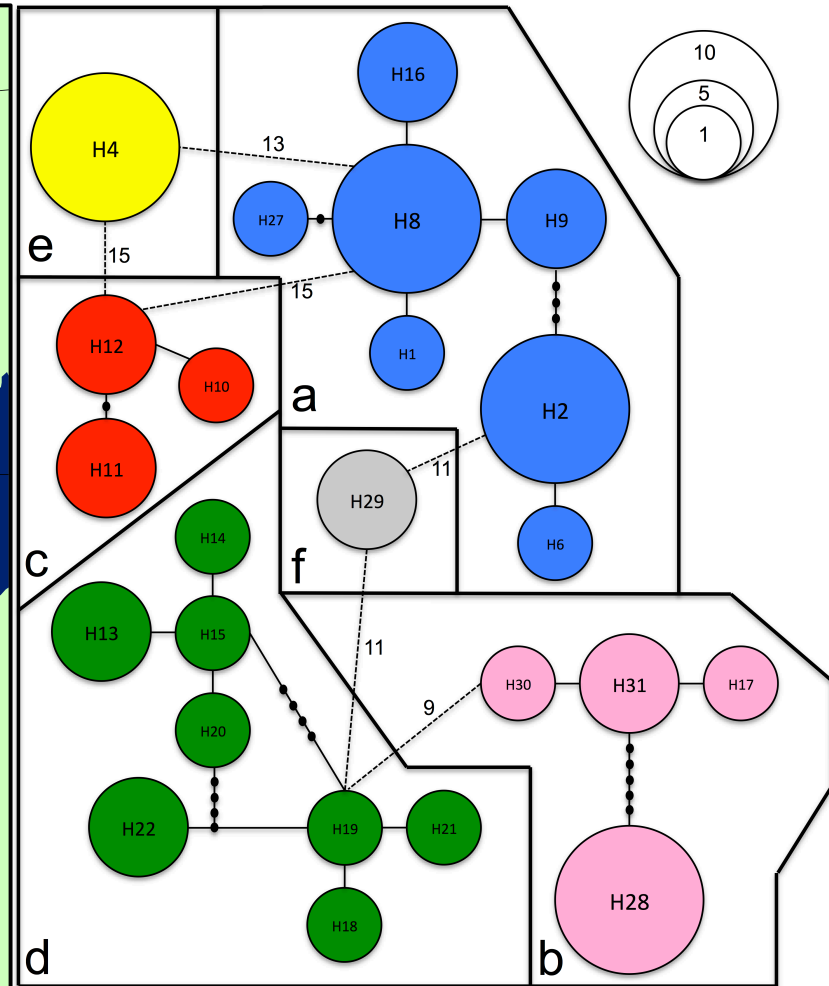
We developed a set of primers that amplified four mitochondrial (total of 895 bp of control region [CR], cytochrome *b* [Cyt *b*], tRNA-Pro and tRNA-Thr) genes and 12 nuclear intron (exon priming intron crossing [EPIC]; Table S4.1) loci. Following Bonferroni correction, none of the nuclear loci were found to be in linkage disequilibrium or to be out of Hardy-Weinberg Equilibrium (HWE). For HWE testing, sampling regions 1-4 (Figure 4.1) were analysed separately. A summary of nuclear SNP variation is given in Table S4.1. Seventy-two samples could be used to produce mitochondrial CR data, 61 samples could be used to produce a concatenated sequence of mitochondrial Cyt *b*, tRNA-Thr and tRNA-Pro data and 28 samples could be used to produce nuclear intron EPIC (Aitken et al., 2004) sequences.

Figure 4.1 (next page). Okapi samples used in the present study, with the colour relating to the adjacent network (Clement et al., 2000), based on 473 bp of the CR. For the network, TCS connected alleles with a 95% confidence limit, those that did not fall within that limit are connected with dotted lines (with numbers corresponding to the number of mutations) and separated by solid black lines. These haplotype groupings are then used as haplogroups for description in the text. Sampling locations are arbitrarily labelled 1-4 for reference in the text. The cluster of samples labelled as RFO corresponds to the Okapi Faunal Reserve. Key protected areas are labelled A (Rubi-Tele Hunting Reserve), B (Okapi Faunal Reserve, RFO) and C (Lomami National Park).

1



2



4.4.2 Sequence analysis

The networks (TCS) of mitochondrial CR is shown in Figure 4.1. Six distinct haplogroups were recovered (number of pairwise differences was always higher between haplogroups than between any two haplotypes within a haplogroup), with some appearing to be geographically restricted while others were not. The most common CR haplogroup (a) was present in the northern regions (i.e. sampling regions one and two) but not southern regions (i.e. sampling regions three and four), as were haplogroups b and f. Haplogroup d was however detected in all four sampling regions (and therefore also on both sides of the Congo), as was haplogroup e. In contrast, haplogroup c was much more common southwest of the Congo River (sampling region four; 42.9% of all samples from this sampling region) compared to the northeast side (4.3% of all samples from sampling regions one, two and three). For the concatenated sequences, all of the samples from the southwest side of the Congo possessed either the haplotype H11 or H12, which were equally frequent. These two haplotypes were highly differentiated and both also present on the northeast side of the Congo River, but at much lower percentages (10.6% and 4.3% respectively).

Haplotype and nucleotide diversities, and number of polymorphic sites for the mtDNA genes investigated in the present study are given in Table S4.2. CR nucleotide diversity was compared to that of a number of other African species (see Table 4.1). Based on a common 275 bp of CR sequence, the combined bushbuck (*Tragelaphus scriptus spp.*) dataset showed the highest haplotype diversity (0.151), with bushbuck ecotypes showing highly variable nucleotide diversity estimates (0.007 – 0.092; Table 4.1). Nucleotide diversity in okapi (0.045) was slightly lower than the combined giraffe dataset (0.052), and very similar to the African buffalo (0.045), and higher than the eland antelope (0.038).

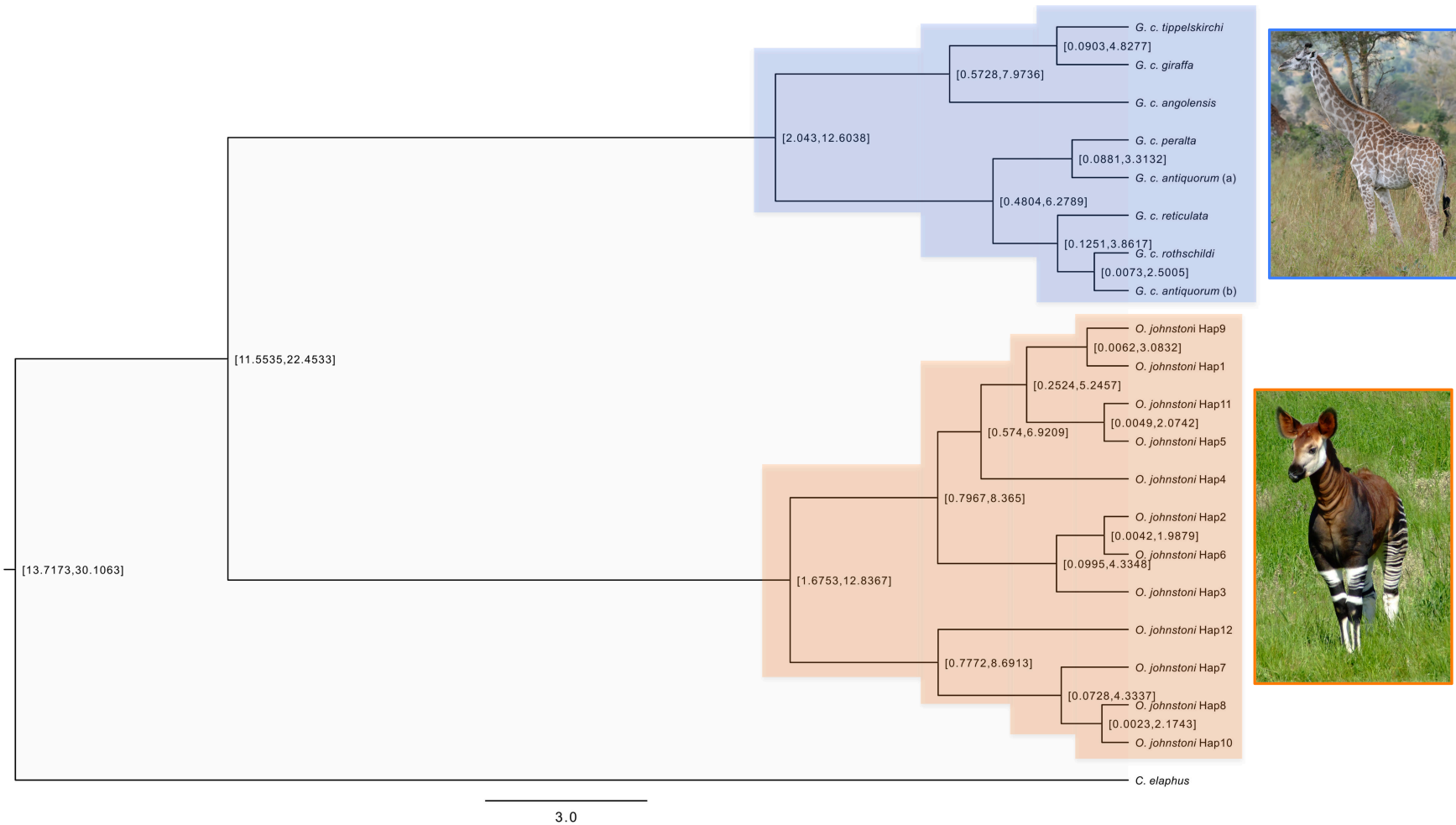
Table 4.1. Nucleotide diversity in 268 - 275 bp of homologous CR sequences of African ungulates, sorted on pi. Key taxa are shown in bold. These often correspond to the combined calculations for several species or subspecies.

Taxon	pi	SD	Number of haplotypes
<i>Tragelaphus scriptus</i> spp	0.151295	0.072901	197
<i>Tragelaphus stricptus barkeri</i>	0.092222	0.044905	5
<i>Alcelaphus buselaphus</i> spp.	0.085421	0.041796	92
<i>Tragelaphus stricptus roualeyni</i>	0.075681	0.037075	13
<i>Tragelaphus stricptus massaicus</i>	0.069069	0.033945	18
<i>Alcelaphus buselaphus lichtensteini</i>	0.058252	0.028857	56
<i>Tragelaphus stricptus scriptus</i>	0.056282	0.027889	28
<i>Tragelaphus stricptus sylvaticus</i>	0.053813	0.02672	11
<i>Giraffe camelopardalis</i> spp.	0.051564	0.025654	29
<i>Giraffa camelopardalis rothschildi</i>	0.050059	0.024946	4
<i>Hippotragus equinus</i> spp.	0.046574	0.023565	43
<i>Alcelaphus buselaphus swaynei</i>	0.04647	0.02328	4
<i>Syncerus caffer caffer</i>	0.045189	0.022629	60
OKAPIA JOHNSTONI	0.045015	0.022561	26
<i>Alcelaphus buselaphus lelwel</i>	0.043154	0.021669	5
<i>Alcelaphus buselaphus major</i>	0.041606	0.020936	6
<i>Giraffa camelopardalis giraffa</i>	0.041507	0.020894	4
<i>Tragelaphus stricptus diana</i>	0.038925	0.019665	4
<i>Giraffa camelopardalis reticulata</i>	0.03833	0.019387	8
<i>Taurotragus oryx</i>	0.037956	0.019322	50
<i>Tragelaphus stricptus knutsoni</i>	0.037463	0.018972	23
<i>Tragelaphus stricptus fasciatus</i>	0.037455	0.018967	5
<i>Tragelaphus strepsiceros strepsiceros</i>	0.035514	0.01807	24
<i>Tragelaphus stricptus heterochrous</i>	0.034188	0.017418	3
<i>Tragelaphus stricptus meruensis</i>	0.033939	0.0173	3
<i>Alcelaphus buselaphus caama</i>	0.033708	0.017214	10
<i>Alcelaphus buselaphus cokei</i>	0.030588	0.015737	11
<i>Giraffa camelopardalis tippelskirchi</i>	0.029038	0.014982	11
<i>Tragelaphus stricptus sassae</i>	0.024374	0.012757	8
<i>Hippotragus equinus cottoni</i>	0.021818	0.011541	3
<i>Tragelaphus stricptus pictus</i>	0.021818	0.011541	3
<i>Tragelaphus stricptus decula</i>	0.01705	0.009268	8
<i>Tragelaphus stricptus meridionalis</i>	0.014545	0.00807	3
<i>Tragelaphus stricptus bor</i>	0.013857	0.00774	5
<i>Tragelaphus stricptus meneliki</i>	0.012165	0.006927	9
<i>Giraffa camelopardalis angolensis</i>	0.011679	0.006697	5
<i>Tragelaphus stricptus cottoni</i>	0.007273	0.004552	7
<i>Tragelaphus stricptus dama</i>	0.007273	0.004552	3
<i>Tragelaphus stricptus johannae</i>	0.007273	0.004552	5
<i>Tragelaphus stricptus phaleratus</i>	0.007273	0.004552	4
<i>Damaliscus pygargus</i>	0.004902	0.003371	3
<i>Giraffa camelopardalis antiquorum</i>	0.004866	0.003356	3

4.4.3 Partitioning of genetic diversity

Partitioning of mitochondrial and nuclear DNA sequence variation was investigated across the geographic range, with initial groups being defined as either sampling region 1 and 2 *versus* region 3 and 4 (i.e. north *v.* south), or region 1, 2 and 3 *v.* region 4 (i.e. each side of the Congo River). For the north *v.* south comparison for CR sequences, percentage of variation among groups was 19.28% (AMOVA; $p = 0.334$), among populations (i.e. sampling regions) within groups 3.91% ($p = 0.127$) and within populations was 76.81% ($p < 0.001$). Comparing each side of the Congo River for CR sequences, percentage of variation among groups was 12.30% ($p = 0.248$), among populations within groups 10.66% ($p < 0.001$) and within populations was 77.03% ($p < 0.001$). For the north *v.* south comparison for the concatenated sequences, percentage of variation among groups was 18.56% ($p = 0.337$), among populations within groups 9.14% ($p = 0.081$) and within populations was 72.30% ($p < 0.001$). Comparing each side of the Congo River for the concatenated sequences, percentage of variation among groups was 29.65% (AMOVA; $p = 0.253$), among populations within groups 7.49% ($p = 0.005$) and within populations was 62.87% ($p < 0.001$). For the north *v.* south comparison for nuclear SNPs, the percentage of variation among groups was 2.30% ($p = 0.329$), among populations within groups 4.22% ($p = 0.129$), among individuals within populations 45.16% ($p < 0.001$), and within individuals was 48.31% ($p < 0.001$). Comparing each side of the Congo River for SNPs, percentage of variation among groups was zero ($p = 0.509$), among populations within groups 6.77% ($p = 0.029$), among individuals within populations 46.06% ($p < 0.001$), and within individuals was 49.27% ($p < 0.001$).

Figure 4.2. Giraffidae phylogeny drawn in BEAST v1.7.5 (Drummond & Rambaut, 2007), with red deer (*Cervus elaphus*) as an outgroup. Numbers on branches correspond to 95% HPD confidence intervals.



Significant F_{ST} values for CR sequences were found between all sampling regions except 3 v. 4 ($F_{ST} = 0.013$, $p = 0.373$), all other F_{ST} values were significant at $p < 0.01$, except between regions 1 and 2 which was significant at $p = 0.038$, and between regions 1 and 4 which was significant at $p < 0.001$. Significant values ranged from 0.054 (regions 1 and 2) to 0.342 (regions 1 and 4). For concatenated sequences, F_{ST} values were significant between regions 1 and 3 and between 2 and 4 at $p < 0.01$, and between region 1 and 4 at $p < 0.001$. Values were not significant between regions 1 and 2 ($p = 0.166$) and between regions 3 and 4 ($p = 0.082$). F_{ST} values ranged from 0.029 (regions 1 and 2) to 0.441 (regions 1 and 4; followed by $F_{ST} = 0.329$ for regions 2 and 4). For SNPs, F_{ST} values between regions 1 v. 2 ($F_{ST} = 0.185$), 1 vs 3 ($F_{ST} = 0.115$) and 1 vs 4 ($F_{ST} = 0.246$) were all significant ($p < 0.05$), whereas the other comparisons were not.

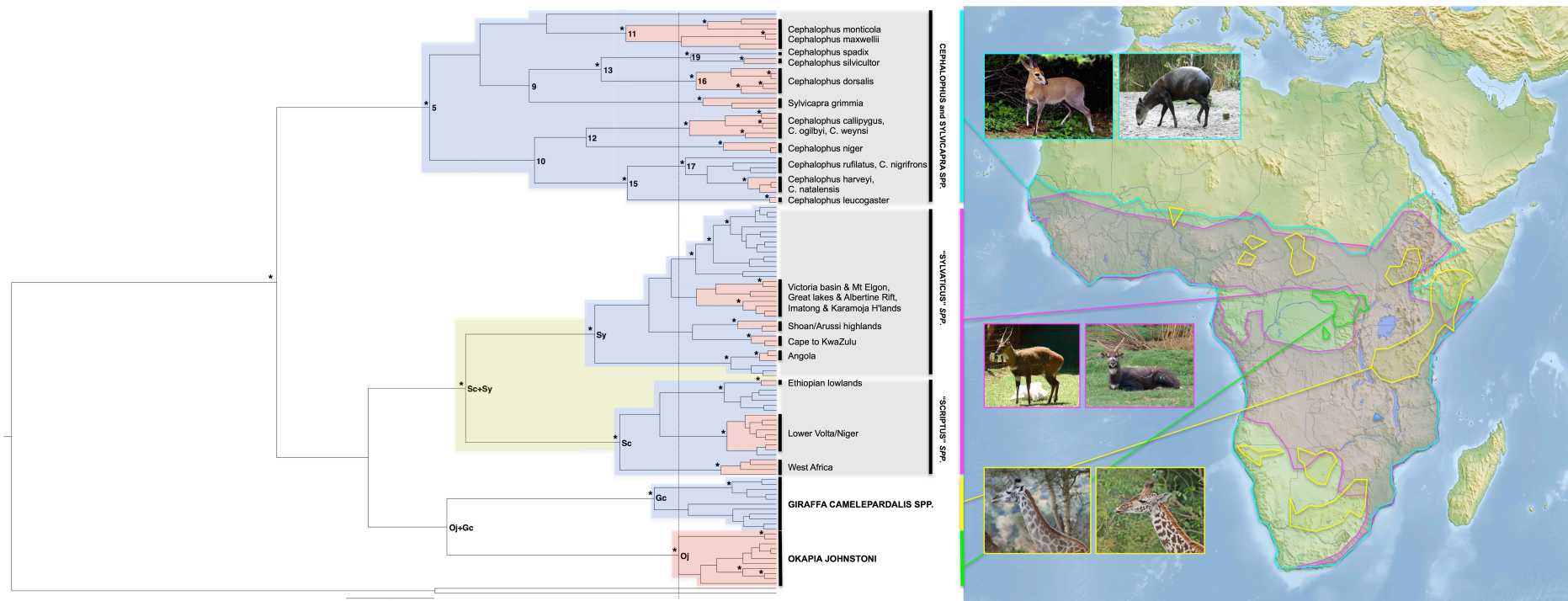
Table 4.2. 95% HPD intervals for dates of divergences (mya) for Figure 4.4 of the present study, and from the original studies (nodes 5-17, Johnston and Anthony [48]; nodes Sc/Sy, Moodley and Bruford [14]).

Node	Dates (Previous study; mya)	Dates (present study; mya)
5	6.27 - 11.43	10.47 - 20.91
9	4.16 - 7.78	6.60 - 16.44
10	3.58 - 6.69	6.50 - 16.20
11	2.68 - 5.31	3.71 - 10.93
12	2.52 - 4.97	4.53 - 13.67
13	2.53 - 4.93	4.38 - 12.52
15	2.13 - 4.27	3.16 - 11.75
16	1.74 - 3.54	1.51 - 6.62
17	1.18 - 2.38	1.87 - 7.30
19	0.80 - 1.91	1.35 - 7.28
Sc	2.0 - 3.0	4.01 - 10.97
Sy	2.0 - 3.0	4.65 - 12.65
Sc+Sy	3.9 - 6.5	9.46 - 19.01

4.4.4 Population and sequence divergences

To investigate phylogenetic relationships between haplotypes, BEAST (Drummond & Rambaut, 2007) was used to clarify phylogenetic relationships and to infer divergence times of the lineages. A phylogeny was constructed for okapi and giraffe, using 505 bp of homologous concatenated mitochondrial Cyt *b* sequence, tRNA-Pro and tRNA-Thr genes (Figure 4.2), using a lognormal relaxed tree, calibrated to the divergence between Giraffidae and Cervidae. The phylogeny identified several deep lineages within okapi, including one ancient divergence that divides okapi mtDNA into two monophyletic groups. BEAST analysis estimated the most ancestral okapi divergence as occurring at 1.7 – 12.8 (95% HPD; mean, 6.8; posterior probability of 0.96) mya. Six of the ten other okapi divergence events were also estimated at greater than one million years old (lower estimate was greater than one million years). The giraffe phylogeny showed divergence events of a similar magnitude, with the most ancestral divergence estimated at 2.0 – 12.6 (95% HPD; mean, 6.3; posterior probability of 0.99) mya.

Figure 4.3. Okapi (*Okapia johnstoni*), giraffe (*Giraffa camelopardalis*), bushbuck (*Tragelaphus scriptus* spp.) and duiker (Cephalophinae spp.) tree drawn in BEAST v1.7.5. Posterior probabilities of >0.8 are highlighted with an asterisk. Dotted line indicates the most ancestral divergence within okapi. The shading on the tree shows when taxonomic units can be monophyletically grouped, with the different colours corresponding to different levels of inclusiveness for these groupings. For example, for bushbuck, Victoria Basin & Mt Elgon, Great Lakes & Albertine Rift and Imatong & Karamoja Highlands ecoregions could be grouped monophyletically, and are shaded red. The next monophyletic taxonomic grouping are the “scriptus” species (shaded blue), and then all bushbuck (shaded yellow).



In order to further understand the okapi phylogeny, trees for okapi, giraffe, duiker and bushbuck, were also reconstructed jointly (Figure 4.3). This was done in an attempt to address some of the discrepancies that can be encountered when using dated phylogenies, such as faulty calibration points (Graur & Martin, 2004), rate heterogeneity among lineages (Bromham & Penny, 2003), and time dependent of rates of evolution (Ho & Larson, 2006; Ho et al., 2011). The comparative approach addresses these issues by simply providing relative divergence estimates using a single methodology, rather than trying to estimate absolute dates using different methodologies. This phylogeny (relaxed lognormal) gave estimates of TMRCA (Time to Most Recent Common Ancestor) for okapi of 2.0 - 7.9 mya and for giraffe of 2.7 - 9.3 mya. The topology of the section of the tree containing bushbuck and duiker species was broadly concordant with phylogenies of these species created in previous studies (Moodley and Bruford (2007); Johnston and Anthony (2012) respectively). The 95% confidence intervals of the divergence times of the duiker species in the phylogeny from the present study all overlap with the intervals in Johnston and Anthony (2012). However, the inferred dates of the coalescent events of the Cyt *b* lineages for bushbuck, and the *T. scriptus* and *T. sylvaticus* lineages in this study were considerably higher than Moodley and Bruford (2007) (Table 4.2). Based on this joint phylogeny, the divergence of the two most divergent okapi lineages predates the divergence of several major duiker species, including *C. jentinki* from *C. dorsalis* (Figure 4.3, node 16); *C. rufilatus* (node 17); *C. nigrifrons* from *C. harveyi* (node 17); *C. natalensis* (node 17); and *C. spadix* from *C. silvicultor* (node 19). These duiker lineages (nodes 16, 17 and 19) have previously been estimated to have diverged between 1.74 – 3.54, 1.18 - 2.38 to 0.80 – 1.91 mya respectively (Johnston & Anthony, 2012). The divergence of the okapi lineages also appears to pre-date the emergence of many of today's described bushbuck subspecies, for example *T. sylvaticus sylvaticus* from *T. sylvaticus meneliki* and *T. sylvaticus powelli*, and approximately twice as old as the emergence of both *T. scriptus decula* and *T. sylvaticus ornatus*. The TMRCA for the okapi is similar to that of all the giraffe subspecies.

Table 4.3. Values with the highest posterior probabilities for the parameters investigated for the popABC analysis, with comparisons between regions one, two and four (R1, R2 and R4). Parameters investigated are average mutation of the sequence (AvMutS), migration into each population (mig1 and mig2), effective population size of each population (Ne1 and Ne2), effective population size of the ancestral population (NeA1) and time of population splitting (t1).

Parameter	R1vR4	R1vR2	R2vR4
AvMutS	$5e^{-4}$	$<1e^{-5}$	$1.5e^{-2}$
mig1	0.2	4	1
mig2	0.2	2	0.1
Ne1	200	500	3000
Ne2	175	75	<50
NeA1	6000	4500	14000
t1	$2e^5$	$2e^5$	$2e^5$

PopABC (Lopes et al., 2009) was used to infer divergence times, migration rates, and present-day and historic effective population sizes of pairwise combinations of samples from sampling regions one, two and four, using both mitochondrial DNA and nuclear loci (Table 4.3; for posterior distributions see Figures S4.2-4.22; samples from sampling region three were excluded due to low sample number). Migration rate was inferred to be consistently lower when comparing populations northeast versus southwest of the Congo River compared to the same side, and three of the four inferred migration rates across the Congo were an order of magnitude lower than the two migration rates on the same side. In every instance ancestral effective population size (NeA1) was considerably higher than any of the inferred present-day effective population sizes (Ne1 and Ne2), implying a reduction in population size since these populations became separated. Time since divergence of all the populations was inferred at approximately 200 kya, and interestingly, was the same for all population comparisons. Inferred mutation rates, however, varied substantially among pairwise comparisons, as did the effective population size for region 2.

4.5 Discussion

4.5.1 Okapi genetic diversity and evolutionary history

Paleontological records of *Okapia spp.* are virtually non-existent, with no known fossils predating the Pleistocene, except *Okapia stillei* (Dietrich (1942) in Van der Made and Morales (2012)), which has since been reclassified as Giraffa (Harris et al., 2010). Giraffidae are first known from the late early Miocene in Africa, and by the Late Miocene giraffids were very widespread and diverse. During the Early Pliocene they became rare in Eurasia, but remained diverse in Africa (Gentry, 1990; Harris, 1991; Bonis et al., 1997; Harris et al., 2010). The late Miocene, early Pliocene is thought to be a period of gradual cooling and increasing climatic variability in Africa. Okapi and giraffe are thought to share a common ancestor approximately 16 mya (Fernandez & Vrba, 2005; Hassanin et al., 2012). Based on the okapi and giraffe phylogeny, the present study estimates the most ancient divergence within okapi mitochondrial lineages to be minimally 1.7 mya (with maximum sequence divergences of 7.10% and 3.49% for CR and Cyt *b* sequences respectively). This result implies that okapi mitochondrial DNA haplotype divergence dates to at least the early Pleistocene. Sequence divergences of this magnitude are more consistent with divergence dates detected between African species or subspecies (e.g. divergence of *Elephas* and *Loxodonta* elephant genera (Maglio, 1973); the *Phacochoerus africanus massaicus*, *P. a. sundevallii* and *P. a. africanus* warthog subspecies divergences (Muwanika et al., 2003); spotted hyena divergences (Rohland et al., 2005); and the *Scriptus* and *Sylvaticus* bushbuck species divergences (Moodley & Bruford, 2007)), yet there is no suggestion that okapis comprise more than one taxon. This estimate of intra-specific divergence time for okapi is also at the upper limit for what has previously been estimated for the emergence of the extant giraffe subspecies (0.54 – 1.62 mya (Brown et al., 2007)).

The present study constructed a phylogeny (Figure 4.3), and calculated genetic diversities (Table 4.1) in a comparative manner, that included multiple ungulate taxa. Based on the combined phylogeny, the divergence of the two most ancestral okapi mitochondrial lineages predates the divergence of several major duiker lineages, which have previously been estimated to have diverged between 0.80 –

3.54 mya. This gives further support to a divergence of at least 1.7 mya for the most ancestral okapi mtDNA lineage. The divergence of okapi is again estimated to be similar to that of all the giraffe subspecies, as well as the emergence of many of the bushbuck subspecies (e.g. *T. scriptus decula* and *T. s. ornatus*). This is a surprising result, particularly when one considers the morphological and geographic variation that is contained within these giraffe, duiker and bushbuck taxa (Brown et al., 2007; Moodley & Bruford, 2007; Johnston & Anthony, 2012). The results of the CR nucleotide diversity comparison showed similar results. Okapi nucleotide diversity was similar to the combined giraffe subspecies and African Cape buffalo, and higher than the eland antelope. This comparative methodology provides a much more useful and meaningful means of comparing interspecific genetic diversity than simply stating genetic diversities out of context. Table 4.1 shows okapi to be one of the more genetically diverse of the ungulate species investigated in this study, implying an evolutionary history that is rich (i.e. has been influenced by considerable climatic change) and diverse (i.e. long periods where there have been opportunities for genetic differentiation) evolutionary history.

4.5.2 Evolutionary biogeography of the Congo Basin

The most ancestral mitochondrial DNA divergence in okapi is dated at greater than 1.7 mya (Figure 4.2). The Congo River is a likely candidate for the cause of the split of the most ancestral mtDNA sequence lineages in okapi, however, it is not possible to prove this definitively due to the possibility of retention of ancestral polymorphism in populations either side of the Congo River. The mitochondrial CR network shows six distinct lineages (Figure 4.1), and divergences of several of the other major okapi mtDNA CR lineages from the BEAST phylogeny are also dated at greater than one million years ago (the divergence of both of the monophyletic clades is *at least* 0.8 mya, Figure 4.2). These dates may be explained by the Congo Basin fragmenting into refugia at various stages throughout the Pleistocene. This is consistent with a hypothesis of increases in African climate variability and aridity near 2.8 ma, 1.7 ma, and 1.0 mya (DeMenocal, 2004)). Okapi are however known to be highly selective folivores and currently occupy a disjunct distribution within the Congo Basin.

Refugia may therefore have provided isolated regions of suitable forest type, rather than simply comprising patches of forest separated by savannah. Cowling et al. (2008) simulated the paleovegetation of Central Africa and LGM simulations indicate that although tropical broadleaf forest may not have been severely displaced by expanding grassland in central Africa, the structure of the forests may have been very different from today (with forests characterized by lower leaf area indices, lower tree heights and lower carbon content).

The inferred Approximate Bayesian relative divergence times and migration rates between the okapi sampling regions were relatively consistent. As would be expected, migration rates of sampling regions one versus two (same side of river) were consistently higher than regions one versus four and two versus four (opposite sides of river). Interestingly however, divergence times between all population comparisons were the same (~200 kya). When taken together, these results imply that although populations on the same side of the Congo River maintained much higher gene-flow since the time of divergence, they nonetheless diverged around the same time to those on opposite sides of the river. This would suggest that initial population divergence between okapi populations either side of the Congo River was primarily linked to the same biogeographic process that separated those populations on the same side of the river. A possible explanation for this could again be forest fragmentation, linked to repeated glaciation events during the Pleistocene (DeMenocal, 1995; DeMenocal, 2004; Maslin & Christensen, 2007). The Congo River is thought to have existed in approximately the same formation for tens to hundreds of millions of years (Anka et al., 2009). This estimate of population divergence is considerably more recent than the estimates of sequence divergence discussed earlier (>1.7 mya). This implies greater than one population fragmentation event, again suggestive of repeated cooling events.

These results are also in accordance with the distribution of mtDNA haplotypes. Deep genetic divergences with the presence of these haplotypes on each side of the Congo River suggests relatively recent gene-flow since these haplotypes diverged. This has also been found in other phylogeography studies in central

Africa, although the majority of these studies are on savannah organisms. A dichotomy among genetic haplotypes between West and Central Africa and Southern and East Africa has been observed in African bovids (Arctander et al., 1999; Nersting & Arctander, 2001; van Hooft et al., 2002), African elephants (*Loxodonta africana*) (Eggert et al., 2002), cheetahs (*Acinonyx jubatus*) (Freeman et al., 2001), black rhinoceros (*Diceros bicornis*) (Brown & Houlden, 2000), roan (*Hippotragus equinus*) (Alpers et al., 2004) and eland (*Taurotragus oryx*) (Lorenzen et al., 2010) antelopes and giraffe (*Giraffa camelopardalis*) (Brown et al., 2007). Bonobos and chimpanzees provide a particularly interesting comparison to okapi, as their combined range spans the Congo River. Bonobos and chimpanzees are estimated to have diverged ~1 mya (Won & Hey, 2005; Caswell et al., 2008; Hey, 2010; Kawamoto et al., 2013), with chimpanzees restricted to the northeast side of the Congo River, whereas bonobos are restricted to the southwest side. The diversification of chimpanzee sub-species, and bonobo haplogroups are explained by fluctuations in climate during the Pleistocene and the associated changes in forest cover (Fischer et al., 2011; Kawamoto et al., 2013). Taken together, these studies suggest periods of Pleistocene forest expansion that genetically differentiated southern, eastern and western populations in large numbers of savannah taxa. This would imply that okapi have at some point in the past been able to either cross or go around the Congo River, allowing admixture between mitochondrial lineages, whereas chimpanzees and bonobos have not. Intuitively this may sound unlikely due to the Congo River's considerable size, and as it is likely to have existed roughly in the same formation for tens or even hundreds of millions of years (Anka et al., 2009). However, geomorphic mechanisms do exist that may make this possible. Neck cutoff and oxbow lake formation could theoretically allow populations of organisms to move to the opposite side of a river without actually crossing it. A second possible explanation, anastomosis, is a common mechanism where the path of a river is broken into islands with channels of much smaller size. This process could have led to each of the individual channels being surmountable when the entire width of the river is not.

4.5.3 Partitioning of present-day genetic diversity

We show that deeply divergent mitochondrial haplotypes are ubiquitous across the okapi's range. This suggests that historic biogeographic processes have shaped the structure of genetic diversity in this species, and these processes pre-date the present-day distribution of okapi. Nonetheless, present-day geography also contributes to the structuring of genetic diversity in okapi. AMOVA consistently showed a very high percentage of genetic differentiation between groups (up to 29.65% for concatenated sequences when comparing either side of the Congo River). There were high F_{ST} values for mitochondrial DNA sequence data between populations either side of the Congo River, particularly between populations 1 and 4, and high and significant F_{ST} values when comparing SNPs between population 1 and the other populations, in particular population 4. The consistently high genetic differentiation between population 4 versus all other populations based on nuclear loci, and higher mitochondrial genetic differentiation across the Congo River, compared to between sampling regions on the same side highlights the importance of the Congo River in structuring present-day genetic diversity in okapi. In comparison, F_{ST} values between populations 1 and 2 were either low, or non-significant for mitochondrial sequences. The level of present-day population genetic differentiation seen in okapi is within the range of what is seen among chimpanzee populations (Fischer et al., 2011). No known morphological or behavioral differences separate these populations, however they are regarded as separate sub-species (Braga, 1995; Uchida, 1996), again emphasizing the remarkable genetic diversity seen within okapi.

The findings presented here therefore add to the evidence that the Congo River appears to be the most important factor structuring contemporary genetic diversity of large mammals in the Congo Basin (Won & Hey, 2005; Caswell et al., 2008; Hey, 2010; Kawamoto et al., 2013), although this has clearly not always been the case. However, interestingly, the Lomami River (a tributary running parallel southwards with the upper stretches of the Congo River) is the feature that delineates the range of the okapi population on the southwest side of the Congo River. This river has recently been shown to limit the range of a recently

described primate, the “lesula” (*Circopithecus lomamiensis*, (Hart et al., 2012)) and has also recently been shown to be the only river to be a strong barrier to gene-flow in bonobos (Kawamoto et al., 2013). A future avenue for research could therefore involve a multi-taxon analysis of the combined role of the Congo and Lomami Rivers in structuring species and genetic diversity in this area.

4.6 Acknowledgements

We gratefully acknowledge the assistance of the *Institut Congolais pour la Conservation de la Nature* (ICCN), in particular Director J J Mapilanga for allowing the field collection and export of samples. We also thank all the museum curators that sent historic samples or allowed us to collect them. In particular, from the Chicago Field Museum: Bill Stanley and Lawrence Heaney (with the help of Keith Dobney, Aberdeen Univ. and Greger Larson, Durham Univ); from Copenhagen NHM: Hans J. Baagøe, Kristian Gregersen and Mogens Andersen; from Paris NHM: Joséphine Lesur-Gebremariam; from the Royal Museum for Central Africa, Tervuren, Belgium: Wim Wendelen (with the help of Floris van der Sman). We thank Gilman International Conservation, the Okapi Conservation Project, the Frankfurt Zoological Society, the Wildlife Conservation Society, the Lukuru Foundation/TL2 project, the Zoological Society of London (ZSL) and the considerable number of people who assisted in various ways with sample collection. In particular, we would like to thank John Fataki Bolingo, Bryna Griffin, Terese Hart, Chrysostome Kaghoma, Luaison, John Lukas, Kambale Magloire, Ephrem Mpaka, Stuart Nixon, Linda Penfold, Elise Queslin, Alex Quinn, Rosemary Ruf and Ashley Vosper. This project was jointly funded by the UK Natural Environment Research Council (NERC) and ZSL (ZSL as NERC CASE industry partner and via an Erasmus Darwin Barlow Expedition grant). We also acknowledge the financial support provided by the US Fish and Wildlife Service, the UK’s Darwin Initiative and the Mohamed bin Zayed Species Conservation Fund.

4.7 Supplementary

Table S4.1. EPIC primers designed and tested in this study. Observed and expected heterozygosities were generated using the same 28 individuals as the AMOVA and F-statistic analysis in the present study. Multiple SNPs occurring on a single sequence have been notated with a suffixed letter.

SNP Locus	Primers (F, forward; R, reverse)	H _o	H _e	Fragment size (bp)	Reference
ADH2 (Alcohol dehydrogenase)	F: AGG TGA TAA AGT CAT CCC G; R: TAC TCA CTC GTG TTT AAG GC	0.056	0.049	49	Lyons et al. 1997
ALDOB (Aldolase B, fructose-bisphosphate)	F: ACA GGT AGC CAA CAC TTC; R: TGA GGA GAG AAG GGA TTT AG	0.181	0.159	127	Venta et al. 1996
B2M _a (Beta-2 microglobulin)	F: ACC CTC TGA GAG AGC TG; R: GAA GCC TTG GGA GTG C	0.125	0.094	103	Lyons et al. 1997
B2M _b	(as above)	0.6	0.386	(as above)	(as above)
B2M _c	(as above)	0.746	0.454	(as above)	(as above)
B2M _d	(as above)	0.021	0.103	(as above)	(as above)
BGN (Biglycan)	F: GAC GGG GAA GAG AGA GG; R: GGT GTG ACC CTA GAC AAG	0.227	0.372	68	Lyons et al. 1997
CHRNA1 (Cholinergic receptor, nicotinic, alpha 1)	F: CAC GTT GGG ATT TAA ATT GTC; R: CTA GGG AGT CAT TAG GCA C	0.232	0.178	47	Lyons et al. 1997
CHY _a (chylous ascites)	F: TGC AGA CAC CTT TGG AAG; R: ACC ACA TGG GAC CGA G	0.1	0.08	83	Venta et al. 1996
CHY _b	(as above)	0.357	0.349	(as above)	(as above)
CKMM (Creatine kinase muscle)	F: GTT GGG AAA ACA TAG GCA C; R: GAA GCT GTC TCC TGC C	0.125	0.094	71	Venta et al. 1996
COL10A1 (Collagen, type X, alpha 1)	F: GGT ACT CCT ATC CCA TTC G; R: AGA GAA GTA ATA TAT CCC TGG AA	0.479	0.459	102	Venta et al. 1996

CYP19 _a (Cytochrome P450, family 19)	F: TGG AAA TAA AAT GAG AAG AAA TGA; R: ACC TCA ATT AAG AAC AAG GTG	0.263	0.267	118	Lyons et al. 1997
CYP19 _b	(as above)	0.256	0.216	(as above)	(as above)
GH1 (Growth hormone)	F: GGA TGG TGT TGT GGG TC; R: AAG AGA GAG CAG GGT TCT	0.167	0.236	58	Lyons et al. 1997
GLB1 (Galactosidase, beta 1)	F: GAA TTC TAT ACT GGC TGG CTG; R: CCC TCC AAA AAT TAA CAC TCA	0.372	0.23	90	Venta et al. 1996
GLUT2 (Glucose transporter 2)	F: TGG GAC TCT GTT TAT TCT CTG; R: TCC TTT GGT TTC TGG AAC TTT A	0.267	0.228	62	Venta et al. 1996

Table S4.2. Nucleotide and haplotype diversities, and number of polymorphic sites for the mtDNA genes used in the present study

Locus	Mean nucleotide diversity	Mean haplotype diversity	Polymorphic sites
CR	0.035 (SD 0.018)	0.932	54
Cyt <i>b</i>	0.011 (SD 0.006)	0.798	18
tRNA-Pro	0.002 (SD 0.004)	0.161	1
tRNA-Thr	0.020 (SD 0.014)	0.831	6

Table S4.3. Study names and genbank IDs of sequences used in the comparative analysis of CR nucleotide diversity. CR section refers to the DNA fragment available for use from genbank, with numbers referring to the position of the fragment, relative to the start of the CR (0) based on the genbank annotations.

Study	Taxon	CR section	Number of genbank sequences	Genbank accession no.
Moodley and Bruford (2007)	Bushbuck (<i>Tragelaphus scriptus</i> spp.)	0 - 452	485	EF138117-EF138601
Alpers et al (2004)	Roan antelope (<i>Hippotragus</i>	0-444	103	AF049373–AF049375, AF049378, AF068839,

	equinus spp.)			AF068840, EF138117-EF138601, AY575568-AY575605
Arctander et al (1999)	Hartebeest (<i>Alcelaphus buselaphus</i> spp.); Bontebok (<i>Damaliscus pygargus</i>)	0-317	112; 3	AF167720-AF167978; AF176682-4
Nersting and Arctander (2001)	Greater kudu (<i>Tragelaphus strepsiceros strepsiceros</i>)	0-396	89	AF301621.1-AF301712.1
VanHooft et al (2002)	African cape buffalo (<i>Syncerus caffer caffer</i>)	31-373	195	AF313151-AF313345
Brown et al (2007) and Hassannin et al (2007)	Griaaffe (<i>Giraffa camelopardalis</i> spp.)	0-428	47	EF442263-EF442274; EU088317-EU088351
Arctander et al (2010)	Common eland (<i>Taurotragus oryx</i>)	0-395	122	GQ388117-GQ388238

Table S4.4. Study names and genbank IDs of sequences used for the 274 bp phylogeny of bushbuck, duiker, giraffe and okapi, including pig and collared peccary outgroups. Original* and final* haplotypes refer to number of haplotypes in the study in which those haplotypes were originally sequenced, and haplotypes based on 274 bp sequences in the present study respectively.

Taxon	Study	No. individuals / original* haplotypes	No. final* haplotypes	Genbank ID
Bushbuck spp.	Moodley and Bruford (2007)	161/90	55	EF137956.1 - EF138116.1
Collared peccary	Hassanin et al. (2012)	1/1	1	JN632682.1
Duiker spp.	Ntie et al. (2010)	92/54	39	FJ807571.1 - FJ807579.1; FJ807584.1 - FJ807593.1; FJ807595.1 - FJ807598.1; FJ807601.1 - FJ807603.1;

				FJ807605.1 - FJ807628.1; FJ888512.1; FJ959386.1 - FJ959388.1
Giraffe	Brown et al. (2007)	266/35	11	EU088317.1 - EU088351.1
	Hassanin et al. (2007)	23/12		EF442263.1 - EF442274.1
Okapi	Present	60/13	11	TBC
Pig	Hassanin et al. (2012)	1/1	1	FJ236991.1

Table S4.5. Table of prior values for popabc analyses. All priors (Ne, effective population size; NeA, ancestral effective population size; t, divergence time; mig, migration rate) used a uniform distribution, except mutAvS (average sequence mutation rate), which PopABC only gives the option of a normal or lognormal distribution. Priors were determined by carrying out preliminary runs, and altering the prior value until all posterior distributions were distinct from the prior distributions.

Prior	Region 1v2	Region 2v4	Region 1v4
Topology	Uniform	Uniform	Uniform
Ne1	10,000	5,000	5,000
Ne2	10,000	5,000	5,000
NeA	30,000	15,000	20,000
t1	1,000,000	1,000,000	1,000,000
mig1	15	1	2
mig2	15	1	1
mutAvS	Normal distribution (mean of mean, 0.01; stdev of mean, 0.01; mean of Sdev, 0.001; stdev of stdev, 0.0)	Normal distribution (mean of mean, 0.01; stdev of mean, 0.01; mean of Sdev, 0.001; stdev of stdev, 0.0)	Normal distribution (mean of mean, 0.01; stdev of mean, 0.01; mean of Sdev, 0.001; stdev of stdev, 0.0)

Figure S4.1-4.7. Prior (black) and posterior (blue) distribution for PopABC (Lopes et al., 2009) analysis of region 1 verses region 2 (Figure 4.1). Parameters investigated are mutation rate (mut rate), migration into populations one and two (mig1 and mig2; regions 1 and 2 respectively), effective population size of populations one and two (Ne1 and Ne2; regions 1 and 2 respectively), effective population size of the ancestral population (NeA) and time since divergence of populations one and two (t1; regions 1 and 2).

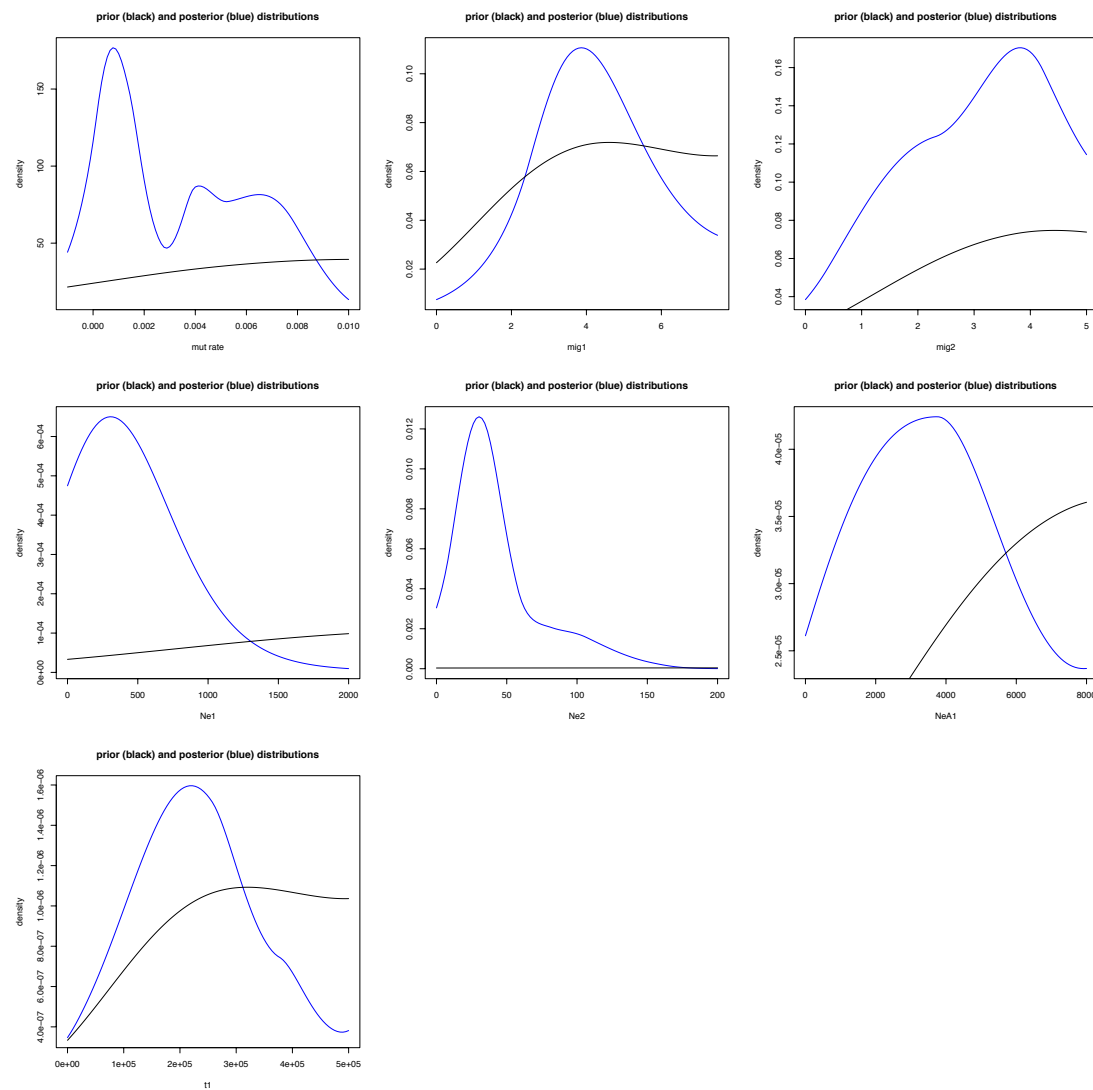
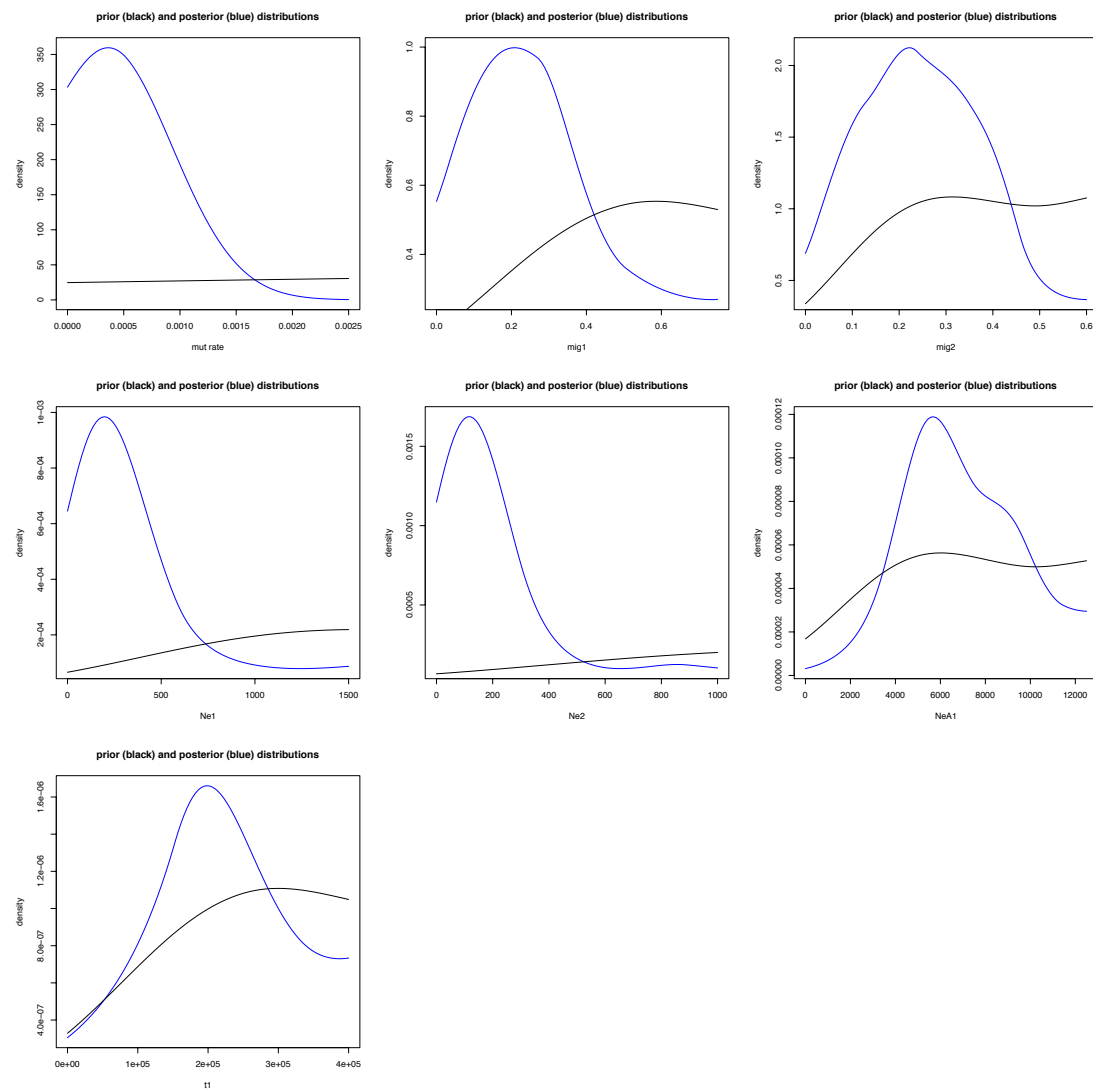
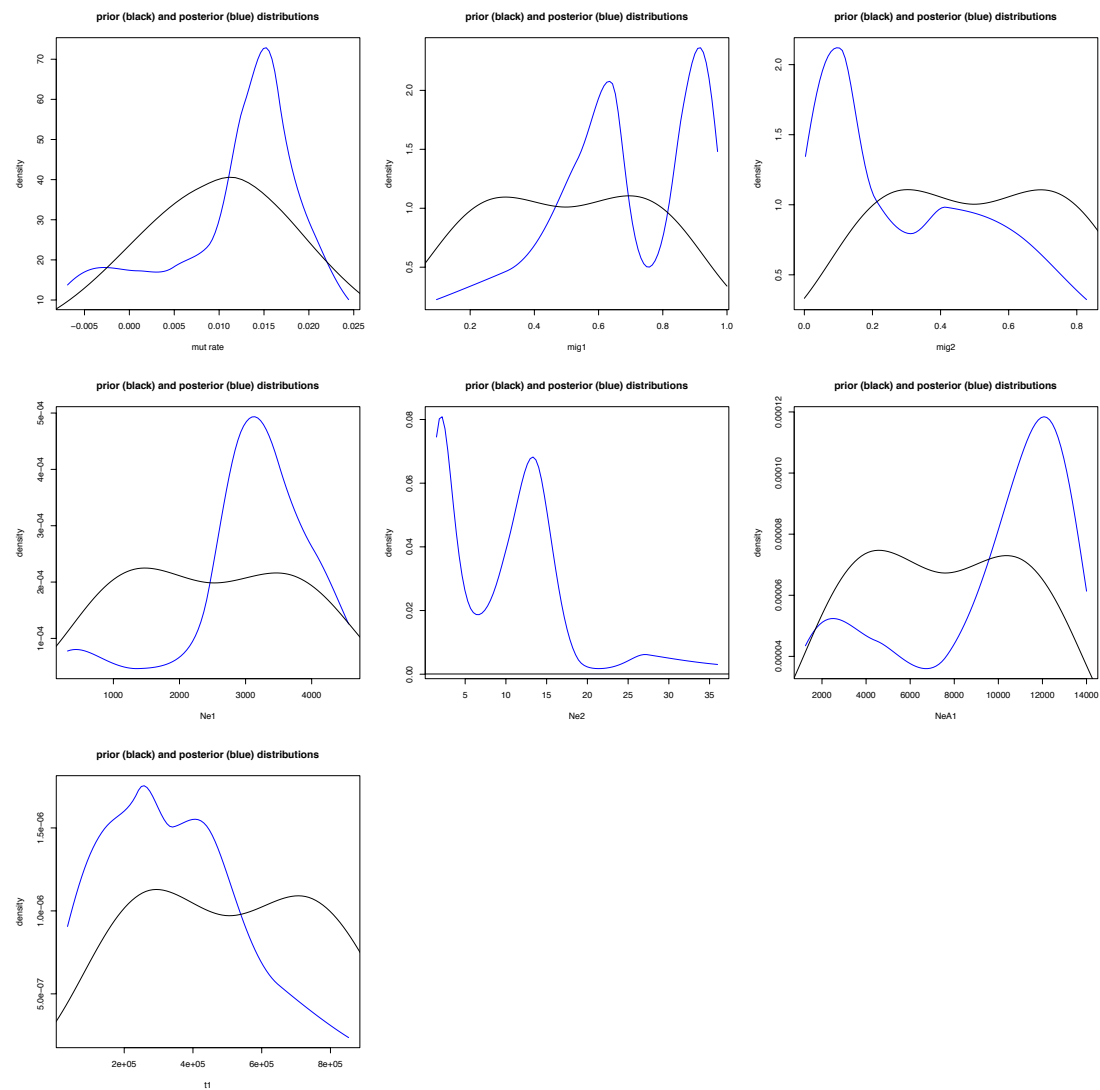


Figure S4.8-4.14. Prior (black) and posterior (blue) distribution for PopABC analysis of region 1 verses region 4. Parameters investigated are mutation rate (mut rate), migration into populations one and two (mig1 and mig2; regions 1 and 4 respectively), effective population size of populations one and two (Ne1 and Ne2; regions 1 and 4 respectively), effective population size of the ancestral population (NeA) and time since divergence of populations one and two (t1; regions 1 and 4).

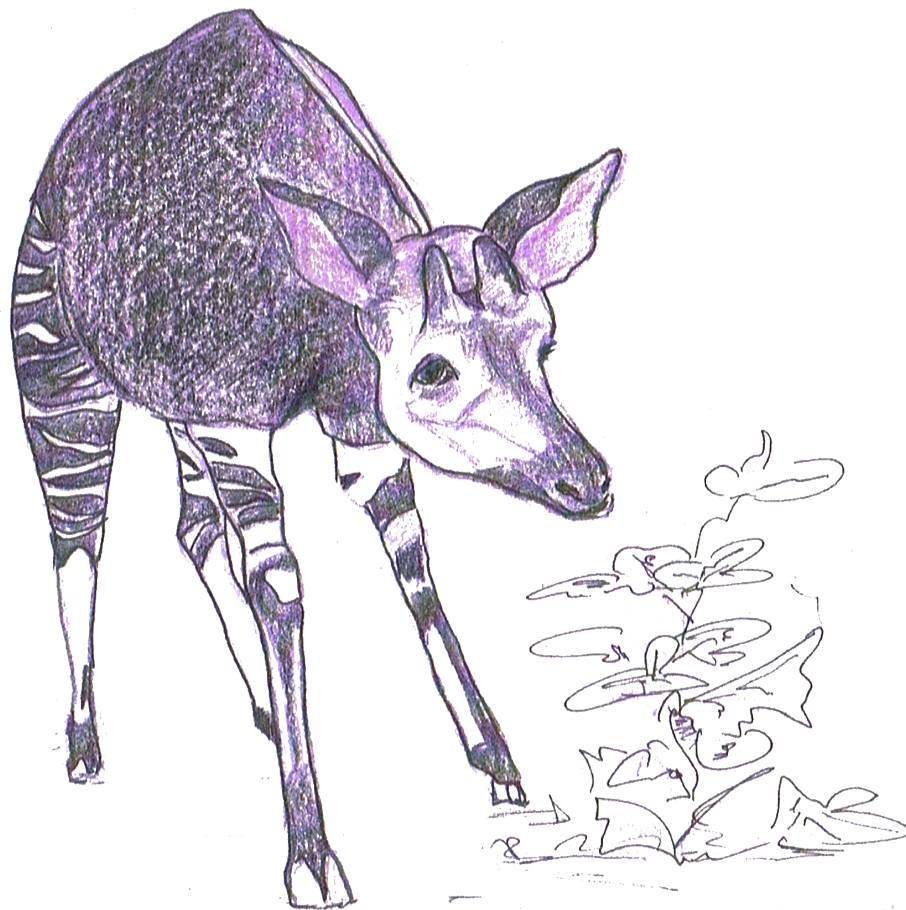


S15-21. Prior (black) and posterior (blue) distribution for PopABC analysis of region 2 versus region 4. Parameters investigated are mutation rate (mut rate), migration into populations one and two (mig1 and mig2; regions 2 and 4 respectively), effective population size of populations one and two (Ne1 and Ne2; regions 2 and 4 respectively), effective population size of the ancestral population (NeA) and time since divergence of populations one and two (t1; regions 2 and 4 respectively).



5

CHAPTER FIVE – Okapi ecology



"I really wonder what gives us the right to wreck this poor planet of ours."

Kurt Vonnegut

CHAPTER FIVE - Enhancing knowledge of the ecology of a highly elusive species, the okapi, using non-invasive genetic techniques

David W. G. Stanton^{1*}, John Hart², Noëlle F. Kümpel³, Ashley Vosper⁴, Stuart Nixon³, Michael W. Bruford¹, John G. Ewen^{5,§}, Jinliang Wang^{5,§}

Submitted to: Animal Conservation (26th March 2014)

§Joint senior authors

¹*School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom*

²*Lukuru Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), 1235 Ave Poids Lourds, Kinshasa, DRC*

³*Conservation Programmes, Zoological Society of London, London NW1 4RY, United Kingdom*

⁴*Wildlife Conservation Society (WCS), Great ape and human impact monitoring, Okapi Faunal Reserve, DRC*

⁵*Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom*

* Lead author contribution: All labwork and all data analysis except for the program COLONY, which was run by Jinliang Wang. Preparation of first manuscript draft, and modifications of manuscript based on co-author comments.

5.1 Abstract

Okapi (*Okapia johnstoni*) are an even-toed ungulate in the family Giraffidae, are endemic to the Democratic Republic of Congo (DRC). Very little is known about okapi ecology in the wild. We used non-invasive genetic methods to examine the social structure, mating system and dispersal for a population of okapi in the Réserve de Faune à Okapis, DRC. Okapi individuals appear to be solitary, although there was some evidence of genetically similar individuals being associated at a very small spatial scale. There was no evidence for any close spatial association between groups of related or unrelated okapi for either sex, but we did find evidence for male-biased dispersal. Okapi are genetically polygamous or promiscuous, and are also likely to be socially polygamous or promiscuous. An isolation by distance pattern of genetic similarity was present, but appears to be operating at just below the spatial scale of the area investigated in the present study. We therefore here provide new ecological information about a species that has recently been recognised by the IUCN as Endangered, and is a potentially important flagship species for Central Africa.

5.2 Introduction

The key to protecting and managing species of conservation concern is a good understanding of their ecology, including knowledge of their dispersal, sociality and mating system (Schaller, 1993; Primack, 2000). This information can have a considerable and very real impact on conservation (e.g. Veit et al., 2005; Martins et al., 2006; Archie & Chiyo, 2012).). However, measuring these factors in wild animals by direct observation can often be very difficult, especially for elusive mammals, or those inhabiting difficult terrain (Hájková et al., 2008; Anuradha Reddy et al., 2012).

There is a vast amount of variation in social structure, mating systems and dispersal strategy amongst mammals, even among those that are taxonomically and geographically similar (Hewison & Gaillard, 1999; Klaus-Hügi et al., 2000; Starin, 2001; Magliocca et al., 2002; Cameron, 2004; Sheldon, 2004; Bowland & Perrin, 2009). This variation makes predictions of ecological and socio-genetic processes difficult for any poorly studied mammal species. In terms of social structure, mammals that utilise densely forested habitats tend towards forming a smaller social unit, putatively because the coordination of a social group is difficult in a forest especially if the animal is large (Eisenberg & Lockhart, 1972). Also, animals at greater risk of predation are more likely to adopt a hiding strategy (Eisenberg & Lockhart, 1972) and be predominantly solitary to reduce social interaction and therefore detection probability (Geist, 1974).

Mating systems are even more diverse (Clutton-Brock, 1989) and difficult to predict. For example, the extent of polygamy can be affected by predation pressure (Kelly et al., 1999), social group composition (Webster et al., 2004) and phylogeny (Emlen & Oring, 1977). Due to this complex interaction, mammals show a diverse array of mating systems (Clutton-Brock, 1989). Dispersal (specifically natal dispersal (Howard, 1960)) also often varies between sexes, with some degree of sex-biased dispersal, usually male (Greenwood, 1980), being virtually ubiquitous in mammals (Lawson Handley & Perrin, 2007). Due to this lack of predictive power of habitat and taxonomy, other methods are clearly

needed to accurately elucidate the ecology of elusive, or otherwise difficult to observe animals.

Non-invasive genetic methods are increasingly being used to investigate questions such as dispersal, mating systems and social structure in wild animals (Taberlet et al., 1997; Goossens et al., 2005; Zhan et al., 2007). These methods may therefore provide a means of investigating the ecology of elusive animals without actually observing them. The okapi is a highly elusive even-toed ungulate, endemic to the Democratic Republic of Congo (DRC). Although widely distributed throughout the DRC, it occurs at low density across its range (Nixon & Lusenge, 2008). Okapi appear to only be present in dense forest, away from human presence (Hart pers. comm., Bodmer & Rabb, 1992). Determining aspects of behavioural ecology using observations is therefore difficult for this species. Only two *in situ* ecological studies of okapi have been published (Bodmer & Gubista, 1988; Hart & Hart, 1989). However the studies are somewhat equivocal, are lacking in detail, and tell us nothing of okapi mating systems or dispersal. Non-invasive genetic methods therefore potentially provide a useful tool for the study of this species.

We hypothesised that okapi are mostly solitary, due to their utilisation of dense rainforest, and the likelihood of them having a high predation pressure (Hart & Hart, 1989). In captivity, okapi males are rotated among females and sire multiple offspring (Galbusera, 2013). We hypothesised that this would also be true in the wild, with okapi showing evidence of genetic polygamy, or promiscuity. We also hypothesised that okapi would demonstrate male-biased dispersal, due to its higher incidence in mammals. The above hypotheses will be tested using dung samples from okapi in a population in the okapi faunal reserve (Réserve de Faune à Okapis, RFO), DRC.

5.3 Methods

5.3.1 Study species and site

Okapi are an even-toed ungulate in the family Giraffidae, separated from the giraffe by an estimated ~16 million years of independent evolution (Hassanin et

al., 2012). The limited number of long-term ecological studies that have been carried out on okapi have been based in the RFO (Bodmer & Gubista, 1988; Hart & Hart, 1989) and this reserve was also chosen for the present study (Figure 5.1). Four teams sampled the park, between December 2010 and February 2011, and collected 208 putative okapi fecal samples. These samples were collected as part of a great ape and human monitoring survey (Vosper et al., 2012). Surveys comprised a total of 164X one km variable width line transects. Transect sampling locations were determined randomly using DISTANCE 6.0 (Thomas et al., 2010), and fecal samples were collected while walking transect and while travelling between sampling locations (Vosper et al., 2012). All dung samples were geo-referenced using handheld GPS units. Each transect was walked once.

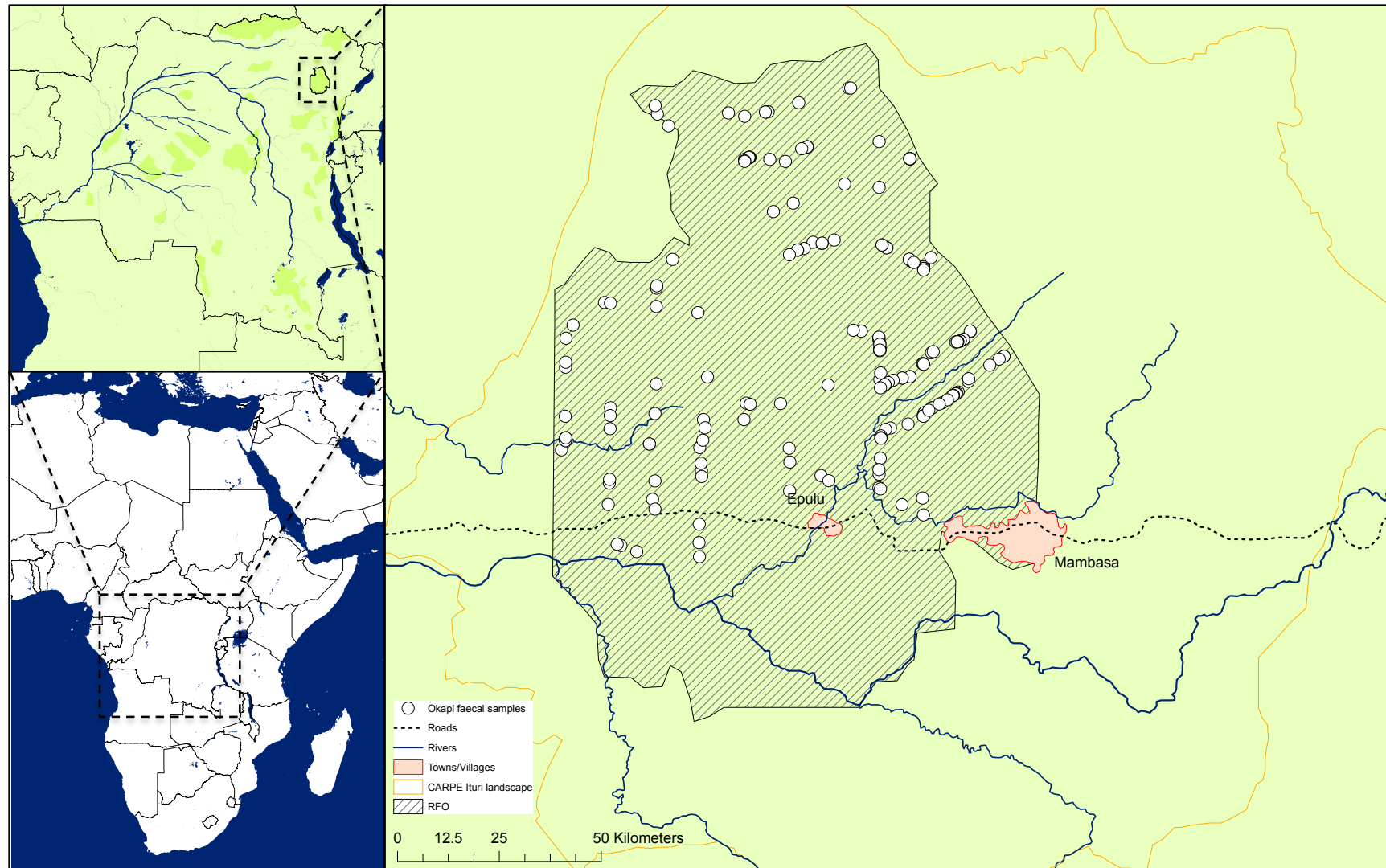
5.3.2 DNA extraction and amplification

DNA was extracted from faecal samples (stored in 100% ethanol for 24 hrs and then silica) using a QIAamp DNA Stool Mini Kit (Qiagen). Thirteen microsatellite loci were amplified using the primers Oka-01–13 and PCR conditions from Stanton et al. (2010). Primers Oka-02, 10 & 11 were excluded from the analysis due to low PCR amplification success rate. From the 208 faecal samples, consensus genotypes were generated for 105. These 105 samples were confirmed to be okapi based on the following: 1) Correct species identification from this survey was 100% based on mitochondrial DNA analysis of a subset of samples (Stanton et al (submitted)). 2) Genetic structure and distance analysis of microsatellite data in the present study did not identify any unusually different genotypes within the 105 genotyped samples.

The primer sequences SRY 1 (5' CTCATTGTGTGGTCTCGTG 3') and SRY 2 (5' CGGGTATTTGTCTCGGTGTA 3'; Wilson and White [1998]) were used to amplify a fragment in 5 blood samples from captive male okapi. Internal primers OJSEX-F (5' CGTGAACGAAGACGAAAG 3') and OJSEX-R (5' TCAATATCTGTAAGCCTTTTCC 3') were designed to amplify a shorter 101 bp fragment in non-invasive okapi samples. Sexing primers were multiplexed with an internal control, Oka-01 (forward: 5' AAGAGAGACTGCACTGTGGACC 3', reverse: 5' GCTCTTGTGTCTGACATGTTCTC 3', (Stanton et al., 2010)). PCR was carried out in

a 6.5 µl volume with 2.5 µl Multiplex Mix (Qiagen), 4 µg BSA, 2 nmol OJSEX primer, 0.8 nmol Oka-01 and 2 µl DNA. The PCR was carried out twice for each of the samples that had been successfully genotyped, always with two negative controls. A sample was accepted as a female if both reactions showed the absence of a band from the sexing primers.

Figure 5.1. Map showing study area, the Okapi Faunal Reserve (Réserve de Faune à Okapis, RFO) in the DRC. All 105 samples with genotypes used in the present study are shown as white open circles



Primers Mt 1 – 5 (Stanton et al. (submitted)) were used to amplify a fragment of the mitochondrial DNA control region (mtDNA CR), and cytochrome b, tRNA-Thr and tRNA-Pro genes in individuals with sexing information, using the conditions from (Stanton et al. (submitted)). A 325 bp fragment was amplified in 20 individuals (females $n = 9$, males $n = 11$), and a 543 bp fragment (that included the 325 bp fragment above) was amplified in a further 15 individuals.

5.3.3 Data validation

A preliminary genotyping error rate study was carried out using the programs PEDANT (Johnson & Haydon, 2007) and GEMINI (Valière et al., 2002) on 14 okapi faecal samples, comparing two genotyping repeats of each sample. GEMINI indicated that 2-3 repeats would be required to be able to accept a consensus genotype with >95% confidence, and PEDANT calculated an allelic dropout rate for each locus at between 0.0170 and 0.1645 (mean 0.0779), a false allele rate of between 0 and 0.0718 (mean 0.0170). The confidence converged on 100% with approximately three repeats. Therefore, for caution, at least four repeats (and up to eight) for each of the samples in the full study were carried out. Genotyping error rates were then recalculated on the full dataset. The allelic dropout rate for each locus was between 0 and 0.0429 (mean 0.0161) and false allele rate was between 0 and 0.0055 (mean 0.0010), demonstrating that the four repeats carried out were sufficient to give reliable consensus genotypes at the 95% level.

5.3.4 Spatial autocorrelation

To test the hypothesis of low social structure in okapi, the relationship between proximity of okapi dung samples and genetic distance was investigated. This was to determine if related individuals are spatially more closely associated than unrelated individuals, and was carried out using spatial autocorrelation analysis (SAA). Spatial autocorrelation measures the degree of dependency of observations, for example genetic distance, across space. Significant positive autocorrelation (in the example of genetic distance) indicates that genetically similar individuals are closer together than one would expect by chance, whereas significant negative autocorrelation indicates that individuals are arranged to maximize genetic distance between them (Manel et al., 2003). We

carried out SAA using GenAlEx v6.4 (Peakall & Smouse, 2006; Peakall & Smouse, 2012), with significance assessed using 95% confidence interval and 9999 permutations. SAA was carried out on males ($n = 27$) and females ($n = 29$) separately, and on the combined dataset ($n = 83$) at distance intervals of (i) 2 km across 20 km, and (ii) 10 km across 120 km. The analysis was carried out on the combined dataset only ($n = 83$; there was insufficient data to analyse males and females separately) at distance intervals of 0.2 km across 2 km.

5.3.5 Patterns of relatedness

To further describe sociality of okapi in the study site, and to complement the SAA, the association between spatial proximity and genetic relatedness was investigated. Pairwise relatedness was estimated using the program COANCESTRY v1.0.1.2 (Wang, 2011), which implements seven methods for estimating pairwise relatedness from individual multilocus genotypes. Duplicate genotypes were removed from the dataset and the spatial proximity of related dyads in the remaining individuals ($n = 83$) was described. This was done by investigating if there were significant differences between average spatial proximity of dyads with a relatedness greater than 0.5 verses less than 0.5, and greater than 0.25 verses less than 0.25, using t-test tests in R (R Development Core Team). This was carried out for all seven estimators. A rarefaction analysis was also carried out on the microsatellite genotypes using the program RERAT (Schwacke et al., 2005) to investigate the ability of the 10 markers used in the present study for inferring relatedness.

5.3.6 Multiple dung piles

Eight multiple dung piles (greater than one dung pile ≤ 2 m apart) were found in the study site. Duplicate genotypes were identified, and genetic relatedness was described for these samples, to investigate if these multiple dung piles represent social groups, or single individuals. Multilocus genotypes different at most at only one locus (to account for genotyping errors) were regarded as from a single individual.

5.3.7 Mating system

To investigate the mating system of okapi the relative numbers of half versus full siblings were estimated using the program COLONY (Jones & Wang, 2010). COLONY considers the two-generation full-pedigree of all sampled individuals, and assigns sibship and parentage jointly. As the method implemented in this study is effectively using offspring genotypes at autosomal loci, it is unable to determine the polygamous sex. When few half siblings are detected in the COLONY analysis, the mating system is inferred as monogamous for both sexes. Otherwise, it is inferred that either males, females, or both are polygamous. No prior was used for average sibship size, and the defaults for other parameters were accepted in the analysis.

5.3.8 Duplicate genotypes

A direct measure of movement was estimated using identical genotypes, identified in the dataset as dyads with zero or one allele different. Distance between identical dyads was measured, and classified as a natal dispersal event if the distance was greater than the current maximum recorded okapi home-range size (females: 5.1 km², males: 10.5 km²; (Hart & Hart, 1989)). All identical dyads less than this distance were classed as 'movement' events.

5.3.9 Spatial genetic structuring

To detect any hidden genetic structure and barriers to okapi movement/dispersal in the reserve, we carried out a Bayesian clustering analysis, and tested for isolation by distance and spatial autocorrelation. Bayesian clustering analysis was performed using the program STRUCTURE 2.3.4 (Pritchard et al., 2000), with 500,000 MCMC iterations, a burn-in of 50,000, correlated allele frequencies and K set at 1-5. Isolation by distance analysis was carried out in R (R Development Core Team) using a mantel test to assess the correlation between geographic distance and genetic distance, calculated using GenAlEx (Peakall & Smouse, 2006; Peakall & Smouse, 2012). Spatial autocorrelation analysis was also carried out in GenAlEx, using the methods described above.

5.3.10 Sex-biased dispersal

Sex-biased dispersal can be detected by a differences in mitochondrial haplotype diversity (Eriksson et al., 2006; Apio et al., 2010), $mAIC$ and $vAIC$ (Goudet et al., 2002; Lawson Handley & Perrin, 2007), F_{ST} values (Wright, 1931), relatedness estimates (Queller & Goodnight, 1989) and genetic structure (Lu et al., 2001; Zhan et al., 2007) between males and females. In all sex-biased dispersal analyses, only individuals that had been assigned as either male ($n = 27$) or female ($n = 29$), after duplicate genotypes had been removed, were used. Populations for the F_{ST} analysis were the northern half of the RFO verses the southern half, and the western half of the RFO verses the eastern half, with F_{ST} calculated separately for males and females. Pairwise relatedness (Queller and Goodnight method (Queller & Goodnight, 1989)) was calculated for all individuals in the dataset described above ($n = 56$), using GenAlEx (Peakall & Smouse, 2006; Peakall & Smouse, 2012). Significant differences were then tested between males and females in R (R Development Core Team) using a t-test. Normality was confirmed visually using histograms and qq plots. Haplotype diversity was calculated in i) all 35 individuals for the 325 bp fragment, and ii) the 15 individuals for which 543 bp of sequence data was available for, using DNAsp v5 (Librado & Rozas, 2009). Bayesian clustering analysis was performed using the program STRUCTURE 2.3.4 (Pritchard et al., 2000), and the settings described above, separately for males and females to investigate if any differences in dispersal can be detected in differences in genetic structure. FSTAT v2.9.3.2 (Goudet et al., 2002) was used to investigate if there were differences in $vAIC$ and $mAIC$ for males and females in the dataset. A one-sided test was run with 10,000 permutations. Assumptions of the program are that dispersal occurs at the juvenile stage, before reproduction, and that individuals are sampled post-dispersal. This first assumption is reasonable, however it cannot be determined if our dataset contained pre-dispersal individuals. The power of these statistical descriptors may therefore be lower than expected.

5.4 Results

5.4.1 Spatial autocorrelation

Using the 2 - 20 km distance category, we found consistent positive autocorrelations at 4 km ($p < 0.05$) for males, females and the combined dataset. There was also negative autocorrelations in males and females at 14 km and 18 km respectively. When considering the 10 - 120 km distance category: There was a negative autocorrelation at 20 km ($p < 0.05$), 110 km ($p < 0.05$) 80 km ($p < 0.01$) in the female, male and combined datasets respectively. Unexpectedly there was also a *positive* autocorrelation at 50 km ($p < 0.05$) for the male dataset. When considering the 0 - 2 km distance category: There was a positive autocorrelation at 0.2 km ($p < 0.01$) and 1 km ($p < 0.05$), and a negative autocorrelation 0.6 km ($p < 0.05$). When considering the 2 - 20 km distance category: There was a positive autocorrelation at 4 km for males, females and the combined dataset ($p < 0.05$ in all cases). There was also a negative autocorrelation at 14 km for males, and 18 km for females ($p < 0.05$ in both cases). Spatial autocorrelation graphs are shown in Figures 5.2-5.4 (males and females combined), Figures 5.5 & 5.6 (males only) and Figures 5.7 & 5.8 (females only).

5.4.2 Patterns of relatedness

For all seven estimators using COANCESTRY, geographic distance was lower for dyads with a relatedness value greater than 0.5. This difference was significant using some estimators, but not others (LREst: 45.5 km vs 50.2 km, $t = 0.816$, $p = 0.425$; TrioEst: 44.7 km vs 48.8 km, $t = 1.165$, $p = 0.250$; WEst: 39.3 km vs 48.7 km, $t = 1.865$, $p = 0.826$; REst: 47.5 vs 48.7, $t = 0.138$, $p = 0.893$; MEst: 42.5 km vs 48.9 km, $t = 2.126$, $p = 0.037$; LLEst: 38.8 km vs 48.8 km, $t = 2.236$, $p = 0.038$; QGEst: 40.4 km vs 48.7 km, $t = 1.17$, $p = 0.264$). There was no significant difference in average geographic distance between dyads with an estimated relatedness greater than 0.25, compared to those with an estimated relatedness value less than 0.25 for any of the estimators. A rarefaction analysis using RERAT described the ability of the 10 microsatellite markers used in the present study for accurately estimating relatedness (Figure 5.9). This analysis showed that change in relatedness had decreased to 0.038 using all 10 markers. A trend line,

based on a power relationship (change in relatedness = $0.272 * n_{\text{loci}}^{-0.858}$; $R^2 = 0.999$) indicated that increasing the number of loci to 20 would only decrease change in relatedness to 0.021, and increasing the number of loci to 100 would decrease change in relatedness to 0.005 (assuming loci had a similar level of polymorphism to the 10 loci used in this study).

5.4.3 Multiple dung piles

Of eight multiple dung piles, six contained only a single identical genotype. Of the two that were different, COLONY identified one of the dyads to be a first order relative (although couldn't distinguish between sibling or parent-offspring), and the other dyad to be a half-sibling.

5.4.4 Mating system

Mating system was investigated using the program COLONY to estimate relative numbers of half and full-siblings. Number of full siblings was one ($p = 0.999$) and number of half-siblings was 207 and 175 for posterior probability likelihoods of greater than 0.95 and greater than 0.80, respectively. This is highly indicative of a species that exhibits polygamy and or promiscuity.

5.4.5 Duplicate genotypes

All but one pairwise distance between identical genotypes was less than 1 km. The dyad that was greater than 1 km constituted two dung piles 25.5 km apart. Average distance between identical genotypes was 0.655 km (pairwise $n = 36$), or 0.103 km excluding the pair 25.5 km apart (pairwise $n = 35$). When classifying multiple dung piles as a single genotype, average distance between identical genotypes was 2.271 km (pairwise $n = 13$), or 0.337 km excluding the pair 25.5 km apart (pairwise $n = 12$).

5.4.6 Spatial genetic structure

STRUCTURE 2.3.4 (Pritchard et al., 2000) was unable to assign individuals to more than one population (data not shown). In addition, a mantel test was unable to detect any isolation by distance in the study area ($p = 0.462$, $r^2 = 0.000979$, scatterplot shown in Figure 5.10). These results show that the

sampling area effectively constitutes a single random mating population without apparent subdivision.

Figure 5.2. Spatial autocorrelation for all individuals (not including identical genotypes, $n = 84$) at 0.2 km size classes between 0 – 2 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

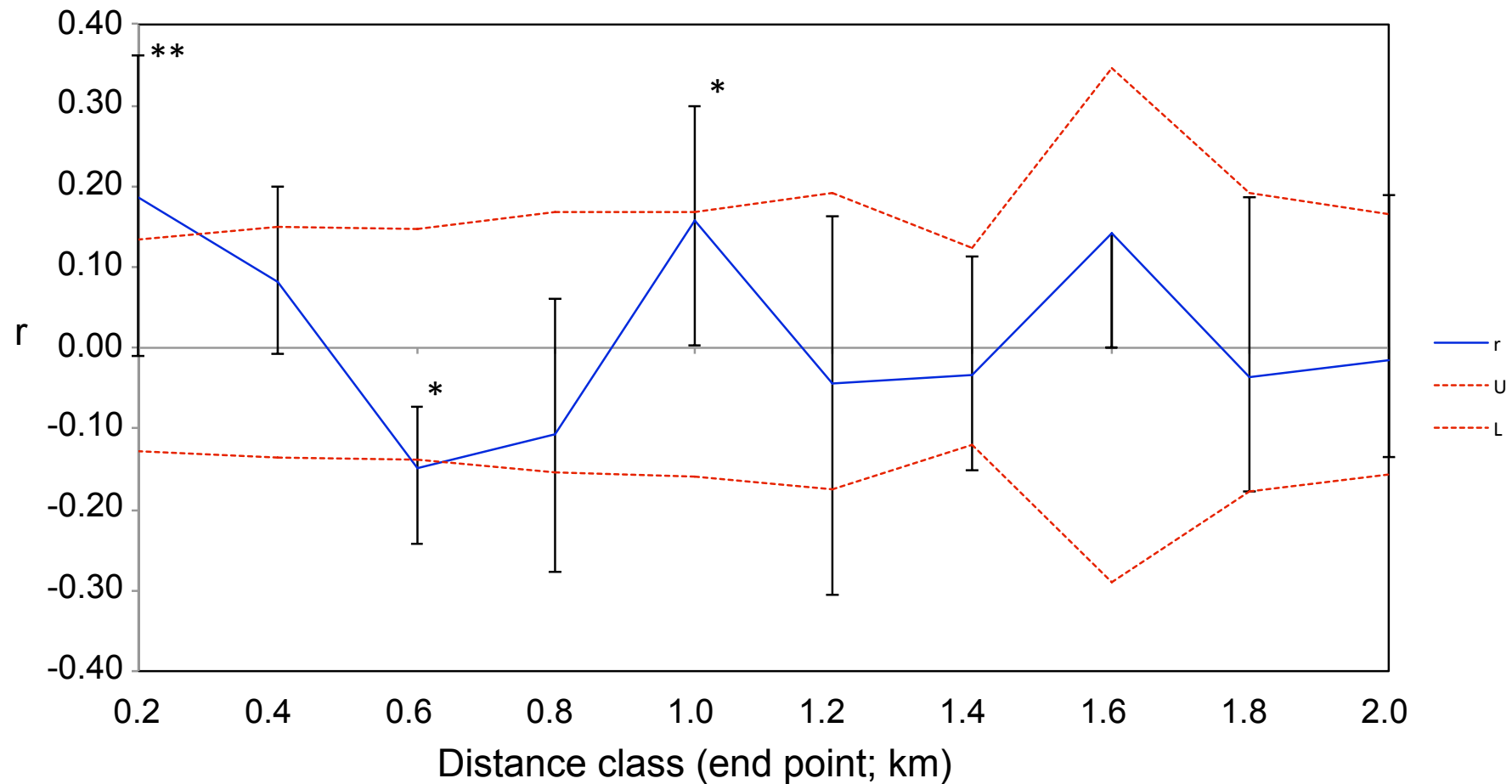


Figure 5.3 Spatial autocorrelation for all individuals (not including identical genotypes, $n = 84$) at 2 km size classes between 0 – 20 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

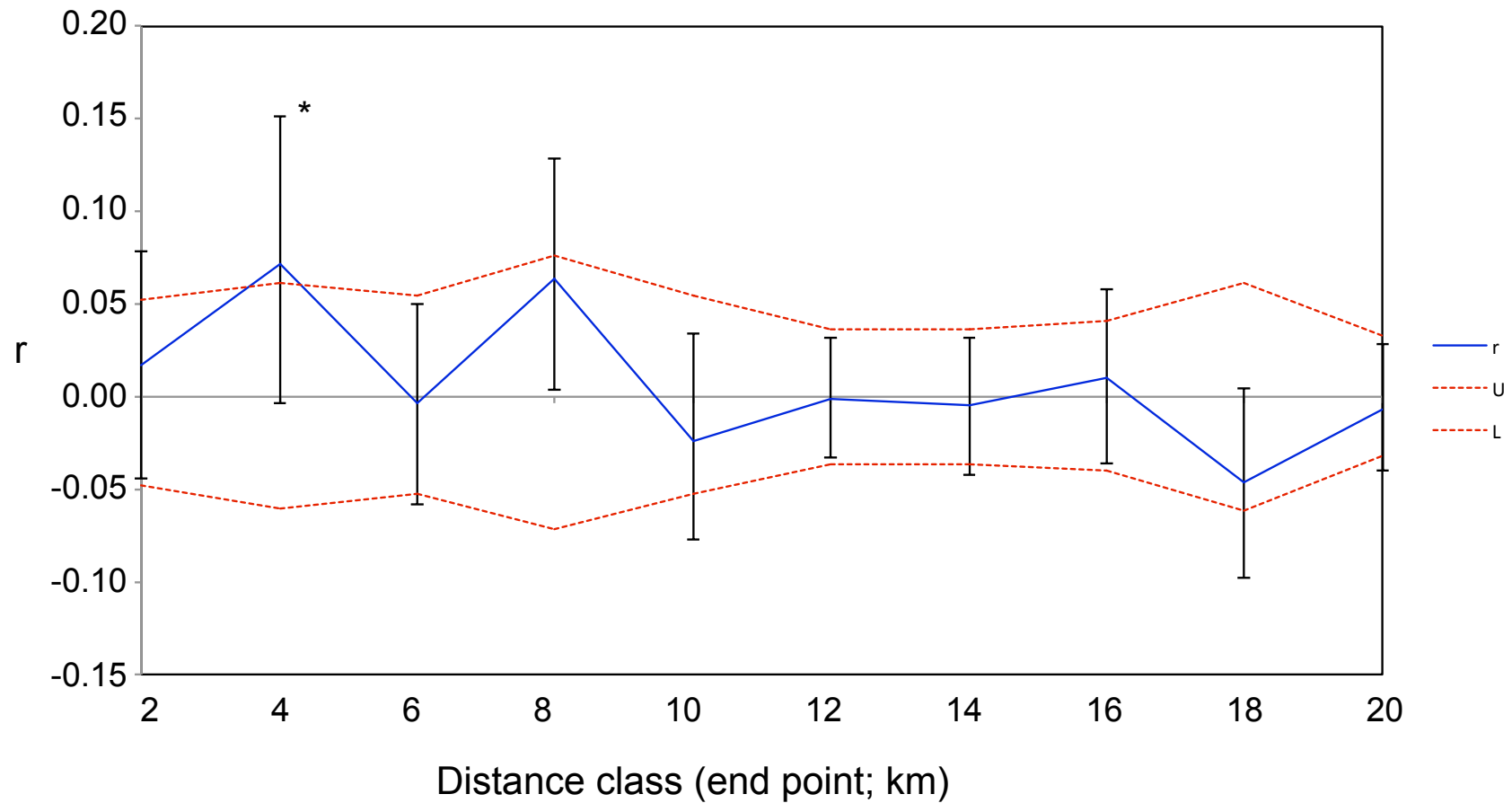


Figure 5.4 Spatial autocorrelation for all individuals (not including identical genotypes, $n = 84$) at 10 km size classes between 0 – 120 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

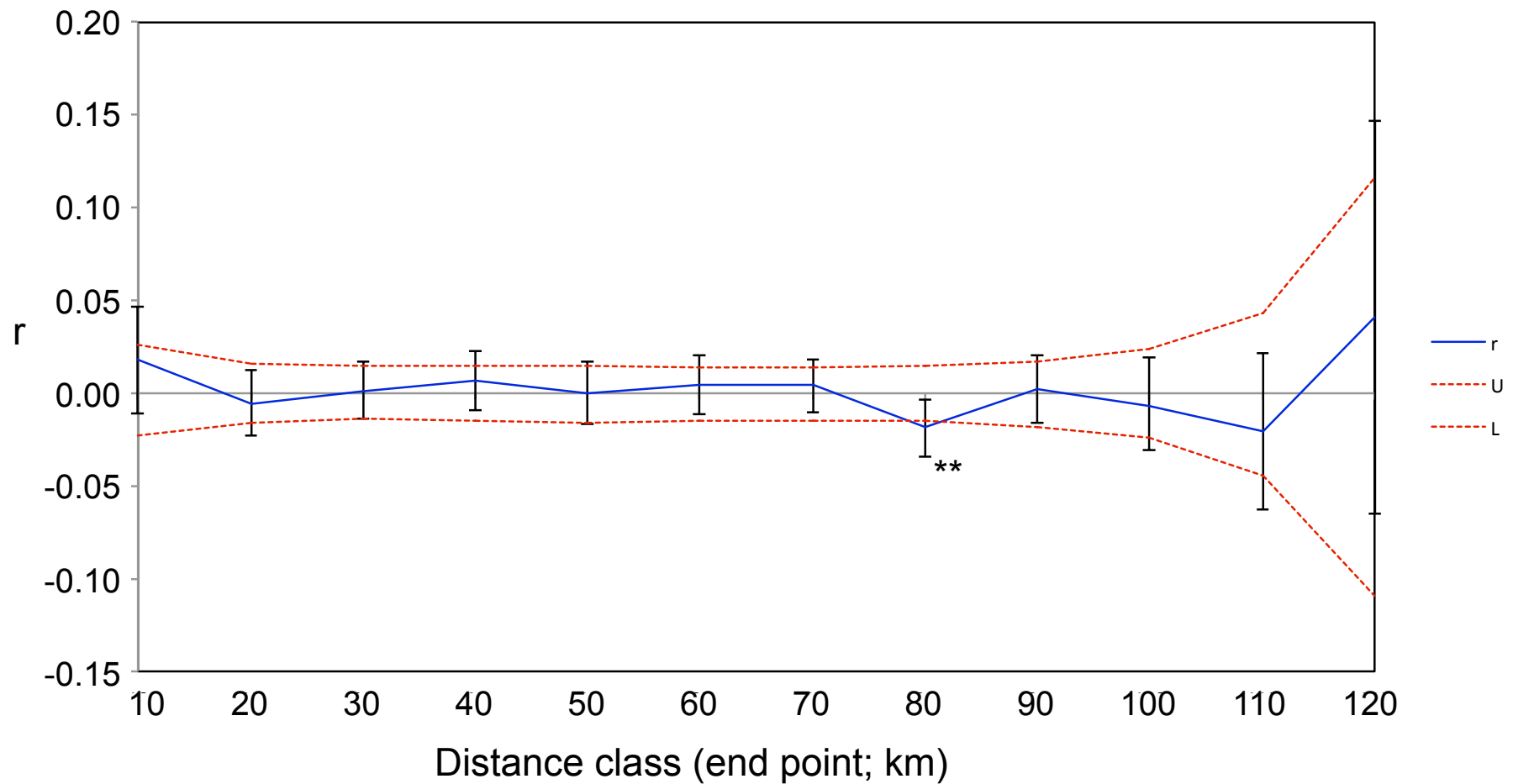


Figure 5.5. Spatial autocorrelation for males (not including identical genotypes, $n = 28$) at 2 km size classes between 0 – 20 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

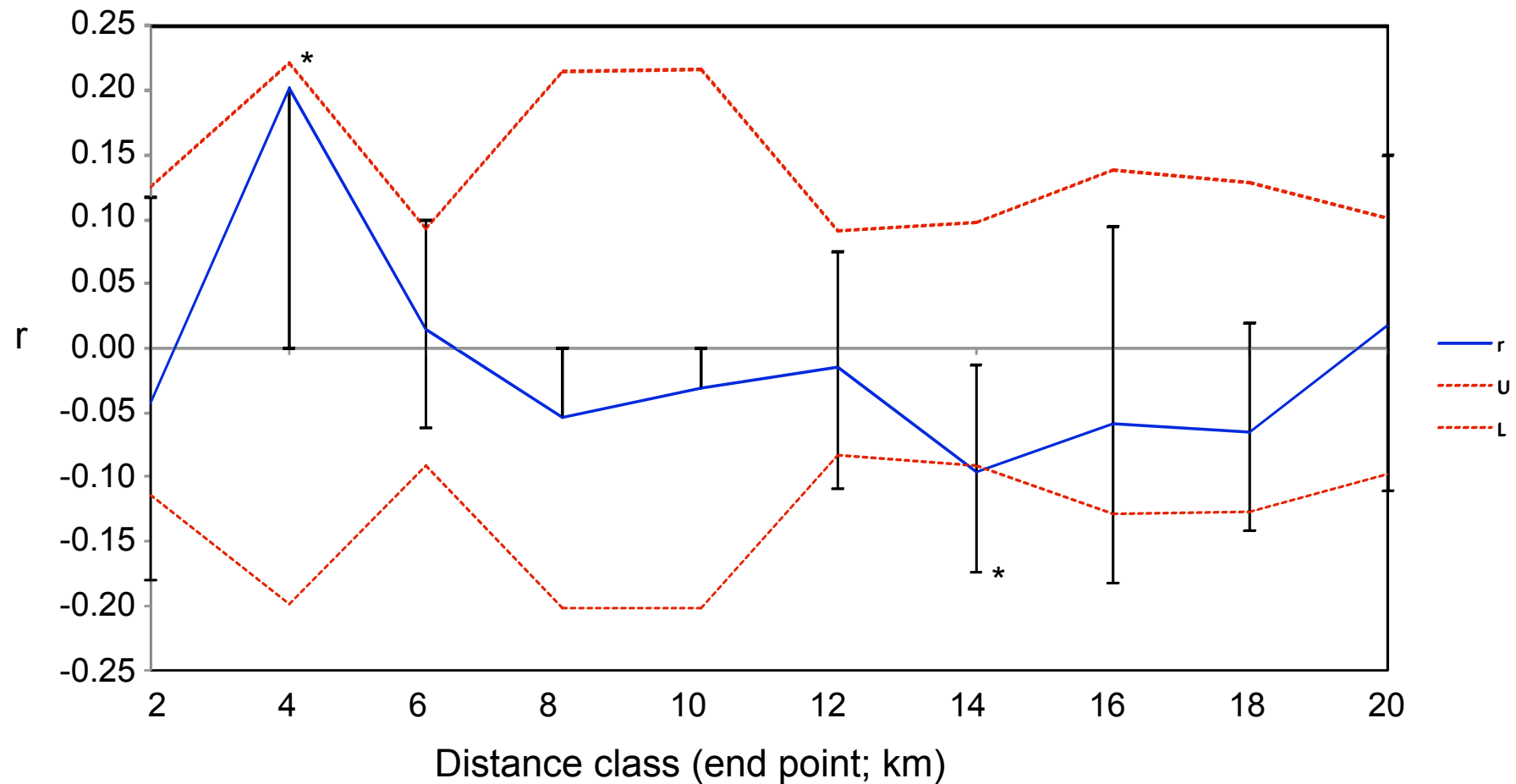


Figure 5.6. Spatial autocorrelation for males (not including identical genotypes, $n = 28$) at 10 km size classes between 0 – 120 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

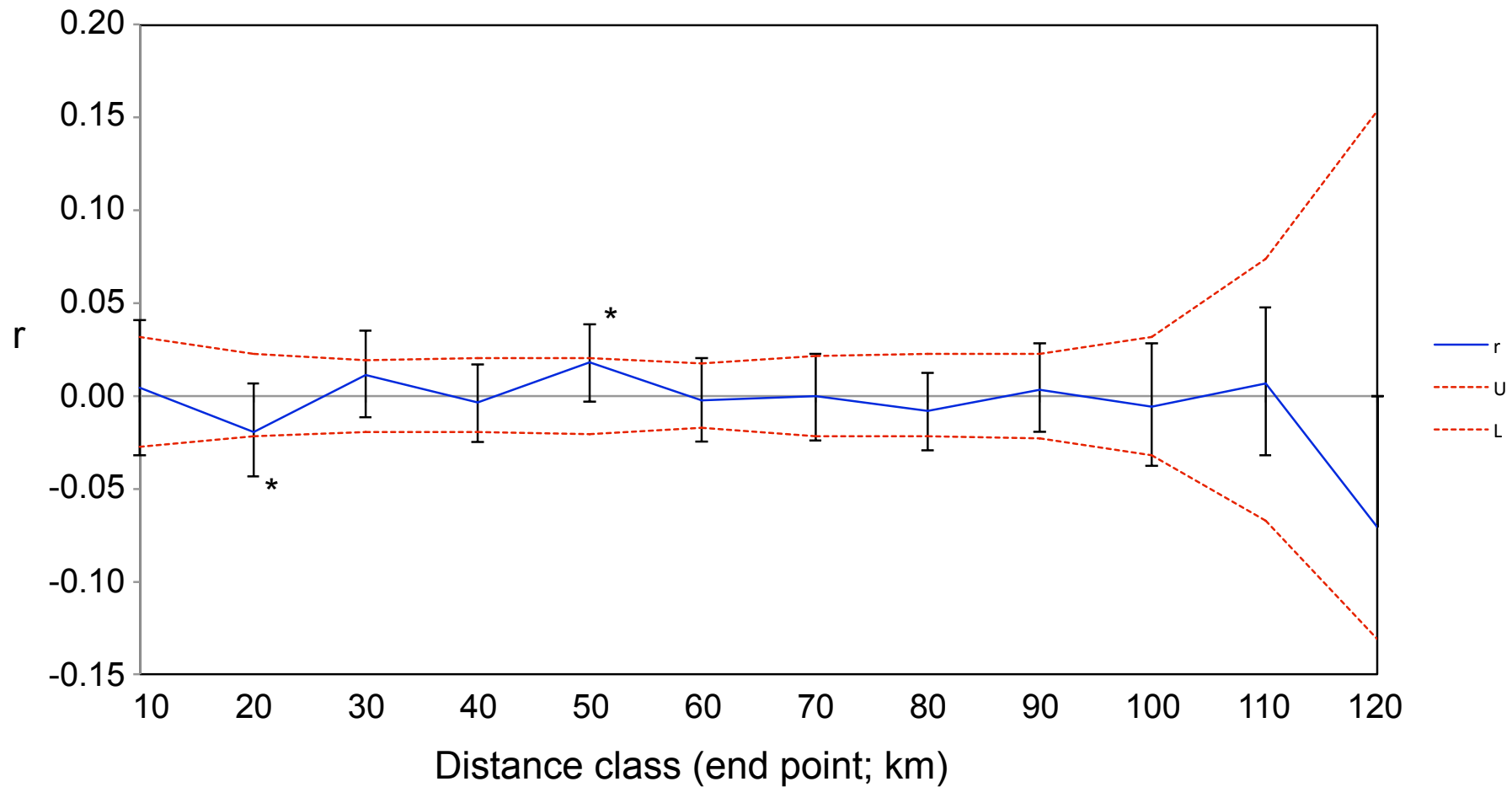


Figure 5.7. Spatial autocorrelation for females (not including identical genotypes, $n = 29$) at 2 km size classes between 0 – 20 km between 0 – 120 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).

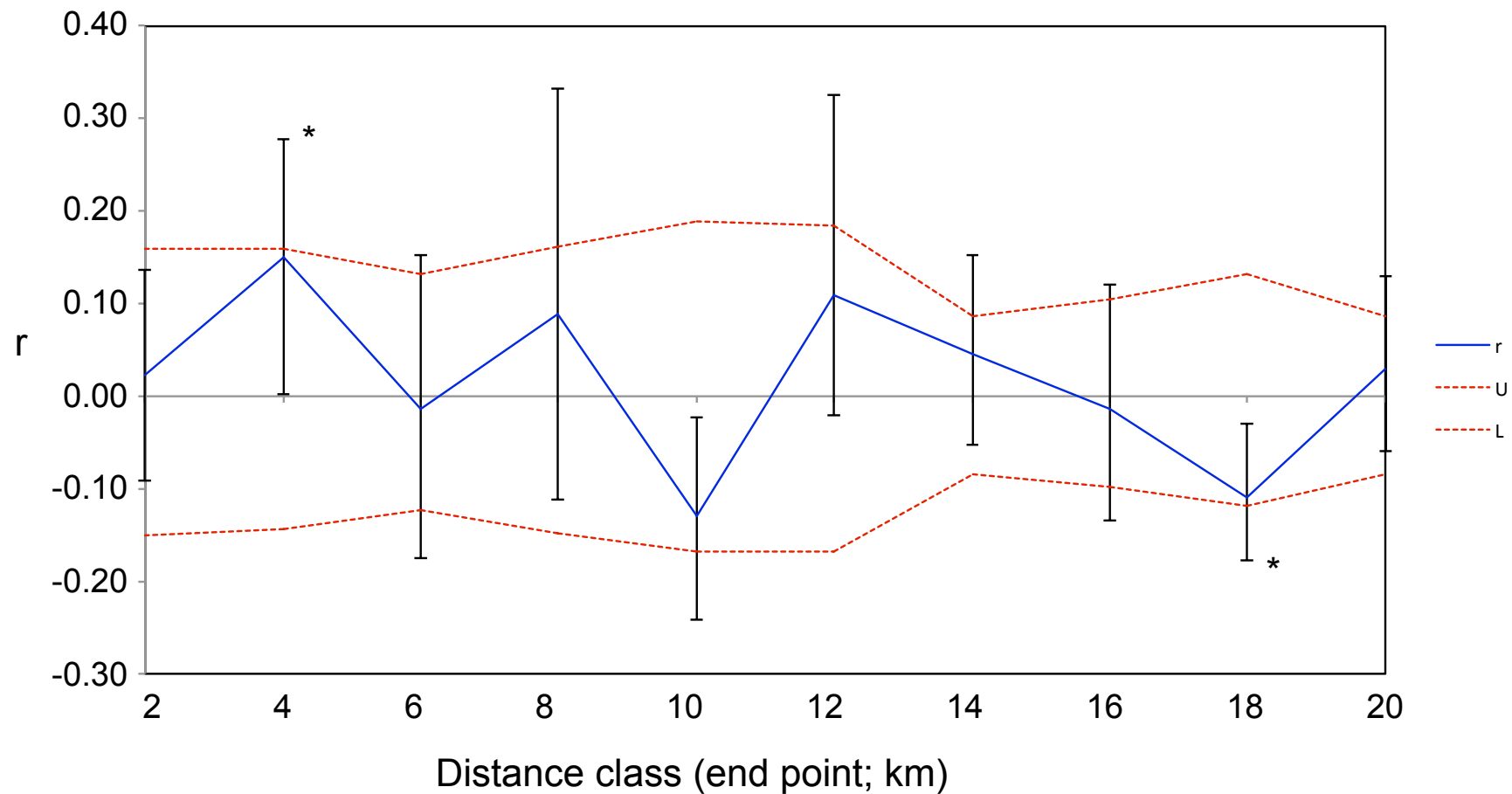
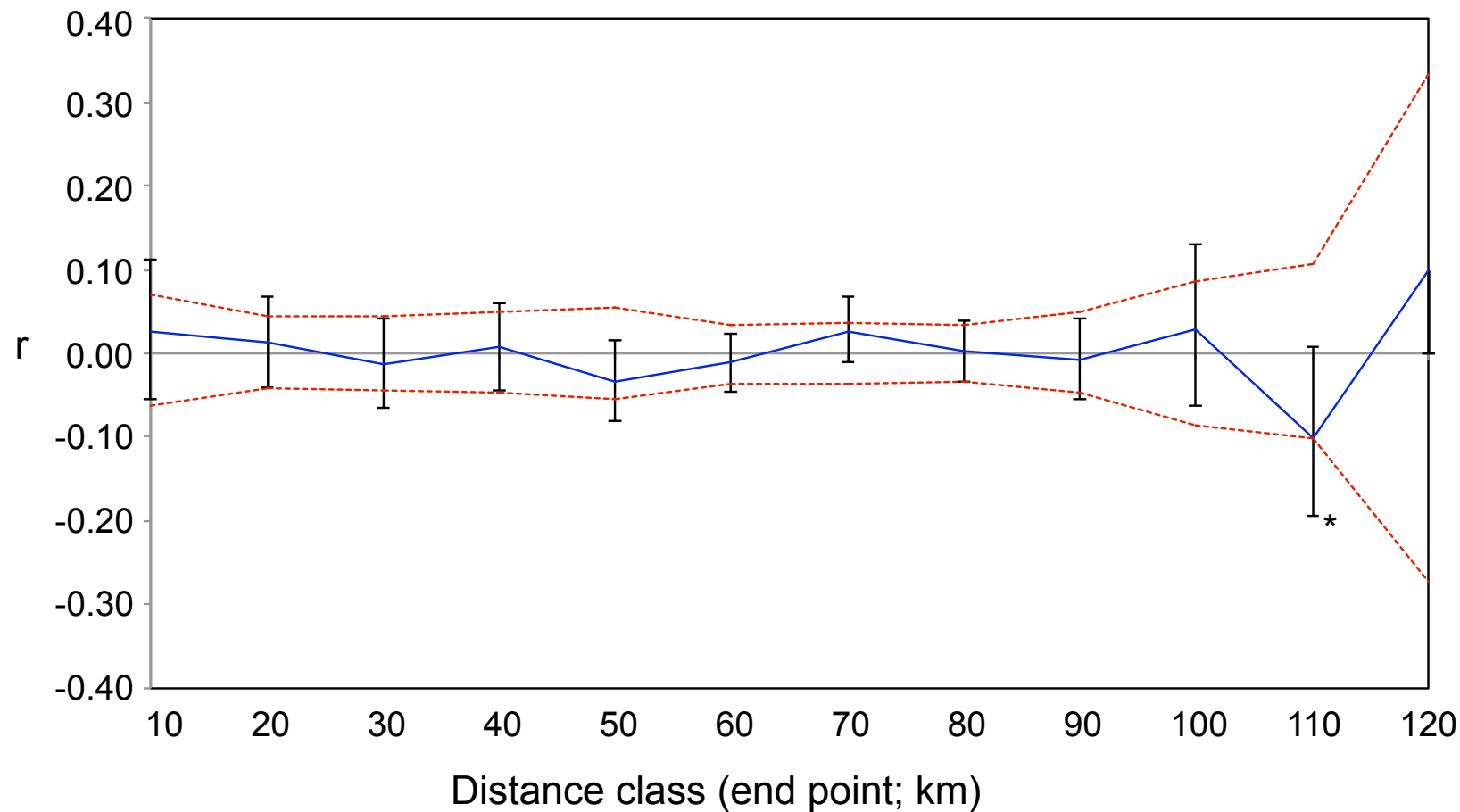


Figure 5.8. Spatial autocorrelation for females (not including identical genotypes, $n = 29$) at 10 km size classes between 0 – 120 km between 0 – 120 km. Significant values are indicated by a star (* < 0.05 , ** < 0.01 , *** < 0.001).



5.4.7 Sex-biased dispersal

There were no significant F_{ST} values between North and South or East and West sides of the study area for either males or females. Mean relatedness in males was significantly lower than in females (males: -0.0478, females: -0.0065, $p < 0.01$, $t = -2.907$), indicating that males were less related than females presumably because of a higher male immigration rate into the study area. Haplotype diversity in males was higher than in females, true for both the 325 bp (males: 0.8772, females: 0.8250) and the 543 bp (males: 0.9286, females: 0.9048) fragments of mtDNA CR. As mentioned above, STRUCTURE 2.3.4 (Pritchard et al., 2000) was unable to assign individuals to more than one population. This was also true when only males or females were considered. $mAlc$ for females was 0.85455, and for males was -0.91785 ($p < 0.05$). $vAlc$ for females was 8.61963 and for males was 10.77515 ($p = 0.2809$).

Figure 5.9. Graph showing change in relatedness as a function of number of microsatellite loci, estimated using RERAT (Schwacke et al., 2005). Changes in relatedness from additional loci is predicted using a trend line (power relationship, $R^2 = 0.999$).

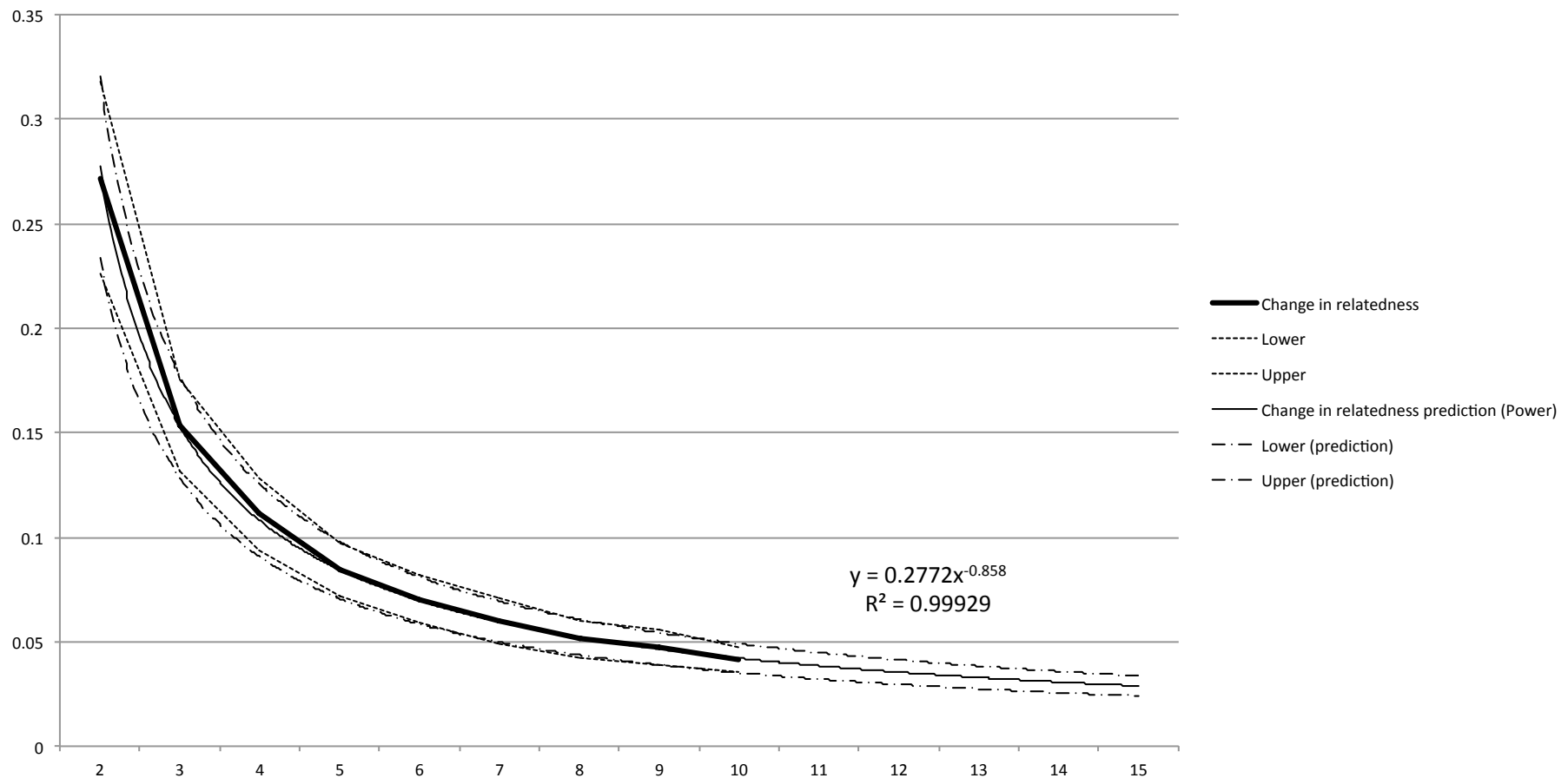
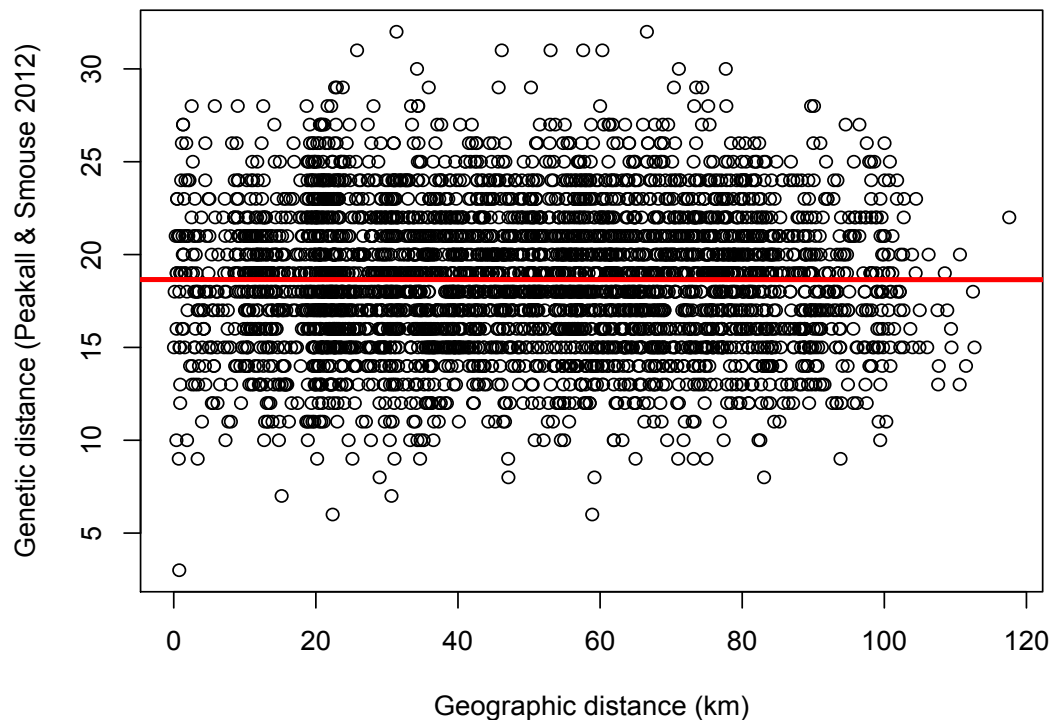


Figure 5.10. Scatter plot of genetic versus geographic distance of okapi in the RFO (not including identical genotypes, $n = 84$).



5.5 Discussion

This study aimed to elucidate information about okapi sociality, mating system and dispersal. Before this study was carried out, the only information available was some mixed reports on sociality (Johnston, 1900; Bodmer & Gubista, 1988; Hart & Hart, 1989). Any information that can be added to the little that is currently known about this species is therefore of great benefit to the species conservation efforts.

5.5.1 Okapi sociality

There is a great deal of variation in social structure amongst ungulates, and even among ungulates sharing a similar distribution to okapi. Blue duikers (*Philantomba monticola*) form permanent pairs, occupying exclusive home-ranges, whereas red duikers (*Cephalophus natalensis*) are solitary with greatly overlapping home-ranges (Bowland & Perrin, 2009). Sitatunga (*Tragelaphus spekii*) are mostly solitary, however, do have a tendency to be gregarious for reasons related to food availability (Magliocca et al., 2002). Bongo (*Tragelaphus eurycerus spp.*) form social groups of approximately 10-20 individuals, and

groups have home ranges measured at between 19-49 km² (Klaus-Hügi et al., 2000). Sociality was investigated in the present study for okapi using a combination of spatial autocorrelation analysis, relatedness estimators, and a description of the pattern of identical genotypes in the dataset. Spatial autocorrelation generally showed a pattern whereby there was negative autocorrelation at the larger distances (in the female, male and combined datasets), and positive autocorrelation at the shorter distances. Unexpectedly, there was a positive autocorrelation at 50 km for the male dataset. A possible explanation for this result could be high male sibling dispersal distances, although this hypothesis would need to be tested in future studies. There was also a negative autocorrelation at 0.6 km. This could be explained by proximity of unrelated male-female mating pairs. Unfortunately, this result (based on the male-female combined dataset) could not be tested directly with males and females separately (at 0 - 2 km), as these datasets were not large enough at this distance class. Our results therefore demonstrate a detectable correlation between geographic and genetic distance at specific distance categories within the RFO (maximum distance between samples 118.7 km). Also, the negative autocorrelations were usually at only the largest pairwise distances, implying a limited effect of isolation by distance operating just below the extent of the study area. The positive autocorrelation at ≤ 1 km for both males and females is evidence of social interaction between relatives at this small spatial scale. As mentioned earlier, this dataset may contain juveniles, and so it is likely that these significant positive values are detecting mother-offspring dyads with a low but detectable level of spatial association, similar to that described in (Hart & Hart, 1989).

Dyads with a relatedness estimate of greater than 0.5 had an average geographic distance that was lower than that of the dyads with a relatedness estimator less than 0.5. This was true for all seven estimators implemented in COANCESTRY, although this difference was only significant in two cases. This finding suggests a relatively weak overall correlation between relatedness and geographic distance, but with significant associations at the highest relatedness values. Although the difference in geographic distance is significant, the magnitude of this difference

is not particularly large (38.8 – 42.5 km vs 48.8 – 48.9 km). Taken together, the results of the spatial autocorrelation and relatedness patterns are indicative of a species where genetic structuring is determined more by relatively high dispersal ability, and a small proportion of spatially proximate dyads (for example mother-offspring) than by a tendency to form tight social groupings.

Only one genotype was detected at six of the eight multiple dung piles from the study site. The other two were found to be relatives. This finding again appears to show that okapi form small temporary social units (consistent with mother-offspring, or adult male-female), with no evidence for larger social groups of extended family members. The COLONY analysis was unable to distinguish between relationship classes for one of the dyads from the multiple dung piles, and the other dyad was a pair of half-siblings. The results from the multiple dung piles seems to indicate that large social stable units appear to be very unlikely to be formed in this species. We can therefore accept our first hypothesis, that okapi are mostly solitary animals. This social strategy has been predicted as a means of animals avoiding predator detection (Eisenberg & Lockhart, 1972; Geist, 1974), consistent with the ecology of okapi, which are known to be predated heavily by leopards (Hart & Hart, 1989).

5.5.2 Okapi mating systems

COLONY assigned one dyad to be full siblings ($p = 0.999$) and 207 and 175 half-siblings with posterior probabilities of greater than 0.95 and 0.80, respectively. We can therefore accept our second hypothesis, that okapi are *genetically* polygamous or promiscuous. This is not unexpected, as monogamous mating systems occur in only ~5-15% of all mammalian species (Kleiman, 1977; Clutton-Brock, 1989; Dobson et al., 2010). Also, even in predominantly monogamous animals, a detectable level of promiscuity often occurs (Gren et al., 1989; Welsh & Sedinger, 1990; Hasselquist & Sherman, 2001; Reichard, 2010). Among the hypotheses advanced for the function of polygamy and promiscuity are that they may function as a means of reducing genetic incompatibility for a particular sex (usually females; Stockley, 2003) or that they may be under selection on a particular sex (usually males) to dominate a large number of

females (King et al., 2011). Some hypotheses suggest that social monogamy may only evolve when males are unable to defend access to more than one female (Lukas & Clutton-Brock, 2013). The findings of Hart and Hart (1989) suggest that this is not the case for okapi, with male home-ranges overlapping with several females in the RFO. We can therefore conclude that the mating system of okapi is most likely to be genetic and social polygyny or promiscuity. Our results cannot rule out social monogamy in okapi, however do make this mating system much less likely.

It is worth mentioning that this mating system is highly dependent on the abundance and distribution of individuals, and relies on there being enough females occupying small enough adjacent territories to be defended by a single male (Emlen & Oring, 1977; Hilgartner et al., 2012). This would be much more likely to be the case in the RFO, a region where okapi density is thought to be relatively high (Nixon & Lusenge, 2008; Quinn et al., 2013), although even in this habitat food appears to be a limiting factor (Hart & Hart, 1989). Mating systems can vary within a species, depending on variations in resource distribution, predation pressure and costs of sociality (Cockburn, 1988; Clutton-Brock, 1989; Kappeler, 1999). These factors are likely to vary greatly across the okapis range, potentially leading to different mating strategies in different regions.

When classing multiple dung piles as a single genotype, average distance between identical genotypes was 2.271 km (pairwise $n = 13$), or 0.337 km excluding the largest movement event detected (25.5 km; pairwise $n = 12$). The duplicated genotypes, excluding the largest movement event, all fall well within even the smallest home-range size previously measured for okapi (Hart and Hart 1989). The movement event of 25.5 km was by a male, and represents the only *potential* dispersal event ever recorded for this species. This is a direct estimate of dispersal, and as such it cannot be determined if this corresponds to a successful dispersal event (i.e. resulted in a mating), or even if this move was a permanent one. Nonetheless, this is valuable information as it clearly gives some indication of the movement potential of okapi.

The spatial autocorrelation analysis in the present study detected genetic structure, whereas IBD analysis did not. It is likely that the spatial scale investigated in this study is not large enough to detect a correlation between genetic distance and geographic distance, which would likely emerge if a larger spatial scale were investigated. The significant spatial autocorrelation results indicate a relationship between geographic and genetic distance that is only acting at certain distance classes. This signal may be lost in the IBD analysis, which simultaneously investigates all distance classes. Other studies have identified local genetic structure that is likely to have caused isolation by distance at large spatial scales (e.g. badgers; Pope et al., 2006).

5.5.3 Sex-biased dispersal

Male-biased dispersal is the norm for mammals (Greenwood, 1980), however, exceptions have been found. A notable example is the study of Zhan et al. (2007) who concluded that giant pandas demonstrate female-biased dispersal, based on $vAlc$ values, mean spatial distances between individuals, and estimates of relatedness, F_{ST} and population genetic structure. We can accept a hypothesis of male-biased dispersal in okapi, based on i) significantly lower pairwise relatedness in males than females within our study site, ii) higher haplotype diversity in males than females, and higher $mAlc$ for females than males. Differences in F_{ST} , microsatellite based genetic structure and $vAlc$ were not significant. The lack of significant difference between F_{ST} values may be due to the limited power of the statistic. It is not unusual for only a subset of these tests to give significant values (e.g. Zhan et al., 2007), as they have variable power depending on demographic parameters specific to the sampled population, for example dispersal rate (Goudet et al., 2002). The hypothesis of male-biased dispersal can still be accepted with confidence due to multiple lines of evidence pointing towards this fact. This information is vital for okapi conservation plans. Dispersal is one of the main drivers in species persistence, especially in spatially structured populations (Bowler & Benton, 2005). This will become an increasingly important factor to consider in okapi conservation plans if deforestation continues at the current rate in the DRC. Notably, the spatial autocorrelation also shows that there is a spatial association between both males

and females at small distances (< 5km), showing that in okapi, both sexes exhibit some degree of social behaviour at small spatial scales. This pattern of positive spatial association for *both* males and females at small distance classes is a relatively common phenomenon (e.g. birds (Rollins et al., 2012), badgers (Pope et al., 2006) and wombats (Walker et al., 2008)), but does not appear to obviate these species from demonstrating considerable sex-biased dispersal.

The present study has made an important first step in describing sociality, mating systems and dispersal for okapi. These ecological features have important evolutionary consequences (Storz, 1999; Ross, 2001; Archie & Chiyo, 2012), and is a requirement for effective conservation management (Lacy, 1993). This information is therefore crucial for the conservation of this elusive, endangered giraffid.

5.6 Acknowledgements

We gratefully acknowledge the assistance of the *Institut Congolais pour la Conservation de la Nature* (ICCN) for allowing the field collection and export of samples. We would like to particularly thank the Wildlife Conservation Society (WCS), Gilman International Conservation (GIC) and the Okapi Conservation Project for their considerable assistance with the fieldwork in this study, and the amazing okapi conservation work that they carry out in the DRC. We thank the Frankfurt Zoological Society, the Lukuru Foundation/TL2 project, the Zoological Society of London (ZSL) and the large number of people who assisted in various ways with sample collection. In particular, we would like to thank John Fataki Bolingo, Bryna Griffin, Terese Hart, Chrysostome Kaghoma, Luaison, John Lukas, Kambale Magloire, Ephrem Mpaka, Stuart Nixon, Linda Penfold, Elise Queslin, Alex Quinn, Rosemary Ruf. This project was jointly funded by the UK Natural Environment Research Council (NERC) and ZSL (ZSL as NERC CASE industry partner and via an Erasmus Darwin Barlow Expedition grant). We also acknowledge the financial support provided by GIC (the Okapi Conservation Project), the US Fish and Wildlife Service, the UK's Darwin Initiative and the Mohamed bin Zayed Species Conservation Fund.

6

CHAPTER SIX – Ex-situ conservation



"Life begins at the end of your comfort zone."

Neale Donald Walsch

CHAPTER SIX - An integrated genetic management assessment for the okapi (*Okapia johnstoni*)

David W. G. Stanton^{1*}, Philippe Helsen², Jill Shephard^{2,3}, Kristin Leus⁴, Linda Penfold⁵, John Hart⁶, Noëlle F. Kümpel⁷, John G. Ewen⁸, Jinliang Wang⁸, Peter Galbusera^{2,§}, Michael W. Bruford^{1,§}

Submitting: Conservation genetics (18th April 2014)

§Joint senior authors

¹School of Biosciences, Cardiff University, Cardiff CF10 3AX, United Kingdom

²Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningen Astridplein 26, 2018 Antwerp, Belgium

³School of Veterinary and Life Sciences, Murdoch University, Perth 6150, Australia

⁴København Zoo, CBSG Europe, Belgium

⁵South-East Zoo Alliance for Reproduction & Conservation, Yulee, FL, USA

⁶Lukuru Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), 1235 Ave Poids Lourds, Kinshasa, DRC

⁷Conservation Programmes, Zoological Society of London, London NW1 4RY, United Kingdom

⁸Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom

* Lead author contribution: All labwork and data analysis. A subsection of the sample collection. Preparation of first manuscript draft, and modifications of manuscript based on co-author comments.

6.1 Abstract

Captive breeding programs increasingly use molecular genetics to inform their management strategies. Molecular approaches can be useful for investigating founder relatedness, and for estimating genetic diversity in a captive population. Genetic data can also be used to evaluate the representation within a captive gene-pool of the past and present genetic diversity found in wild populations. Maintaining a captive population that is genetically representative of its wild counterpart offers a means of preventing adaptation to captivity. We carried out a genetic assessment the *ex-situ* okapi (*Okapia johnstoni*) population, alongside an investigation into the genetic structure of wild populations across their geographic range. Okapi are endemic to the Democratic Republic of Congo, and have recently been reclassified as Endangered by the IUCN. We show that while levels of nuclear genetic variation in the wild, founder and captive okapi populations are similar, mitochondrial genetic diversity within captive okapi is considerably reduced compared to the wild. Further, both nuclear and mitochondrial alleles present in captivity poorly represent the allelic diversity

present in the wild. The translocation of individuals between captive breeding programs may improve representation of wild allelic diversity in captivity, however this must be accompanied by further genetic characterisation of the global captive population.

6.2 Introduction

A fundamental objective of captive breeding programs should be to provide a source of individuals for eventual reintroduction into the wild, for which best practice requires the effective maintenance of genetic diversity (Kozfkay et al., 2008; Lacy, 2013). At present, many captive breeding programs use pedigrees to retain genetic diversity, reducing genetic drift and minimising inbreeding by minimising mean kinship among captive individuals (MK strategies; Rudnick & Lacy, 2007). A complete pedigree is usually more informative, in terms of relatedness, than molecular markers (Baumung & Sölkner, 2003; Fernández et al., 2005), and management of pedigrees is a more accurate way to prevent inbreeding (Santure et al., 2010; Townsend & Jamieson, 2013). However, pedigree management alone does not provide a complete picture of diversity when founder relatedness, genetic diversity and wild origin are unknown (Henkel et al., 2012). Molecular methods are therefore increasingly being combined with pedigree-based *ex-situ* management to assess and monitor captive conservation programs, and to correct for relatedness estimates among founders (Fienieg & Galbusera, 2013). These analyses have usually used nuclear markers such as microsatellites to estimate a simple measure of genetic diversity (such as heterozygosity or allelic diversity) (e.g. Forstmeier et al., 2007; Shen et al., 2009; Gonçalves da Silva et al., 2010; McGreevy et al., 2011), or relatedness (Santure et al., 2010; Townsend & Jamieson, 2013).

However, monitoring genetic diversity alone does not ensure that all alleles of a wild population are represented in captivity. Also, even if the diversity of the wild population can be represented, genetic drift may cause allele frequencies to fluctuate from the present-day or ancestral wild population. This is because equalising founder contributions in practical captive breeding is often very challenging and some founders may become poorly represented or even absent

(Frankham, 2010). *Ex-situ* breeding programs therefore should aim to choose founders that capture as much of the wild genetic diversity as possible, for example by trying to match the genetic profile of the founder population to the wild (Miller et al., 2010). This may be important because the success of reintroduction programs is can be dependent on the extent to which a reintroduced population has adapted to captivity (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Jule et al., 2008), which may be determined by alleles that were deleterious and/or partially recessive in the wild (Frankham, 2008). However, through necessity, founders are often captured opportunistically or are already present in captivity (Hedrick et al., 1997; Russello et al., 2007; Hedrick & Fredrickson, 2008; Ivy et al., 2009). Given that it is becoming less acceptable to found or augment captive populations with wild individuals (e.g. Williams & Hoffman, 2009) it is important to assess whether the founders of a captive population constitute a representative sample of the wild population's genetic diversity (Miller et al., 2010). One approach to tackling this problem a posteriori (as opposed to a priori (Miller et al., 2010)), is via a genetic evaluation of both the wild and the captive and/or founder populations (Witzenberger & Hochkirch, 2011) implemented by an analysis of genetic structure and population assignment (e.g. Pritchard et al., 2000) and/or by using descriptive statistics evaluating levels of genetic differentiation (i.e. F_{ST} statistics) among wild and founder populations.

Mitochondrial DNA (mtDNA) has been used much less frequently than nuclear markers for informing captive breeding programs (e.g. Russello et al., 2007; Benavides et al., 2012), and when mtDNA has been used, analysis has also often been limited to simple measures of genetic diversity (Gautschi et al., 2003; Muñoz-fuentes et al., 2008; McGreevy et al., 2009; Lesobre et al., 2010; Khan et al., 2011; McGreevy et al., 2011). For some species, this is because low genetic diversity in wild and/or founding individuals has limited its utility (Hedrick et al., 1997). However, due to the unique inheritance properties of mtDNA (haploid, non-recombining and maternally inherited), it can also provide information about the phylogeographic origins and ancestral demography of captive individuals that is less easily interpreted when using nuclear markers alone

(Awise et al., 1984; Awise et al., 1987). In certain situations, mtDNA may therefore provide an important complementary perspective to nuclear loci for maintaining a captive population that is genetically representative of the wild.

Here we use both nuclear and mitochondrial DNA to inform captive management, using a case study: the okapi (*Okapia johnstoni*). Okapi are an even-toed ungulate, endemic to the Democratic Republic of Congo and are under threat from habitat fragmentation, human encroachment and poaching. They have recently been reclassified as Endangered by the IUCN (Mallon et al., 2013). Okapi have an ex-situ conservation program that is managed using a well-documented studbook (Leus & Hofman, 2012), which currently lists 173 okapi in captivity. As with many captive populations (e.g. Haig et al., 1992; Geyer et al., 1993; Gautschi et al., 2003; Russello et al., 2007), the wild origin and corresponding genetic structure of the founders are uncertain (Leus & Hofman, 2012).

It is currently unknown how representative the captive okapi population is of wild mitochondrial and nuclear genetic diversity and evolutionary history, information that is particularly important in this species in light of its recent reclassification (Mallon et al., 2013), with wild okapi populations thought to be declining rapidly (Okapi Conservation Workshop, 2013; Quinn et al., 2013). It is therefore highly plausible that genetic diversity will continue to be lost from the wild. Fourteen okapi in the captive okapi station in Epulu, Democratic Republic of Congo, were killed by Mai-Mai rebels in June 2012 (Okapi Conservation Project, 2012). These individuals constituted an unknown component of the genetic diversity of the global captive population, highlighting an added potential risk related to the spatial distribution of genetic diversity in captivity for this species. Knowing the distribution of genetic diversity in wild and captive okapi is therefore essential for ensuring that the diversity in captivity is as representative as possible of the wild population.

Here we use a combination of microsatellites and mtDNA to characterise and compare genetic diversity and structure of wild, founder and captive okapi,

including the individuals from the captive okapi station in Epulu. We aim to investigate if the founder and captive okapi populations were/are genetically representative of the wild. We assess this based on a comparison of the genetic structure at microsatellite loci, and a comparison of haplotypes for mitochondrial DNA. We describe the population genetic processes that may have lead to the pattern that is observed, and discuss the in- and ex-situ conservation implications.

6.3 Methods

6.3.1 Samples

A total of 365 samples, comprising 305 samples from the wild (247 dung samples, 44 museum skins and 14 confiscated skin samples), 33 captive individuals (24 blood or fresh muscle tissue and 9 dung samples), including eight from the captive okapi station in Epulu, DRC and 27 founder samples of the captive okapi population (16 museum tissue samples, 6 dung samples and 5 bone or tooth samples) were analysed. The founder samples were from the 34 individuals that either founded the captive okapi population, or are direct offspring of those that did. Founders were sampled, with permission, from the National Center of Scientific Research, Paris (France), Natural History Museum of Denmark , Copenhagen (Denmark), the Field Museum of Natural History, Chicago (USA), the Royal Museum for Central Africa, Tervuren (Belgium) and the okapi capture station, Epulu, DRC. For simplicity wild samples were classified as originating from one of four broad sampling locations corresponding to those assigned in Chapter 4 (and shown in Figure S6.1). The captive samples analysed here were selected by choosing individuals that represented as many of the founding lineages as possible. This was to try and ensure that the genetic diversity of the captive okapi population was accurately represented. The 33 captive individuals used in this study can trace ancestry back to all but one of the 34 individuals that founded the global captive okapi population *and* have living descendants. The single founding individual that is not represented has a founder representation (Lacy, 1989) of 2.3%. This founder individuals used in this study can therefore be considered representative of 97.7% of the living okapi captive population.

6.3.2 Lab methodology

DNA was extracted from blood, tissue, bone and teeth samples using a Qiagen Blood and Tissue Extraction Kit (Qiagen GMBH, Germany). DNA was extracted from fecal samples using a Qiagen DNA Stool Mini Kit (Qiagen GMBH, Germany). For all museum samples, DNA was extracted in a room dedicated to ancient samples. For these samples, some minor modifications were made to the extraction protocol to increase DNA yield, as follows:

- Samples were initially rehydrated in 9% NaCl solution for 48hrs under refrigeration at 4°C.
- Samples incubated overnight with Proteinase K at 55°C, using glycogen solution.
- Following overnight incubation with Proteinase K, samples were incubated with RNase at 37°C.
- Long term storage at -20°C.

PCR amplification of 13 microsatellite loci was attempted using the 365 samples described above. Primers and conditions described are in Stanton et al. (2010), although locus Oka-11 had to be excluded from analysis due to low PCR amplification rates. MtDNA PCR amplification used the mtDNA primers (OJ1-5) and conditions described in Chapter 4. These primers amplify a 422 bp fragment of the Cytochrome *b* (Cyt *b*), tRNA-Pro and tRNA-Thr genes, and a 473 bp fragment of the control region (CR). The present study used 69 wild, 26 captive and 12 founder individuals to generate the mtDNA dataset. The Cyt *b*, tRNA-Thr and tRNA-Pro sequences were concatenated into a single 422 bp sequence. Due to the high number of samples from the wild in sampling region 2 (Figure S6.1), a subset of 35 samples were chosen at random from this region as representatives for mtDNA analysis.

6.3.3 Data validation

GIMLET v1.3.3 (Valière, 2002) was used to calculate allelic dropout rate (ADR) and false allele (FA) rate for a preliminary subset of 14 okapi faecal samples. GEMINI v1.3.0 (Valière et al., 2002) was used to investigate the number of PCR repeats required to accurately create a consensus genotype for the full dataset. This preliminary analysis indicated that three repeats would create consensus

genotypes with 95% accuracy, and accuracy converged on 100% with four repeats. We therefore used a multitubes approach (Taberlet et al. 1996) with at least four repeats carried out for our full dataset of faecal samples. We typed samples as heterozygous at one locus if both alleles appeared at least twice among the four replicates and as homozygous if all the replicates showed identical homozygous profiles (as Bonin et al., 2004; Zhang et al., 2007). If neither of those cases applied, we treated the alleles as missing data.

Four repeats were also carried out for confiscated skin samples and between four and eight repeats were carried out for museum samples, depending on the amount of DNA extract available. PCR products were run, along with GeneScan ROX 350, or GS-400 HD LIZ, in a Prism 3700 Genetic Analyzer (Applied Biosystems) and analysed using the Genemarker© software package (version 1.9.1, SoftGenetics, LLC, State College, Pennsylvania). Once the full dataset was genotyped, GIMLET was used again to quantify ADR and FA rate for (i) dung samples and confiscated skin samples, and (ii) museum samples. There were five samples whereby only two repeats could be carried out due to insufficient DNA extract. GIMLET was used to calculate ADR and FA rate in these samples separately.

Duplicated genotypes (all alleles identical, or all but one allele identical, with missing data considered missing at that locus across all comparisons) were removed from the dataset and excluded from all future analyses. KINGROUP v2 (Konovalov et al., 2004) was used to identify 1st order relatives (at $p < 0.05$) in the dataset and these (wild) individuals were also excluded from the any further analyses.

6.3.4 Genetic diversity

Observed and expected heterozygosity for microsatellite data were calculated using GENALEX (Peakall & Smouse, 2012), and significance assessed using t-tests in R (R Development Core Team, 2011). Allelic richness was calculated using FSTAT v2.9.3 (Goudet, 1995).

MtDNA haplotype sequences were aligned in Sequencher 4.9 (GeneCodes) and four mtDNA contig alignments were created. These consisted of 372 bp of the cytochrome b (Cyt b) gene, the tRNA-Pro (66 bp) and tRNA-Thr (69 bp) genes and 328 bp of the control region (CR), and a concatenation of the Cyt b, tRNA-Pro and tRNA-Thr genes. Haplotype diversities were calculated using DNASP v5 (Librado & Rozas, 2009). Mitochondrial DNA spanning networks were drawn in TempNet (Prost & Anderson, 2011) for both the CR and concatenated mtDNA sequences. TempNet was used to create a multi-layered network, with the, captive and wild individuals separated into three different layers.

6.3.5 Genetic structure

Pairwise sample F_{ST} values for the microsatellite markers and AMOVA were calculated between the following sets of samples: 1) Captive (United States), 2) Captive (Epulu, DRC), 3) Captive Europe, 4) Founders (or direct offspring of founders), 5) Wild (sampling region 1; Figure S6.1), 6) Wild (sampling region 2), 7) Wild (sampling region 3) and 8) Wild (sampling region 4), using ARLEQUIN. These groupings were chosen to jointly investigate genetic differentiation between and within the four wild sampling regions, the founders, and the two captive populations.

For the AMOVA analysis, the "populations" were designated as the eight previously described sets of samples. The population grouping that best explained the genetic variance observed were investigated by testing a total of 13 different hypotheses. These hypotheses were grouped into four different "sets", the objective of each being to identify groupings that successively identified a better explanation for molecular variance that is observed in the data (highest significant percentage of variation among groups). The best hypothesis was then compared to other hypotheses in the next set. This is shown visually in Figures S6.2-6.5. Set one tests two hypotheses, with hypothesis one: samples are grouped into founder, captive and wild samples, and hypothesis two: as hypothesis one, except captive samples are split into three separate groups (European, US and Epulu). This first set therefore investigated if more molecular variance was explained by considering the captive populations as three separate groups. Set

two investigated whether the founders grouped with any of the four wild sampling regions more than the rest. This was implemented by testing four hypotheses, with the four wild sampling regions individually grouped with the founders. Set three investigated if the Epulu captive individuals should be considered in the same group as the wild sampling region 2 (the sampling region that those individuals were located). This was implemented using two hypotheses: 1. The Epulu captive individuals were grouped with the founders and wild sampling region two, 2. The Epulu captive individuals were grouped separately. Set four investigated whether any remaining sampling regions should have been classed as a separate group, or if the molecular variance could be better explained by them being combined into a single group. This was implemented by testing between five hypotheses, which included all the possible grouping arrangements of the remaining sampling regions.

Bayesian clustering analysis was performed using the program STRUCTURE 2.3.4 (Pritchard et al., 2000), with 500,000 MCMC iterations, a burn-in of 50,000, correlated allele frequencies and K set at 1-8. Six independent runs were carried out. The number of clusters was estimated using the method of Evanno et al. (2005). STRUCTURE was also run using only the founder and wild samples (n = 109) with K analysed between 1 and 4 (other settings were the same as above). This was done to attempt to assign the founder individuals to a part of the wild okapi's distribution.

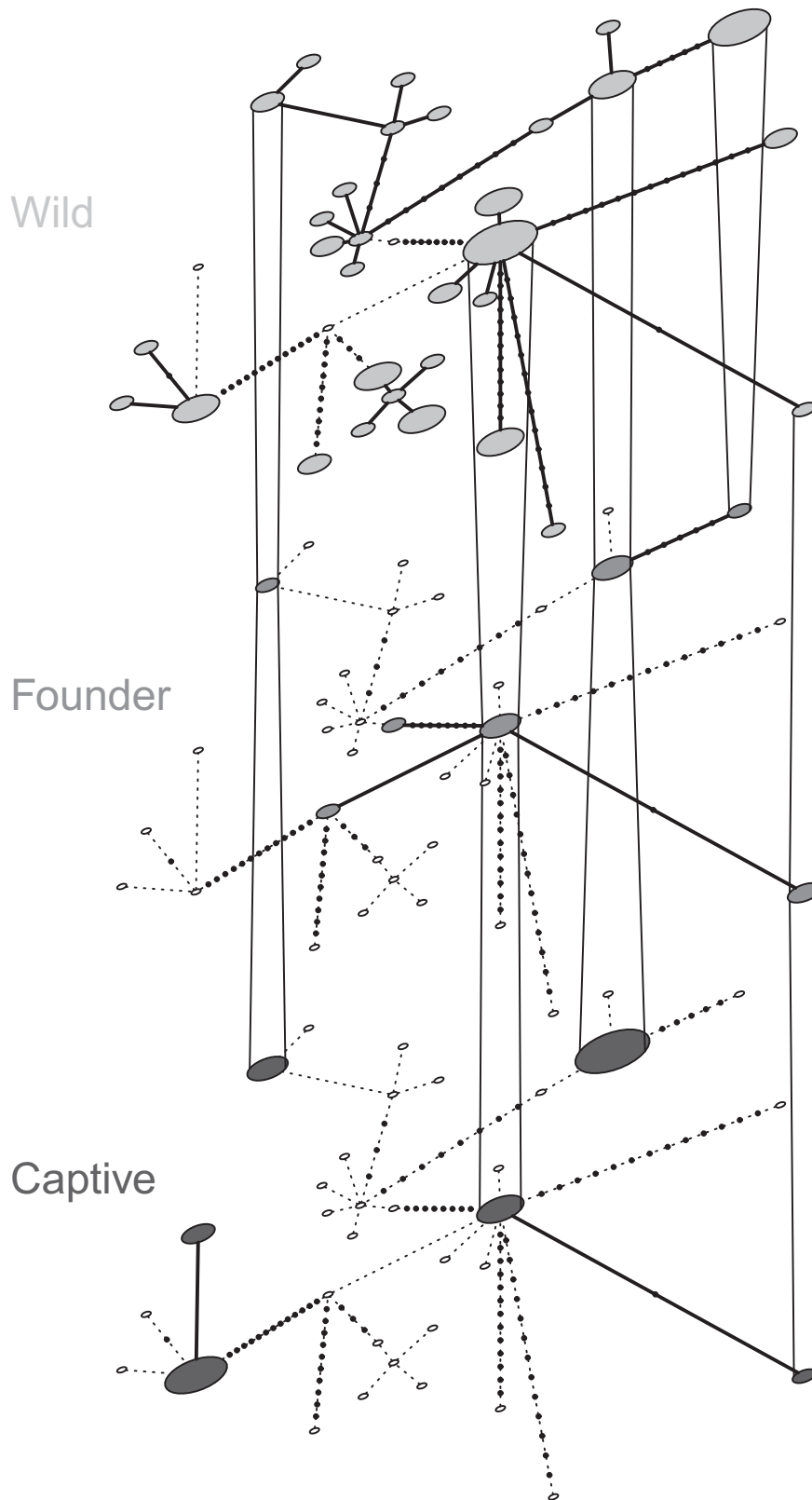
6.4 Results

6.4.1 Samples

After removal of samples based on the quality control described above, 134 wild samples (109 dung samples, 15 museum samples, 10 confiscated skin samples), 33 captive samples (24 blood and tissue samples and 9 dung samples) and 16 founder samples (10 museum tissue samples and 6 dung samples; 6 direct descendants of the founders and 10 founders) provided reliable consensus microsatellite data. After removal of identical genotypes and close relatives, the number of wild samples reduced to 93 (68 dung samples, 15 museum samples and 10 confiscated skin samples). The final mitochondrial DNA dataset contained

69 wild, 26 captive and 12 founding individuals for the CR and 53 wild, 23 captive and 10 founding individuals for the concatenated sequence.

Figure 6.1. 3D network drawn in TempNet (Prost and Anderson 2012) of the mitochondrial CR haplotype changes between wild, founder and captive okapi.



6.4.2 Genotype validation

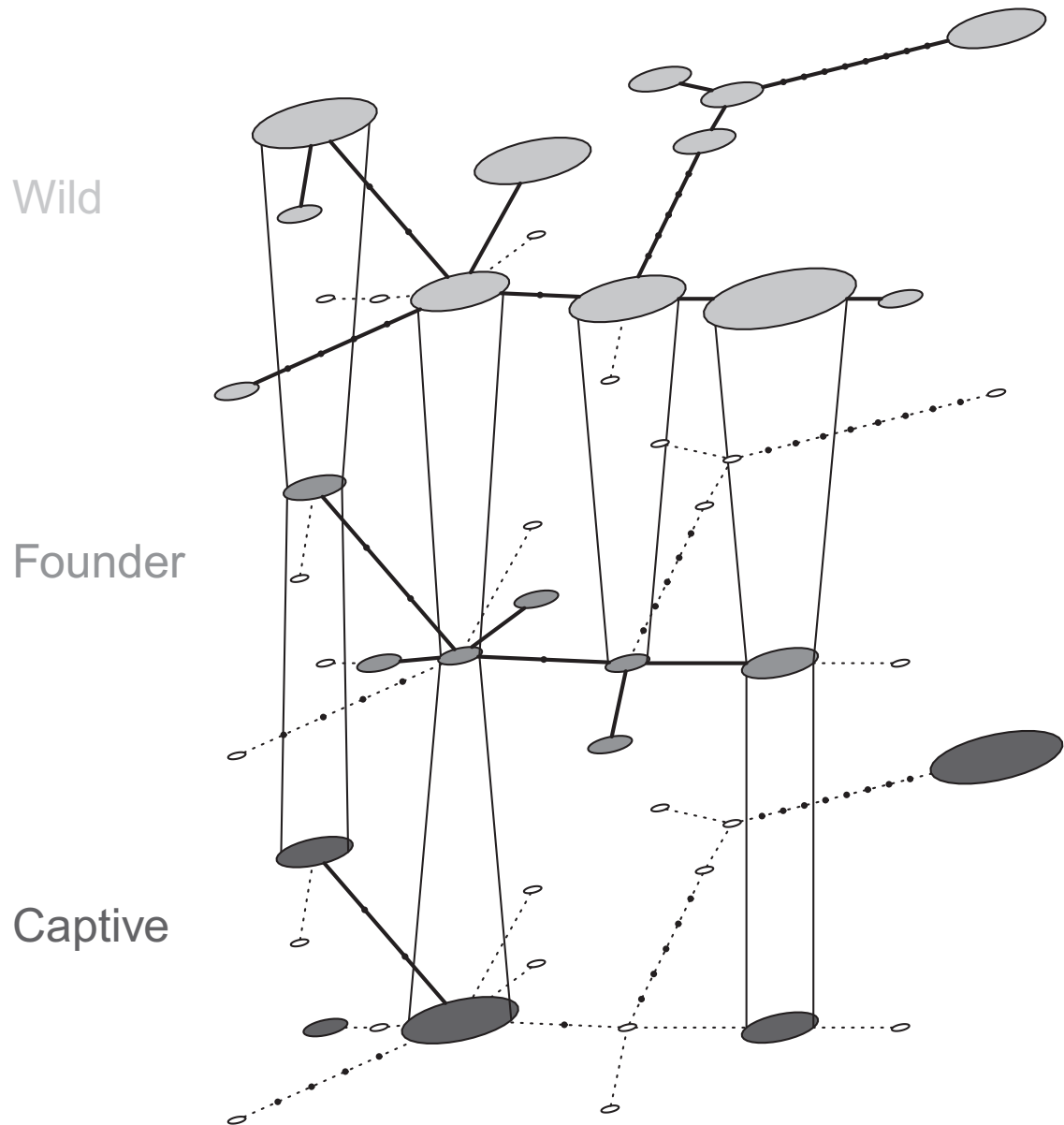
With GIMLET (Valière, 2002) the mean ADR for dung samples and confiscated skin samples was 0.041 (min 0.004, max 0.077) and the mean FA rate was 0.039 (min 0.010, max 0.136). The mean ADR of museum samples was 0.043 (min 0, max 0.259), and the mean FA rate was 0.029 (min 0, max 0.125). GIMLET was also used to estimate genotype error rate for the five samples where only two repeats could be carried out. Of these five samples, only one had an error rate above zero (founding individual, ADR 0.111). The following analyses were repeated excluding this individual, but this did not appreciably alter any of the results (data not shown).

6.4.3 Genetic diversity

For the microsatellite dataset, H_o of captive samples was 0.669 (SE 0.056), founding samples was 0.739 (SE 0.042) and wild samples was 0.688 (SE 0.034). H_e of captive samples was 0.755 (SE 0.025), founding samples was 0.753 (SE 0.023) and wild samples was 0.770 (SE 0.018). H_o and H_e of captive samples was therefore 2.8% and 1.9% lower than wild samples respectively (difference not significant based on t-tests [$p > 0.05$] for all pairwise comparisons). Allelic richness of captive samples was 5.376 (SE 0.398), founding samples was 5.756 (SE 0.413) and wild samples was 5.549 (SE 0.280). Observed heterozygosities of these groups are shown graphically in Figure S6.6. Haplotype diversity for CR sequences of wild samples was 0.932 (SD 0.013), for founder samples was 0.894 (SD 0.063) and for captive samples was 0.783 (SD 0.049). Haplotype diversity for Cyt b sequences of wild samples was 0.885 (SD 0.020), for founder samples was 0.911 (SD 0.077) and for captive samples was 0.723 (SD 0.052).

TEMPNET was used to visualise the mitochondrial haplotype changes between wild, founder and captive okapi. The 3D network is shown in Figure 6.1 for the mitochondrial CR fragment, and in Figure 6.2 for the concatenated fragment. Using the CR fragment, there are 31 haplotypes present in the wild population, 6 in the founder population and 6 in the captive population. Using the Cyt b fragment, there are 12 haplotypes present in the wild population, 7 in the founder population and 5 in the captive population.

Figure 6.2. 3D network drawn in TempNet (Prost and Anderson 2012) of the mitochondrial CB haplotype changes between wild, founder and captive okapi.



6.4.4 Genetic structure

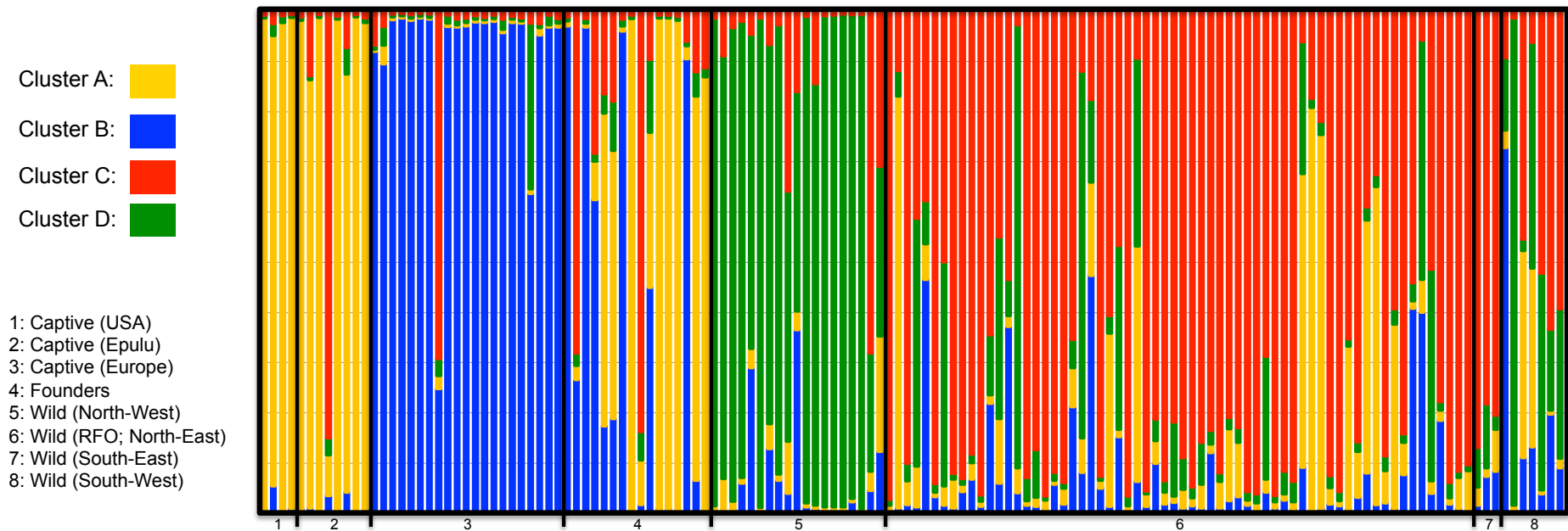
Thirteen hypotheses (arranged into four “sets”, Figures S6.2-6.5) of sample groupings were investigated (see Methods). Set one showed that the molecular variance in the dataset is best explained by considering the captive populations as three separate groups rather than a single one (AMOVA among group variation 0% [$p = 0.887$] versus 2.24% [$p = 0.190$]). Set two showed that of all the sampling regions, molecular variance is best explained when the founders are grouped with sampling region two (AMOVA among group variation 3.11%, $p = 0.030$). Set three showed that molecular variance is best explained when the captive okapi at Epulu are also grouped with the founders and the samples from sampling region two (AMOVA among group variation 3.43%, $p = 0.004$). Set four showed that the molecular variance is best explained when the okapi from sampling regions three and four are grouped together, but separately from sampling region one (AMOVA among group variation 4.63%, $p = 0.002$).

Table 6.1. Table of pairwise F_{ST} values. Numbers 1-8 refer to the following sample sets: 1) Captive (United States), 2) Captive (Epulu, DRC), 3) Captive Europe, 4) Founders (or direct offspring of founders), 5) Wild (sampling region 1; Figure S6.1), 6) Wild (sampling region 2), 7) Wild (sampling region 3) and 8) Wild (sampling region 4). F_{ST} values that are significantly ($p < 0.05$) different from 0 are indicated by an asterisk.

	1	2	3	4	5	6	7	8
1	0.00000							
2	0.04138	0.00000						
3	0.15160*	0.11597*	0.00000					
4	0.03595	0.01404	0.03603*	0.00000				
5	0.10639*	0.10490*	0.09282*	0.07035*	0.00000			
6	0.06870*	0.02741*	0.05679*	0.01637*	0.04628*	0.00000		
7	0.07303	0.04522	0.00465	0.00606	0.04522	-0.02363	0.00000	
8	0.10677*	0.08104*	0.05832*	0.03753*	0.05576*	0.04324*	-0.01489	0.00000

Pairwise F_{ST} values between sample sets 1-8 (described above) are given in table 6.1. Significant F_{ST} values ranged between 0.016 ($p = 0.005$; between the founding individuals and wild [sampling region 2] samples) and 0.152 ($p < 0.001$; between captive [US] and captive [Europe] samples). Other significant F_{ST} values above 0.1 were between captive Epulu and European samples ($F_{ST} = 0.116$, $p < 0.001$), between captive [US] and between wild [sampling region 4] ($F_{ST} = 0.107$, $p = 0.016$), between captive [US] and between wild [sampling region 1] ($F_{ST} = 0.106$, $p < 0.001$) and between captive [Epulu] and wild [sampling region 1] ($F_{ST} = 0.105$, $p < 0.001$).

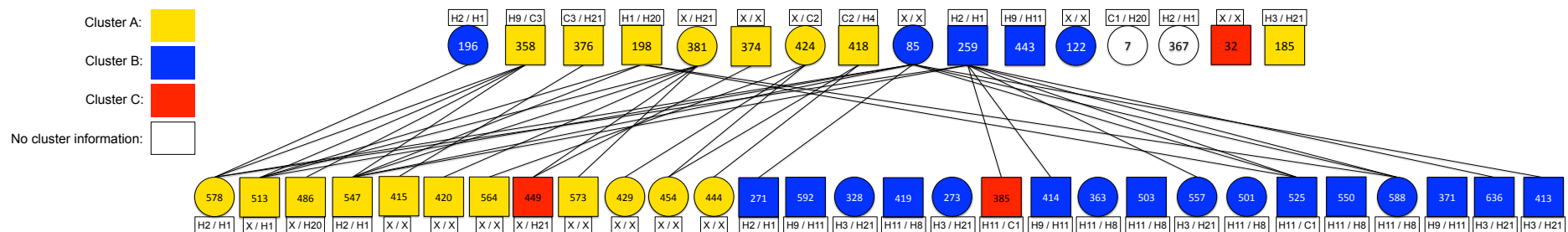
Figure 6.3. Structure plot of founder, captive and wild okapi samples for $K = 4$. Numbers correspond to: 1) Captive (United States), 2) Captive (Epulu, DRC), 3) Captive Europe, 4) Founders (or direct offspring of founders), 5) Wild (sampling region 1; Figure S6.1), 6) Wild (sampling region 2), 7) Wild (sampling region 3) and 8) Wild (sampling region 4).



STRUCTURE was used to further investigate the extent of genetic structure across the founder, captive and wild okapi. All individuals were assigned to the same cluster across all six independent STRUCTURE runs. The Evanno et al. (2005) method (which uses the rate of change (Δ) in Ln Likelihood between successive values of K [ΔK]) estimated the most appropriate number of genetic clusters in the complete microsatellite dataset to be four (Figure S6.7). The STRUCTURE plot for $K = 4$ is shown in Figure 6.3, with samples grouped into the eight sample sets described above. The most appropriate number of genetic clusters for the wild samples only was estimated to be five, although this result was not as clear as it was for the complete dataset, and the majority of individuals did not segregate into any cluster. The STRUCTURE results for only the wild samples are therefore not presented here. STRUCTURE was unable to assign founders to a wild sampling region as there was insufficient genetic structure (based on the STRUCTURE results) between wild populations to assign the founders to.

Figure 6.4 shows the founder and captive individuals used in the present study for which there was mitochondrial haplotype and/or STRUCTURE clustering information. Based on our samples, the European captive population was contributed to predominantly by founders belonging to one cluster (cluster B, from Figure 6.3; 83.3%). The Epulu, DRC captive population was contributed to entirely by founders belonging to a second cluster (cluster A, 100%). The majority of the US captive population was contributed to by founders belonging to the cluster A (57.1%). Of the eight founders belonging to cluster A, 78.6% of their descendants also belonged to that cluster, 14.3% to cluster B and 7.1% to a third cluster (cluster C, Figure 6.3). Of the five founders belonging to cluster B, 33.3% of their descendants belonged to cluster A, 60.0% to the cluster B and 6.7% to cluster C.

Figure 6.4 Pedigree showing okapi founders and captives with STRUCTURE groups and haplotypes. Each individual is represented by either a square (male) or a circle (female), with studbook ID denoted by the central number. STRUCTURE group is shown by the shading of the individual, and corresponds to Figure 6.3 (missing data is shown by a cross hash). Mitochondrial DNA haplotype is shown either above or below the individuals (cytochrome b / control region). Haplotypes prefixed by an “H” correspond to haplotypes previously found in the wild (Chapter 4) and those prefixed by a “C” correspond to new haplotypes found in the present study. If mitochondrial DNA information is missing, the haplotype is denoted as an “X”. Straight lines connect ancestors/descendants, and these are not necessarily first order relatives. Only the data is shown for individuals for which there was at least one of either mitochondrial DNA sequences or STRUCTURE population assignment data.



The okapi studbook was used to estimate contributions of each of the founders used in the present study into the European, US and Epulu captive breeding programs (including individuals that were not sampled in the present study). The founders grouped by the STRUCTURE analysis into the cluster B contributed 83.3% of their descendants into the European captive breeding program (of which 77.0% was from only two individuals [85 and 259, Figure 6.3]). The founders grouped by the STRUCTURE analysis into cluster A contributed 44.6% of the descendants (of the founders used in the present study) into the US captive breeding program. The founders grouped by the STRUCTURE analysis into cluster A contributed all of the descendants (of the founders used in the present study) into the Epulu captive breeding station. The above results demonstrate the uneven contribution of founders, and cluster groups, to each of the breeding programs.

6.5 Discussion

Captive breeding programs can be a crucial part of species conservation (Soule et al., 1986; Kozfkay et al., 2008; Lacy, 2013). Maintaining a captive population that is genetically representative of the wild population at the time it was founded could be an important factor for ensuring the success of any future reintroductions (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Frankham, 2008; Jule et al., 2008). We showed that although levels of genetic diversity were very similar between wild, founder and captive okapi, there was considerable genetic differentiation between captive and the wild populations. In particular, genetic differentiation based on F_{ST} values, was much higher between the European and US captive populations than between any of the four wild sampling regions. Captive populations were also genetically differentiated from the wild populations, based on STRUCTURE results. Mitochondrial genetic diversity was greatly reduced in the founders and their captive descendants compared to the wild, with only a small proportion of the haplotypes present in the wild represented in captivity.

Significant genetic differentiation between wild okapi sampling regions (Table 6.1) confirms what was found in Chapter 4, where genetic differentiation was

detected between present-day wild okapi samples based on mtDNA and nuclear sequences. This information can be of use in in-situ conservation when determining priority areas, identifying barriers to movement, and planning translocations (Pennock & Dimmick, 1997). It can also be of use for identifying the wild origins of founders, potentially useful when planning future translocations from the wild into captivity. Mitochondrial DNA was not of use for assigning origin to founders, due to the ubiquity of haplotypes throughout the okapi's wild distribution (Chapter 4). The STRUCTURE analysis was also unable to assign founders to a wild sampling region, due to minimal genetic structure in the wild based on this particular analysis. The AMOVA analysis however, found strong support for a genetic grouping that included the founders of the captive population with the individuals sampled from sampling region two. AMOVA also confirmed the results from the STRUCTURE analysis that sampling region one is genetically differentiated from the other regions. This information is of use to any future translocations from the wild into captivity, as it would suggest that they should be sourced from areas outside of sampling region two, to represent wild genetic diversity from throughout the okapis range.

Levels of genetic diversity were similar between founder, captive and wild okapi population samples (Figure S6.6). Examples exist in the literature for other captive populations that have both reduced (Forstmeier et al., 2007; Muñoz-fuentes et al., 2008; Shen et al., 2009), and similar (Henry et al., 2009; Nsubuga et al., 2010; McGreevy et al., 2011) nuclear genetic diversity to their wild counterparts. High genetic diversity is important to maintain in captivity, as reduced genetic diversity may cause a decrease in population fitness, and ultimately extinction (Frankham et al., 2002). The captive okapi population is likely to have been able to maintain this high level of genetic diversity due to the species' well-managed studbook, which minimises inbreeding using MK strategies (Mean Kinship strategies; Rudnick & Lacy, 2007; Leus & Hofman, 2012). Genetic diversity based on mtDNA sequences was considerably reduced in captive samples. The importance of mtDNA diversity in captive populations is less well recognised than nuclear genetic diversity (Hedrick et al., 1997). Preserving mitochondrial genetic diversity may be of more importance in okapi

than in other species however, due to the presence of highly distinct and divergent mtDNA lineages in this species in the wild, some of which may represent greater than one million years of independent evolution (Chapter 4). The comparison of mtDNA diversity presented in the present study (Figures 6.1 and 6.2) gives an overview of both the genetic diversity the captive population started with, and an assessment of the breeding program's success in retaining genetic diversity. Future work should investigate if mtDNA could play a more important role in captive breeding management of species like okapi, which have high levels of mitochondrial genetic diversity in the wild.

Genetic diversity *per se* is not the only factor that is important in ex-situ management. Although genetic diversity may be similar between captive and wild populations, allele frequencies may be different, and captive populations may not be representative of the wild (Henry et al., 2009; Nsubuga et al., 2010; McGreevy et al., 2011). This may be of importance because the allele frequencies of a captive population may influence the success of any future reintroductions from captivity into the wild (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Frankham, 2008; Jule et al., 2008). Taken together, the STRUCTURE, AMOVA and mtDNA results demonstrate that the captive okapi population is not representative of the wild. This conclusion is based on, i) genetic differentiation of microsatellite loci, inferred from STRUCTURE analysis, AMOVA and F_{ST} values, and ii) a considerable reduction of mtDNA haplotypes in captivity, relative to wild okapi populations. We define 'not representative' by a detectable difference in allele frequencies populations, important because an increase of alleles in captivity that are rare in the wild may contribute to adaptation to captivity (Frankham, 2008) and consequently affect the success of any future reintroductions (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Jule et al., 2008). Other studies have shown that genetic structure can be introduced rapidly in a captive population due to a founder effect (Hu et al., 2007; Armstrong et al., 2011; Witzenberger & Hochkirch, 2013). There are, however, relatively few studies that use a dataset as comprehensive as ours, with representative samples from the wild, founders and captives (McGreevy et al., 2009; Shen et al., 2009; Gonçalves da Silva et al., 2010).

The cause of this genetic structure is important to ascertain, as an understanding of these population genetic processes can aid future captive breeding programs in obtaining a more representative sample of wild genetic diversity. Figure 6.4 demonstrates the unequal contribution of founders into captivity, particularly evident for two of the founders assigned to STRUCTURE cluster B (individuals 85 and 259; Figure 6.3). These results illustrate how captive populations are highly subject to the process of founder effect and genetic drift, exacerbated by unequal representation of founders, and leading to differentiated ex-situ subpopulations. This also illustrates how these processes can lead to a captive population that is not representative of the wild, despite a captive breeding program that is effectively maintaining high genetic diversity in captivity. It is important to note that the maintenance of high genetic diversity in this ex-situ population is a huge success considering the challenges that zoos can face in breeding okapi at all in captivity (Gijzen & Smet, 1974; Rabb, 1978; Bodmer & Rabb, 1992). This is also not a challenge that is unique to okapi (Snyder et al., 1996), and many captive populations have been shown to demonstrate considerable genetic structure (e.g. Guan et al., 2009; Henry et al., 2009; Armstrong et al., 2011; Witzemberger & Hochkirch, 2013). The captive US samples used in the present study may not be representative of the entire US captive population, as only four individuals have been sampled, however as a whole, the captive samples used in the present study have ancestry that can be linked back to 33 of the 34 individuals that founded the captive okapi population. Nonetheless, future studies would benefit from a more representative sample of the captive US population to investigate if the genetic differentiation between that population and the European one is as strong as indicated by our results.

Future management of the captive okapi population could ensure better genetic representation of the wild population, by focussing on trying to distribute founding genetic lineages throughout the global captive breeding programs. This action would need to be predicated on a genetic assessment of the captive breeding programs outside Europe, to investigate their genetic structure. Consideration also needs to be given to whether it makes good management

sense to disrupt the genetic differentiation currently in captivity. The extent to which captive sub-populations should be fragmented has been debated (e.g. Wang, 2004; Frankham, 2008; Leberg & Firmin, 2008). Both theory (Kimura & Crow, 1963; Robertson, 1964) and empirical studies (Frankham, 2008) however, seem to indicate that genetic structure in captivity helps to increase the fitness of that population when it is reintroduced to the wild. Nonetheless, the present study has shown that genetic differentiation in captivity is already considerably higher than in the wild. Genetic structure is therefore still likely to be able to be maintained despite increased gene-flow between captive breeding programs. The present study provides an important case study to help understand how these population genetic processes affect small, artificially bred populations.

The present study also identified a considerable loss of mtDNA in the captive okapi population relative to the wild. This loss is likely to be rapid due to mtDNA having an effective population size one-quarter that of the nuclear DNA and is therefore more susceptible to the effects of genetic drift (Hartl & Clark, 1997). If the remaining mtDNA haplotypes are not actively preserved in captivity in the future, their rapid loss is likely to continue. Okapi have recently been shown to be a species with a remarkable level of mitochondrial genetic diversity and divergence for an endangered species. Some of the mitochondrial genetic lineages appear to represent hundreds of thousands, or even millions of years of genetic diversification (Chapter 4). The loss of these lineages may therefore be particularly important to monitor in captivity, as they may each represent part of a long, and rich evolutionary history in this species.

6.6 Acknowledgements

We gratefully acknowledge the assistance of the *Institut Congolais pour la Conservation de la Nature* (ICCN), in particular Director J J Mapilanga for allowing the field collection and export of samples. We also thank all the museum curators that sent historic samples or allowed us to collect them. In particular, from the Chicago Field Museum: Bill Stanley and Lawrence Heaney (with the help of Keith Dobney, Aberdeen Univ. and Greger Larson, Durham Univ); from Copenhagen NHM: Hans J. Baagøe, Kristian Gregersen and Mogens Andersen; from Paris

NHM: Joséphine Lesur-Gebremariam; from the Royal Museum for Central Africa, Tervuren, Belgium: Wim Wendelen (with the help of Floris van der Sman). We thank Gilman International Conservation, the Okapi Conservation Project, the Frankfurt Zoological Society, the Wildlife Conservation Society, the Lukuru Foundation/TL2 project, the Zoological Society of London (ZSL) and the considerable number of people who assisted in various ways with sample collection. In particular, we would like to thank John Fataki Bolingo, Bryna Griffin, Terese Hart, Chrysostome Kaghoma, Luaison, John Lukas, Kambale Magloire, Ephrem Mpaka, Stuart Nixon, Linda Penfold, Elise Queslin, Alex Quinn, Rosemary Ruf and Ashley Vosper. This project was jointly funded by the UK Natural Environment Research Council (NERC) and ZSL (ZSL as NERC CASE industry partner and via an Erasmus Darwin Barlow Expedition grant). We also acknowledge the financial support provided by the US Fish and Wildlife Service, the UK's Darwin Initiative and the Mohamed bin Zayed Species Conservation Fund.

6.7 Supplementary

Figure S6.1 Map showing the sampling regions described in the present study. These sampling regions are used in the genetic structure analysis to compare relative differentiation in the wild and captivity. These regions are also described in Chapter 4.

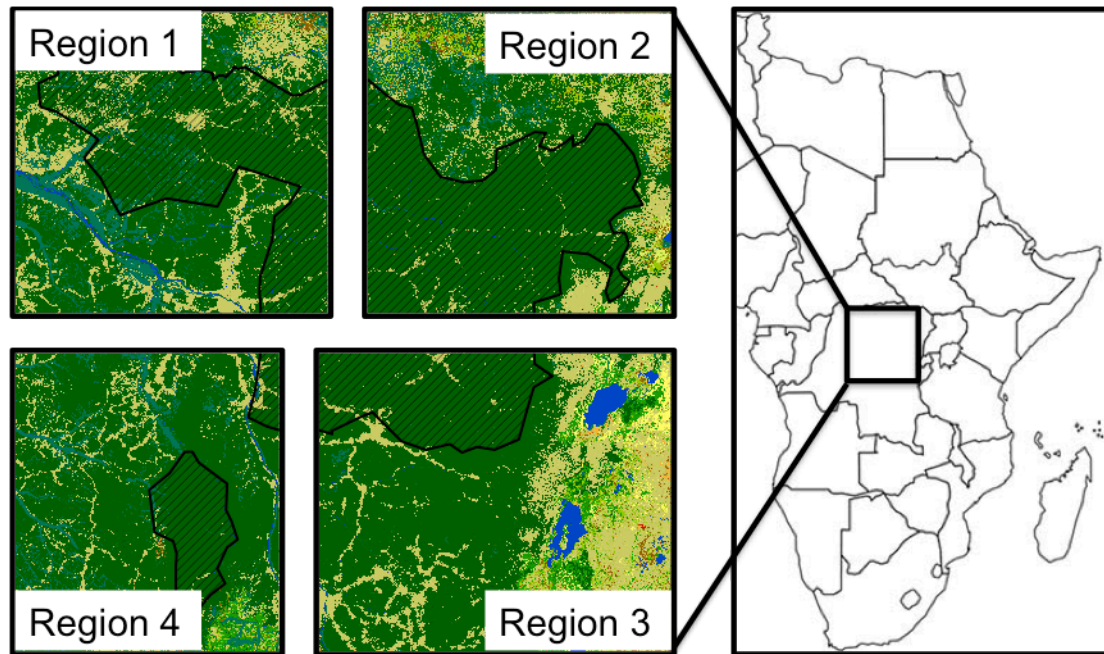


Figure S6.2 Set one, hypotheses one and two for the AMOVA analysis. AMOVA groups are shown in grey boxes, and AMOVA populations are shown with a dashed outline. The AMOVA statistics for the hypothesis with the best explanatory power (highest significant among group variation) are shown in bold and outlined in red.

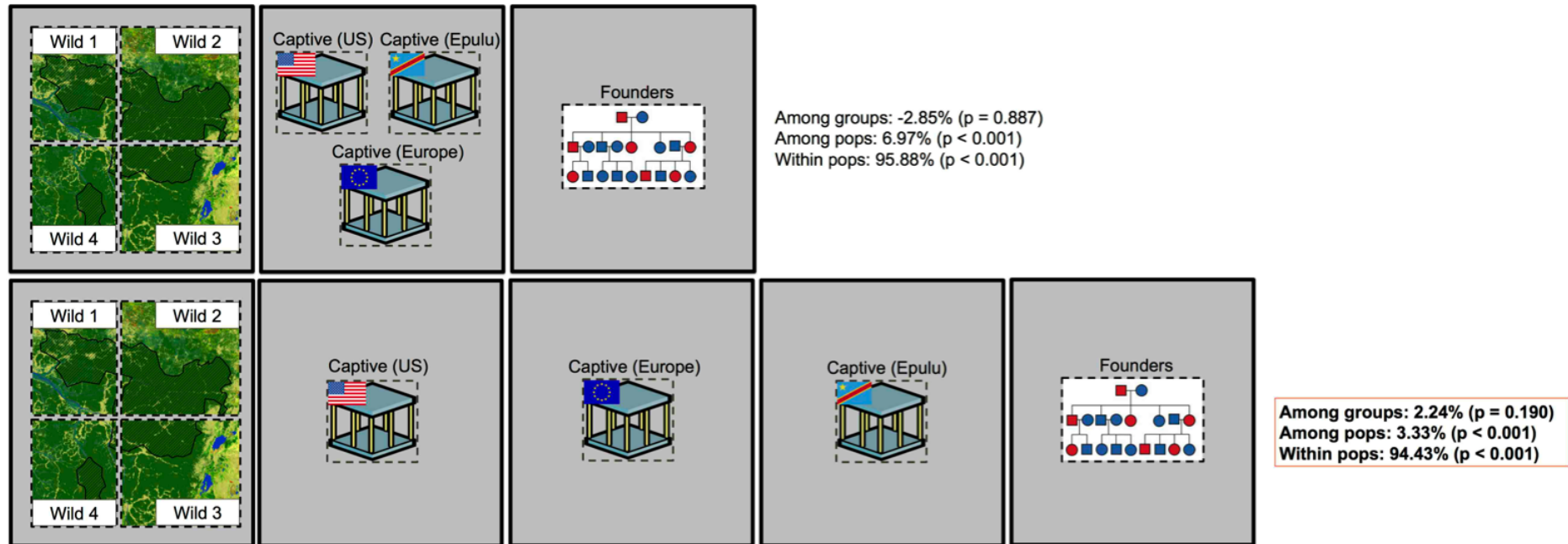


Figure S6.3 Set two, hypotheses three to six for the AMOVA analysis. AMOVA groups are shown in grey boxes, and AMOVA populations are shown with a dashed outline. The hypothesis with the best explanatory power (highest significant among group variation) are shown in bold and outlined in red.

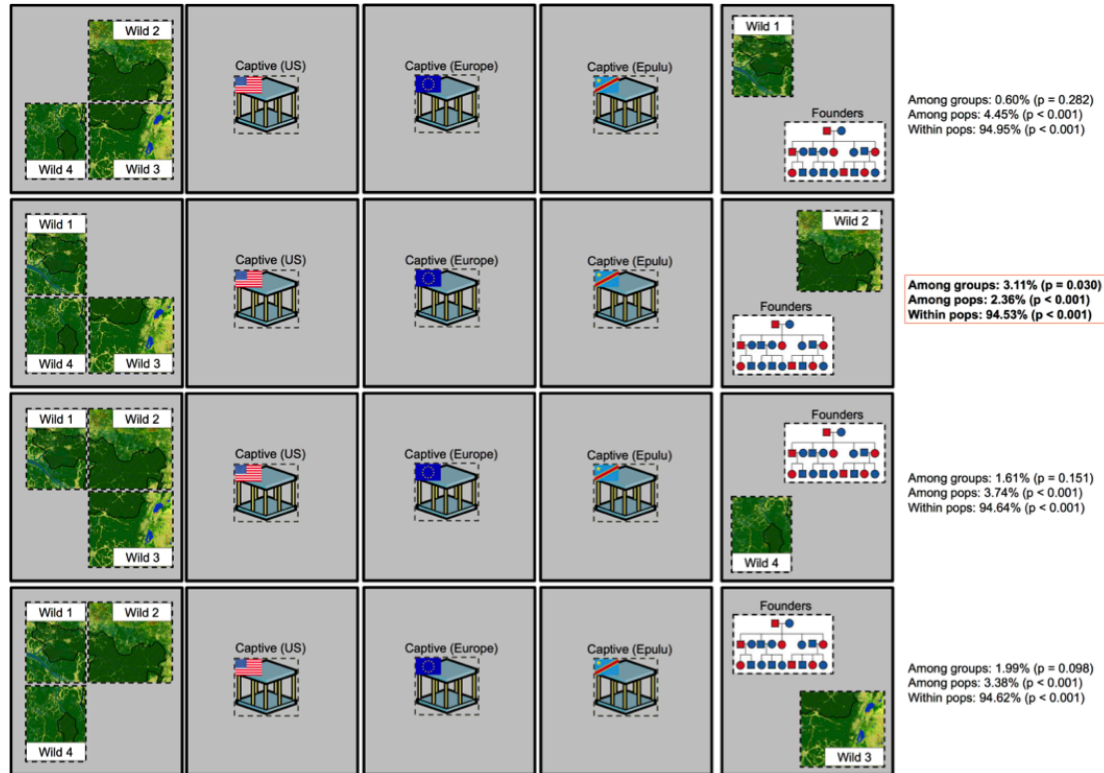


Figure S6.4 Set three, hypotheses seven and eight for the AMOVA analysis. AMOVA groups are shown in grey boxes, and AMOVA populations are shown with a dashed outline. The hypothesis with the best explanatory power (highest significant among group variation) are shown in bold and outlined in red.

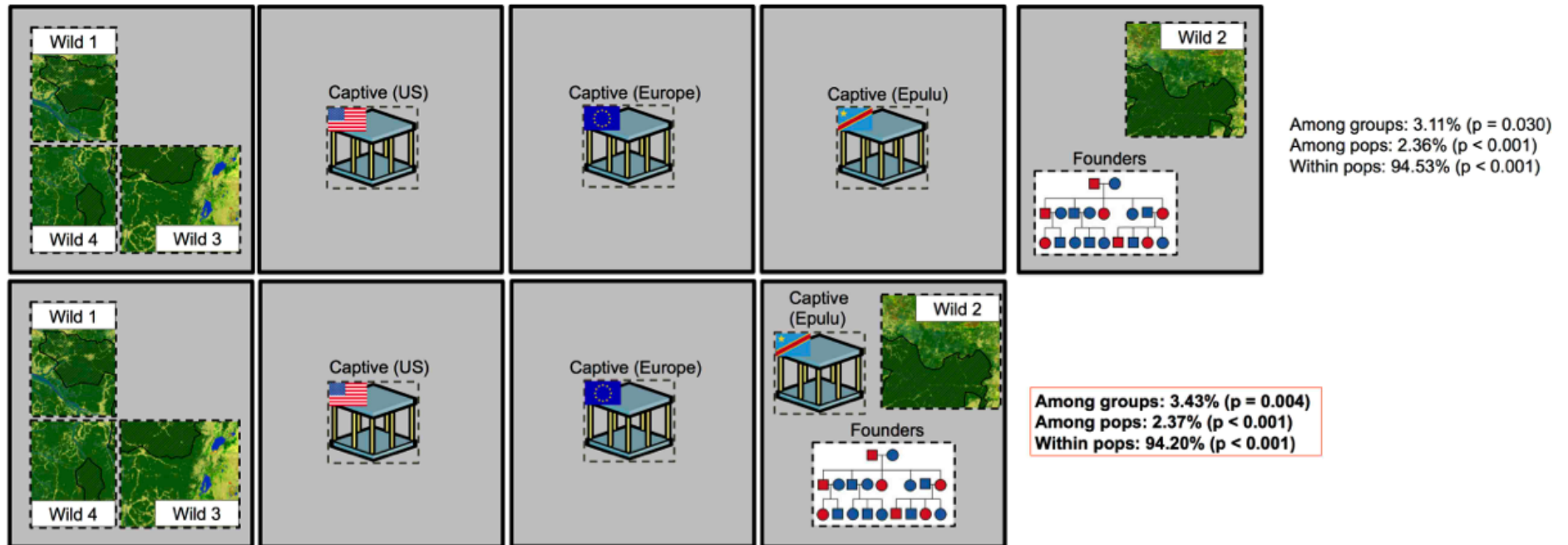


Figure S6.5 Set four, hypotheses nine to 13 for the AMOVA analysis. AMOVA groups are shown in grey boxes, and AMOVA populations are shown with a dashed outline. The hypothesis with the best explanatory power (highest significant among group variation) are shown in bold and outlined in red.

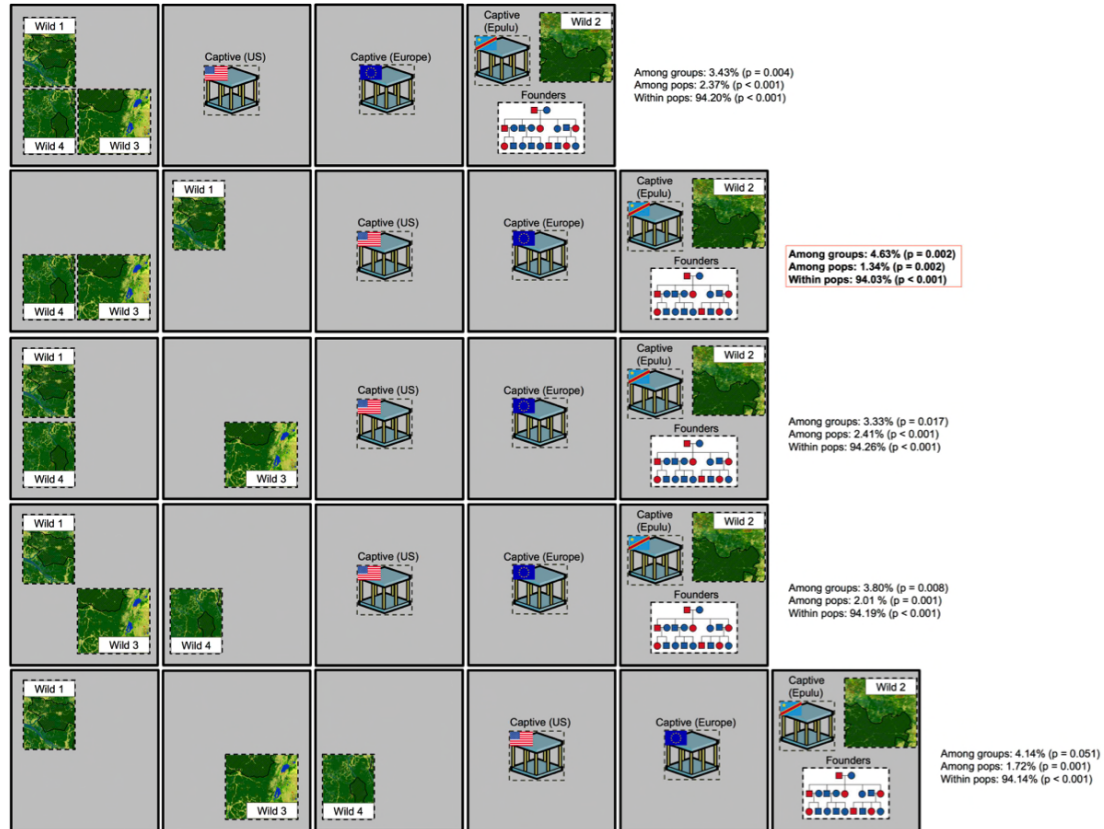


Figure S6.6 Boxplot of observed heterozygosity of founder, captive and wild okapi samples using microsatellite markers. Notches show 95% confidence intervals.

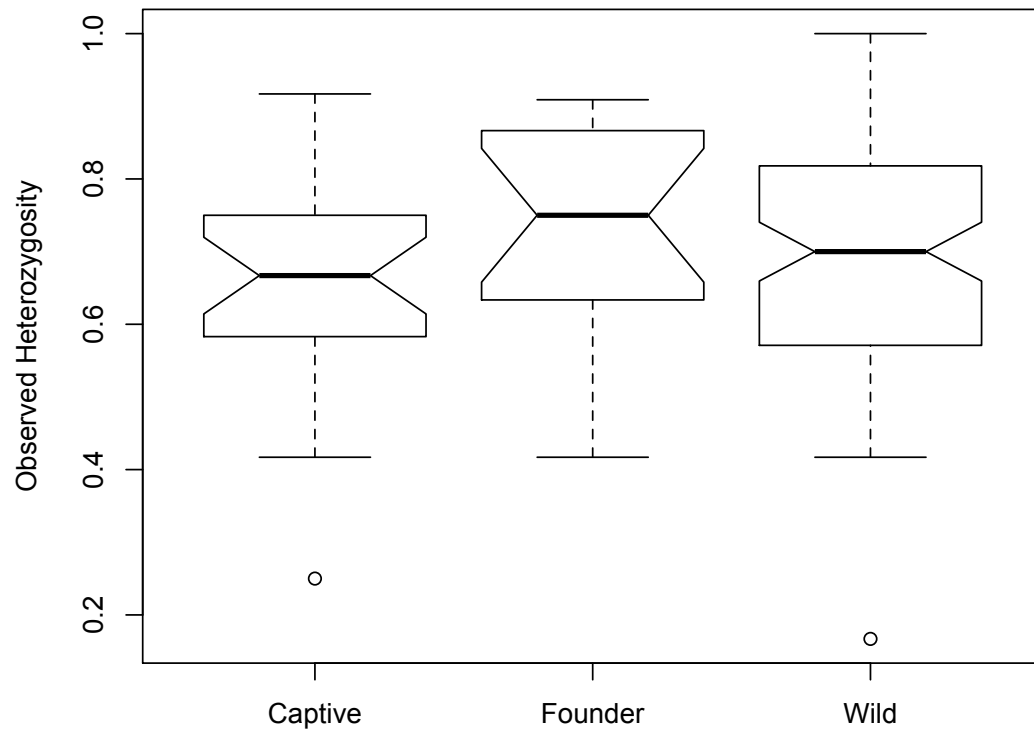
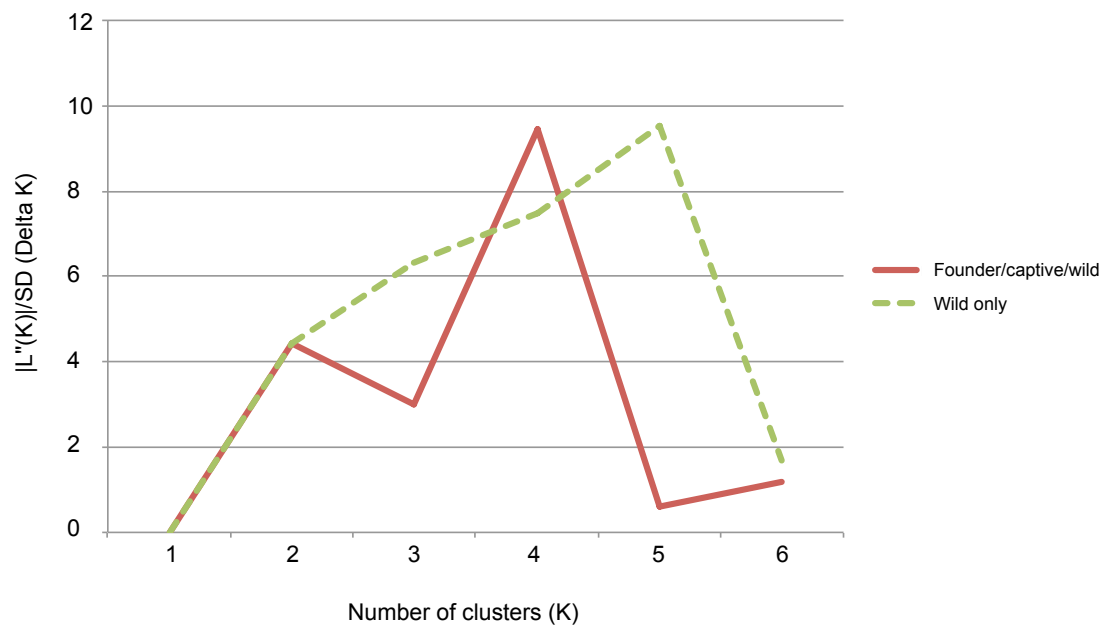


Figure S6.7 Plot to identify the most likely number of genetic clusters, using the program STRUCTURE v2.3.4 (Pritchard et al., 2000) and the method of Evanno et al. (2005). The highest value of delta K indicates the most likely number of clusters. This was carried out for all samples used for the microsatellite analysis (“Founder/captive/wild”), and for the wild samples only (“Wild only”).



7

CHAPTER SEVEN – General discussion



"Eternity is a long time, especially towards the end."

Woody Allen

CHAPTER 7 – General discussion

7.1 Aims

This thesis aimed to use genetic tools to help better describe the geographic range of okapi and their evolutionary diversity across that range; sociogenetic structure and dispersal characteristics of okapi populations (using the most densely populated okapi reserve remaining); and evaluate the captive population of okapi in comparison to their wild ancestors.

7.2 Completion of aims

We successfully developed a set of 13 primers (Chapter 2, Stanton et al. (2010)) that could amplify polymorphic microsatellite loci in non-invasively collected samples (Stanton et al., 2010). A subset of these primers was excluded from some analyses due to low PCR amplification success (Oka-02, -10 & -11, Chapter 5; Oka-11, Chapter 6). Low amplification success in these primers (in the full datasets compared to the partial dataset analysed in (Stanton et al., 2010)) was likely due to either: i) Genetic variation within the primer region present in individuals in the complete datasets, compared to the original dataset tested in; ii) A function of large variation in DNA concentration in samples in the complete datasets, leading to random amplification success; iii) the formation of hairpins; iv) duplexing of primers; v) variation in annealing strength across the primer; vi) A combination of some/all of the above. Low amplification success in these primers was unlikely to be related to PCR product length, as none of the primers with a low success rate amplified the longest PCR product.

Chapter 3 investigated if non-invasive genetic methods could help to confirm species identity from putative okapi fecal samples collected in the wild, and used the information to update on okapi distribution southwest of the Congo River. We showed that okapi definitely do occur on the southwest side of the Congo River, and presented a molecular methodology that could be used to identify putative okapi dung samples quickly and cheaply. This approach will be useful for determining distribution of other elusive species persisting at low densities in widely scattered localities. Unfortunately, this is likely to be the case for an

increasing number of the world's fauna (e.g. Woodroffe & Ginsberg, 1999; Li & Smith, 2005; Paviolo et al., 2008).

Chapter 4 investigated the evolutionary diversity of okapi across their range. We demonstrated that okapi are both genetically diverse *and* evolutionarily distinct, a combination that is perhaps unusual for an endangered species (e.g. Menotti-Raymond & O'Brien, 1993; Lu et al., 2001). We detected six, highly divergent mitochondrial genetic lineages in okapi, with some haplotypes distributed throughout the species' range. Overall, both mitochondrial and nuclear diversity showed a significant, non-random pattern of alleles between different sampling regions. These results are indicative of cyclical periods of population isolation and connectivity, leading to genetic differentiation and admixture respectively. The most recent population divergences events were estimated to have occurred ca. 200 kya, and mtDNA sequence divergences were estimated to have occurred at least 1.7 mya. This is consistent with repeated glaciation events during the Pleistocene (DeMenocal, 1995; DeMenocal, 2004; Maslin & Christensen, 2007). This information is important for okapi conservation because an understanding of genetic structure and evolutionary history of a species, and how this has been affected by geographic and demographic barriers, can influence the designation of management units and prioritise wild populations (Frankham et al., 2002). This information has wider influence than just on one species, as much of the fauna of the region will be subject to similar biogeographic processes.

In Chapter 5 we investigated the mating system, social structure and dispersal of okapi. This information is important for okapi conservation because an understanding of these aspects of ecology are required for effective conservation management, for example population viability analysis (Lacy, 1993; Schaller, 1993; Primack, 2000). We show that okapi appear to be solitary, genetically polygamous or promiscuous, and demonstrate male-biased dispersal. The results presented in this chapter also provide a methodological framework for future studies investigating these ecological questions in other rare, elusive species.

Finally, Chapter 6 investigated if the okapi founder and captive populations were/are genetically representative of the wild. We used the genetic data from Chapter 4 and generated new genetic data from wild okapi samples, to compare to that of a representative sample of individuals from captivity and the founders of the captive population. We show that the captive population has a similar extent of genetic diversity to the wild, but that that genetic diversity is not representative of what is found in the wild in terms of allele frequencies. We argue that this is of considerable importance because it may influence the success of reintroduction programs (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Frankham, 2008; Jule et al., 2008). The genetic diversity based on mtDNA was considerably reduced in captivity however. This loss of mitochondrial genetic diversity constitutes the loss of lineages that represent hundreds of thousands or millions of years of diversification (Chapter 4). The implications of this loss should be a focus for future research of captive breeding management.

7.3 Conservation recommendations

7.3.1 Wild population

Low genetic diversity and population genetic fragmentation are antagonistic to population viability (Frankham et al., 2002; Allendorf & Luikart, 2007), so populations with high genetic diversity and lack of population genetic fragmentation should ideally be prioritised. Also, populations that are genetically distinct should be prioritised due to their potential positive effect on the species' evolutionary potential (Moritz, 1994; Frankham et al., 2002; Stockwell et al., 2003; Grivet et al., 2008). In light of limited conservation funding in the DRC, we recommend that conservation management should prioritise the RFO and the Lomami NP (Chapter 3). The RFO should be prioritised based on its high levels of genetic diversity for both nuclear and mitochondrial loci (Chapter 4; Chapter 5; Chapter 6), and an apparent lack of population genetic fragmentation (Chapter 5). Also, the okapi in the RFO is currently the best protected and most studied of all the populations within the okapis range. Despite a rapid decline in the RFO, populations elsewhere in the range are thought to be declining even more rapidly (IUCN, 2013). The prioritisation of the Lomami NP is based on the estimate that

gene-flow is considerably lower across the Congo River than it is between populations on the same side (Chapter 4). Also, the okapi population on the southwest side of the Congo River showed the highest pairwise differentiation based on AMOVA and F_{ST} values for both nuclear and mitochondrial sequences. Taken together, these results indicate that the Lomami NP potentially contains a high level of unique genetic diversity that needs to be conserved.

Chapter 5 showed that okapi have a high movement and dispersal potential in the RFO, and that isolation by distance is operating at an extent that extends beyond the bounds of the reserve. This implies that a considerable level of immigration and emigration is occurring in this protected area. There is therefore currently no evidence for large-scale fragmentation, and restriction of gene-flow in the Ituri region for this species (Figure 5.1, Chapter 5). Due to their affinity for dense, isolated rainforest, okapi may be particularly susceptible to forest fragmentation (due to them being likely to avoid non-forested areas, and human disturbance). We recommend that as the DRC becomes increasingly deforested, forest corridors could provide a potentially useful way of maintaining viable okapi populations in this part of their range. The okapis demonstrated ability for large-scale movement suggests that forest corridors could facilitate gene-flow between forest fragments. A more complete genetic characterisation of the populations surrounding the RFO would be required in order to propose corridor locations.

7.3.2 Captive population

Chapter 6 found no significant difference in genetic diversity (allelic richness and heterozygosity) between captive and wild okapi, highlighting the remarkable success of the ex-situ okapi breeding program in maintaining high genetic diversity in this artificially-bred, bottlenecked population. Captive US and European populations were shown to be highly genetically differentiated from each other, and from the wild population however. This is because, despite the maintenance of high genetic diversity, allele frequencies differ considerably between these populations. This genetic differentiation poses some interesting management questions. As we argue in Chapter 6, a difference in allele

frequencies between in- and ex-situ populations could lead to difficulties for future reintroductions, due to the potential increase of deleterious alleles in the population (Griffith et al., 1989; Fischer & Lindenmayer, 2000; Wolf et al., 2002; Frankham, 2008; Jule et al., 2008). Reconstructing specific allele frequencies in an established captive population would at best be highly challenging, and at worst impossible. Attempting to do this would imply genetic characterisation of the global captive population and extensive translocations of individuals between international breeding programs, an infeasible task, due to logistical difficulties and cost. Also, maintaining population structure in captivity is often an intentional management decision. Maintaining genetic structure in a captive population has been shown to increase the fitness of that captive population when it is reintroduced to the wild (Kimura & Crow, 1963; Robertson, 1964; Frankham, 2008). A compromise is perhaps needed whereby gene-flow is increased slightly between captive breeding programs to ensure founding lineages are more evenly spread throughout the programs (at present there is a low level of gene-flow between them (Leus & Hofman, 2012)). Genetic structure of okapi in captivity is currently up to three times higher than in the wild (based on F_{ST} values), so gene-flow could be increased in captivity while still maintaining more genetic structure than in the wild.

7.4 Role of genetics in the future of okapi conservation

As demonstrated in this PhD thesis, genetics has already played an important part in okapi conservation. This role of genetics must be continued to ensure effective okapi conservation in the future.

In Chapter 3, we developed and tested two pairs of mtDNA primers that could verify if putative okapi dung samples were, in fact, okapi. The main benefit of these primers was that they could positively identify okapi, with high identification success, without the need for DNA sequencing. This considerably reduces the time and cost of species identification, and could facilitate future wildlife surveys carried out *within* the DRC that are informed by genetics. This should not be underestimated, as surveys of this nature would help build capacity in Congolese universities, and advocate okapi conservation within the

DRC. We identified some key areas that future wildlife surveys could focus on (Figure 3.1, Chapter 3), using these primers, and based on recommendations from the 2013 okapi conservation workshop (Quinn et al., 2013). These are regions where there are anecdotal reports of okapi presence, but no confirmed cases.

Figure 7.3 Abandoned hunting camp found in the Lomami NP (photograph: Dave Stanton)



Chapter 4 developed a set of 12 EPIC loci and 5 pairs of mitochondrial primers for okapi, and demonstrated their utility for non-invasive studies. Despite this relatively low number of nuclear markers, high heterozygosities (max H_e 0.459) give them utility for discriminating among populations (significantly detected an F_{ST} of 0.115 between regions 1 and 2, with $n = 5$ and $n = 14$ respectively). Ecological studies of non-model organisms using SNPs are now becoming commonplace (Garvin et al., 2010; Davey et al., 2011; Haynes & Latch, 2012; Puritz et al., 2012), however, challenges still exist in using large numbers of SNPs in non-invasively collected samples. The reliability of SNPs when used in low quality samples is frequently cited as one of the reasons why SNPs have been

predicted to surpass microsatellites as the marker of choice for ecological studies (Aitken et al., 2004; Creer, 2007; Caballero et al., 2008). However, a number of the studies that have used these markers in non-invasive samples still use far fewer markers than comparative studies that use tissue or blood (e.g. Morin & McCarthy, 2007; Fabbri et al., 2012; Barbosa et al., 2013). The main reason for this is the time and cost of sequencing hundreds of loci in tens, or hundreds, of samples. However, O'Neill et al. (2013) recently developed a pipeline (NextAllele) and demonstrated an approach whereby multiple separate individuals are labelled and combined in a way that allows ~100 loci to be sequenced in ~100 individuals in a single next generation sequencing run. Using this approach in combination with that adopted here, it may finally be feasible to use the full potential of next-generation sequencing in conservation studies that are only able to use noninvasively collected samples. This approach will give added power to be able to detect demographic parameters, such as those investigated in Chapter 4, and investigate even more complex models. Determining *a priori* if 100 loci will be enough to accurately reconstruct demographic history is not simple, as this will depend not only on the number of loci, but also the power of those loci (i.e. their polymorphism) and the magnitude of the effect of the demographic history on the species in question's genome (Wakeley, 2008; Lopes et al., 2009; Csilléry et al., 2010).

Chapter 5 described the genetic structure (or lack of it) in the RFO. Future studies should extend this investigation of genetic structure to include the whole Ituri landscape (Figure 5.1, Chapter 5), and Maiko NP. This would allow the effect of landscape features on genetic structure (i.e. landscape genetics; Manel et al., 2003), and the extent of gene-flow between adjacent protected areas to be investigated. This would help us understand the role of landscape in structuring genetic diversity in okapi, and therefore be able to consider okapi in the design of any future protected areas and wildlife corridors.

Figure 7.1. Okapi skin drums found in a village adjacent to the Rubi-Tele hunting reserve (Figure 4.1; photograph: John Hart).



In Chapter 6 we carry out a genetic comparison of wild, founder and captive okapi, and show that despite high genetic diversity, captive okapi are not genetically representative of the wild. This study focuses on okapi individuals from the European captive population. Future genetic work should focus on other captive breeding programs outside of Europe. We also show that okapi genetic diversity, based on mtDNA is considerably reduced in captivity compared to the wild. The importance of mtDNA is less well recognised than nuclear DNA for measuring diversity in captivity (Hedrick et al., 1997). For a species like okapi however, with mtDNA lineages that potentially represent greater than 1.7 million years of divergence, mtDNA diversity may be more important to preserve. Future research should focus on better understanding the importance of mtDNA in captive breeding management, and develop tools for predicting future loss of mtDNA in captivity. We have gone some way towards these aims in Appendix 2, where we present a preliminary description of a computer package that investigates haplotype survival over time in small isolated populations, using a

stochastic model of genetic drift. This work will be continued, and elaborated on in the future.

Figure 7.2. Okapi skin found in a village adjacent to the Rubi-Tele hunting reserve (Figure 4.1; photograph: John Hart).



7.5 Conservation reality on the ground in the DRC

There are a large number of factors that complicate effective conservation in the protected areas in the DRC. There are many armed militia groups active within the protected areas, including the ADF (Allied Democratic Forces), the LRA (Lord's Resistance Army), ADF-NALU (ADF – National Army for the Liberation of Uganda), Mai-Mai (local self-defence militia, formed on an ad-hoc basis), M23, FDLR (Democratic Forces for the Liberation of Rwanda) and SIMBA-Mai Mai (simba is the Swahili for lion) (Mapilanga, 2013). In addition to the armed groups there are also a large number of armed local and international poachers present within all the protected areas (Figure 7.3). The illegal extraction of natural resources is ubiquitous throughout the DRC's protected areas, including mineral and exploitation, and charcoal burning. Also, the invasion of protected areas by villages and settlements, deforestation for pasture land, and camps utilised by

the poachers and militia is increasing (Mapilanga, 2013). This is thought to lead to an increase in poaching, including okapi (Figures 7.1 & 7.2), although details of the extent of okapi poaching is unknown. The DRC's conservation agency, the ICCN (Institut Congolais pour la Conservation de la Nature) are responsible for attempting to deal with these issues. This is done with a very limited budget (80% of which comes from international NGOs), leaving a shortage for patrols and equipment. Due to the above conflicts, park rangers spend a lot of their time engaged in open fighting rather than being able to focus on wildlife surveys. More than 350 rangers were killed on duty in the last five years (Mapilanga, 2013) (Figure 7.4). These conflicts, and the associated bushmeat hunting, constitute the greatest threat to the persistence of okapi.

Figure 7.4 ICCN ranger training (photograph: Jean Joseph Mapilanga)



The DRC contains some of the greatest biodiversity in the world, however, much of this biodiversity is under threat from extinction (IUCN, 2013). Okapi are a flagship species for the DRC and its biodiversity, appearing on the logo of the ICCN, and on the bank notes for the country's currency. The results presented in this thesis can *aid* okapi conservation efforts, but for long-term conservation

efforts to be effective there will need to be a fundamental change in the 'reality on the ground' in the DRC.

A1

APPENDIX ONE – Computer program concept



“Nibbled to death by an okapi, nibbled to death by an okapi...”

The Damned

APPENDIX ONE - MTfate: A program for modelling the fate of mitochondrial DNA lineages in ex-situ populations (concept and preliminary results)

A1.1 Abstract

Mitochondrial DNA is used much less frequently than nuclear DNA in captive breeding management. The maintenance of genetic diversity in this molecular marker is usually considered of limited use in captivity, however, this assumption has never been justified. For some species with high levels of mitochondrial genetic diversity in the wild, the maintenance of this genetic diversity in captivity may allow for a better representation of a species' evolutionary history. The okapi (*Okapia johnstoni*) is an endangered giraffid, with a well-established captive breeding program, and high levels of mitochondrial genetic diversity in the wild. Some of these mitochondrial genetic lineages may represent over 1.7 million years of genetic diversification. We develop the concept, and test a preliminary version, of a computer program, MTfate, that models genetic drift in captive populations. We predict future changes in mitochondrial genetic diversity of the captive okapi population and show that mitochondrial genetic diversity may be entirely eliminated in the near future if management action is not taken. We argue that mitochondrial DNA should be used wherever possible, in combination with pedigree information, as a means of maintaining captive populations that better represent the evolutionary history of their wild counterparts.

A1.2 Introduction

Mitochondrial DNA has been used much less frequently than nuclear markers in studies of captive breeding programs, probably due to its perceived limited usefulness for this purpose (Russello et al., 2007; Witzemberger & Hochkirch, 2011; Benavides et al., 2012). When it has been used, analysis has often been limited to a simple measure of genetic diversity (Gautschi et al., 2003; Muñoz-fuentes et al., 2008; McGreevy et al., 2009; Lesobre et al., 2010; Khan et al., 2011;

McGreevy et al., 2011). It is certainly true that in certain situations the utility of mitochondrial DNA may be limited due to low genetic diversity in captive and/or wild populations (Hedrick et al., 1997; Hedrick & Fredrickson, 2008; Henry et al., 2009; Lesobre et al., 2010). However, we argue that it may be greatly overlooked in others. In fact, it is due to the unique inheritance properties of mtDNA that suggests it can provide information about ancestral demography that is not available when using nuclear markers alone (Avisé et al., 1984; Avisé et al., 1987). MtDNA may therefore provide an effective means of maintaining an ex-situ population that is more representative of the wild than when using nuclear markers alone (Helsen et al In Prep).

This study presents the concept of a program, MTfate, written VBA (Microsoft), which stochastically models genetic drift and the associated rate of extinction of mitochondrial genetic lineages. This study also aims to predict future changes in mitochondrial genetic diversity of okapi, and discusses the utility of mtDNA in conservation management in okapi. We extend this to infer the utility of mtDNA to ex-situ conservation management in general, and investigate if this marker should be used more, in combination with pedigree information as a means of maintaining captive populations that are better representatives of the corresponding wild populations.

A1.3 Methods

A1.3.1 Model

MTfate is a program written in VBA (Microsoft) that models the survivorship of mtDNA haplotypes in a captive population. MTfate v1.0 uses a haploid, Wright-Fisher model, with non-overlapping generations, and the results in the present study are based on this version. Since the present study was carried out, MTfate has been updated to v1.1. This version allows for overlapping generations (determined by a power function distribution), as well as a number of script optimisations that make the program considerably faster.

A1.3.2 Testing

MTfate was tested on the captive okapi population. Starting haplotype frequencies were taken from the frequencies of CR haplotypes observed in the combined captive and founder population datasets from Chapter 6. To investigate the effect of population structure on the survivorship of haplotypes, simulations were run assuming i) all captive okapi belong to a panmictic population of size 173 individuals, ii) two panmictic populations of size 86 individuals each, and iii) four panmictic populations of size 43 individuals each. 200 iterations were run for each of the population simulations. Mean and standard deviation (SD) of number of haplotypes remaining after 100 generations was plotted as a function of number of iterations (for the one, two and four population models). This was to investigate if the number of iterations carried out was sufficient. The proportion of extinct haplotypes across all iterations was also plotted as a function of generation number. The proportion was weighted to account for the fact that each population cannot have less than one haplotype, which we classed as 100% of haplotypes extinct.

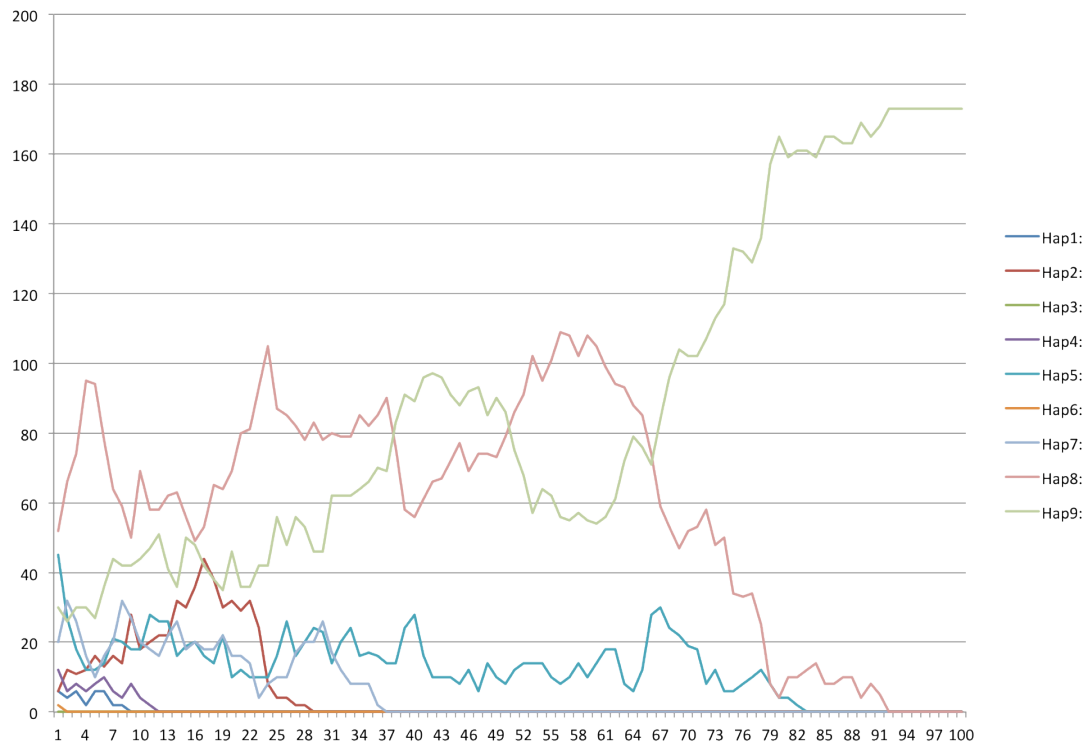
A1.4 Results

The survivorship of okapi mtDNA haplotypes in the captive population was estimated using stochastic models. The mean and SD for number of haplotypes remaining had plateaued by 200 simulations for the one-, two- and four-population models indicating that the number of iterations carried out was sufficient (Figures SA1-3). An example of a single iteration of the model is shown in Figure A1.

For the one-population model (Figure A2[A]), >50% of the haplotypes were lost from the population after 16 generations (approximately 80 years) and 92.3% of haplotypes has been lost from the population after 100 generations. For the two-population model (Figure A2[B]), >50% of the haplotypes were lost from the meta-population after 16 generations, and 84.5% of haplotypes had been lost from the meta-population after 100 generations. For the four-population model (Figure A2[C]), >50% of the haplotypes were lost from the meta-population after

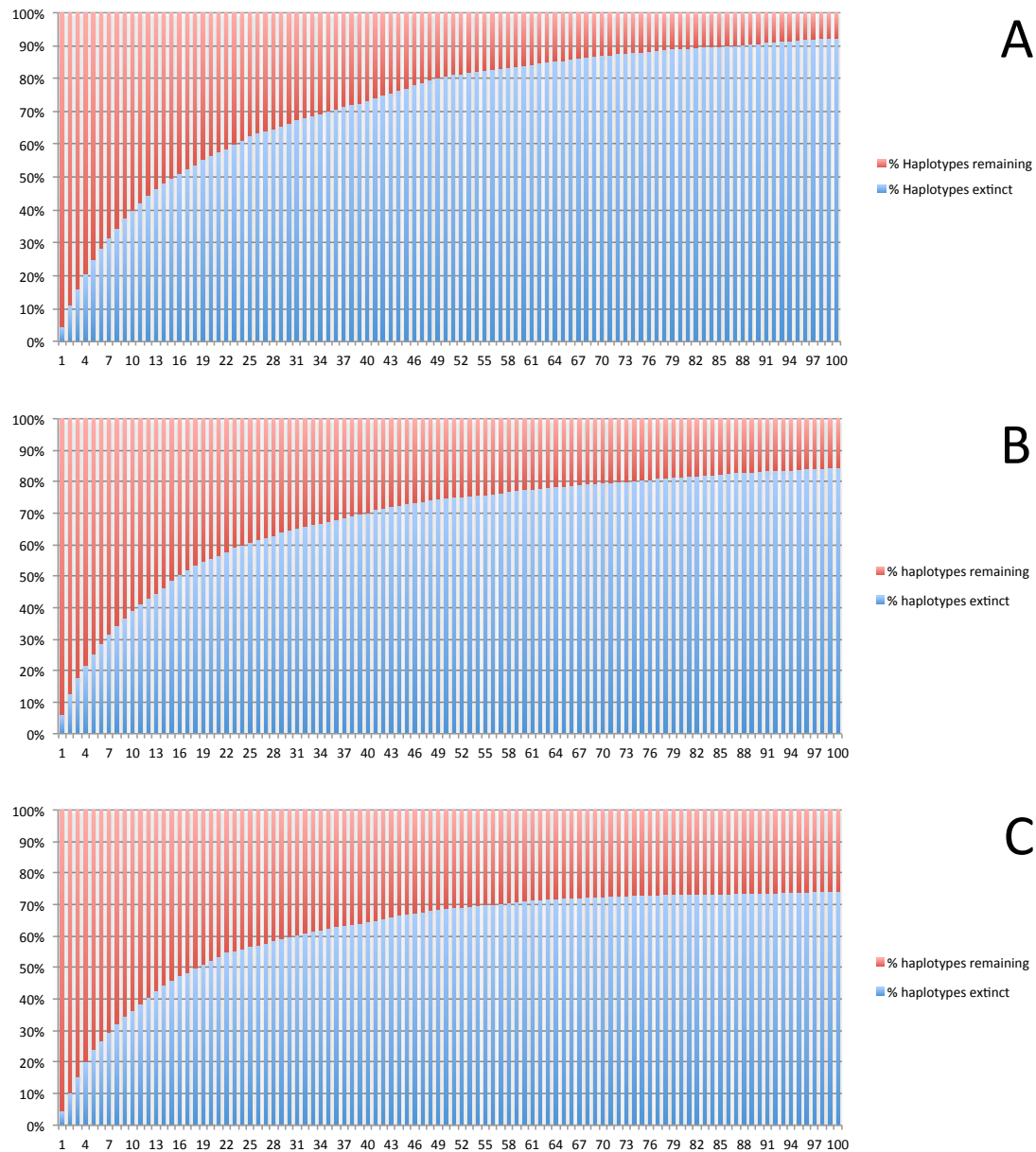
19 generations, and 74.1% of haplotypes had been lost from the meta-population after 100 generations.

Figure A1. Example of a single iteration of the program MTfate. Frequency of haplotype in the population is indicated on the y-axis.



Mean number of haplotypes remaining after 100 generations for the one population model was 1.75 (SD 0.64) and probability of fixation for a single haplotype was 0.36 (SD 0.48). Mean number of haplotypes remaining after 100 generations for the two-population model was 2.20 (SD 0.63; mean of 1.25 [SD 0.44] in each population) and probability of fixation for a single haplotype was 0.10 (SD 0.29; mean of 0.76 [SD 0.43] in each population). Mean number of haplotypes remaining after 100 generations for the four-population model was 3.07 (SD 0.73; mean 1.03 [SD 0.16] in each population) and mean probability of fixation for a single haplotype was 0.02 (SD 0.12; mean 0.97 [SD 0.16] in each population).

Figure A2. Predicted proportion of alleles lost from the okapi captive population over 100 generations for the one (A), two (B) and four (C) population models. The models assume a haploid Wright-Fisher model of drift, with the one population model assuming a panmictic population, and the two and four population models assuming island models with no gene-flow.



A1.5 Discussion

The present study presents a new program, MTfate, and used this program to carry out stochastic simulations to investigate the potential future rate of haplotype extinction in a captive population of okapi. All models used an initial number of haplotypes of nine, based on our conservative estimate of the number

of different haplotypes present in the combined founder and captive populations (true for both the CR and concatenated sequences). We demonstrate that >50% of mitochondrial haplotypes are likely to be lost from the captive okapi population over the next 20 generations (Figure A2). This is true for the one-, two- and four-population models. The loss of mitochondrial genetic diversity was slowed by including more genetically isolated populations to the model. This is predicted by population genetic theory, and well documented in practice (see Margan et al. (1998)). However, even with the four-population model, 74.1% of haplotypes had been lost from the meta-population after 100 generations. The mean number of haplotypes remaining of between 1.75 and 3.07 represents a 82.8% - 93.2% decrease in genetic diversity from the wild for the concatenated haplotypes and a 93.1 – 97.5% decrease in genetic diversity from the wild for the CR haplotypes (adjusting so that fixation for a single haplotype represents a 100% loss in genetic diversity).

Solutions to the problem of maintaining genetic diversity in captive populations are likely to be more effective the simpler they are to implement. Methods exist to choose founders that are representative of the wild, however this is not always practical, or founders have often already been chosen (Hedrick et al., 1997; Russello et al., 2007; Hedrick & Fredrickson, 2008; Ivy et al., 2009). Ex-situ management could use a program like STRUCTURE to try to match captive population allele frequencies to the wild (Miller et al., 2010). However, this requires extensive sampling in the wild. Even considering the present studies' extensive sampling effort, certain parts of the range remain under-represented (e.g. sampling region 3, Figure 4.1, Chapter 4), and so genuinely matching allele frequencies to the wild would be impossible. This is not a problem restricted to okapi. For example, the recent study of Witzenberger and Hochkirch (2013) provides information on the genetic structure of the captive population of the sand cat (*Felis margarita*). Witzenberger and Hochkirch (2013) identified three distinct genetic clusters in captivity, which they attribute to different founder lineages. This information is clearly important for the ex-situ management of this species. However, without genetic information from the wild, it is difficult to

know how these genetic clusters correspond to the wild, and therefore how the information is best implemented in captive management.

Ex-situ management currently uses pedigrees to estimate mean kinships between individuals, and give recommendations of which animals to breed and with whom (Lacy et al., 2012). Genetics may already be used to inform these recommendations, and guidelines exist for incorporating it into management strategies (Fienieg & Galbusera, 2013). By including mtDNA genetic diversity into this management, ex-situ management could potentially ensure the preservation of individuals that represent distinct evolutionary lineages. In the same way that genetic information is currently implemented in captive management, mtDNA information would be weighted so that it is taken into partial consideration for breeding recommendations along with other factors such as relatedness estimated from pedigrees (Lacy et al., 2012; Fienieg & Galbusera, 2013). The extent of this weighting would depend on the species in question, and would have to vary based on the priorities of the ex-situ management plan in question. This mechanism may provide a simple yet hitherto largely overlooked means of ex-situ management maintaining captive populations that are more genetically representative of the wild. This strategy has the benefit that it does not necessarily require samples from the wild. In addition, mtDNA can be used in combination with pedigrees to predict mtDNA haplotypes throughout the pedigree (Głazewska et al., 2013), meaning only a subset of founders and/or captive individuals need to be sequenced.

The captive okapi population's mitochondrial genetic diversity may be entirely eliminated in the near future if management action is not taken. Actively preserving mitochondrial DNA haplotypes would have the effect of maintaining genetic diversity that is more representative of this species' rich and diverse evolutionary history. This is particularly important for okapi considering its recent reclassification to "Endangered" on the IUCN Red List (Mallon et al., 2013). The relative values of mitochondrial and nuclear genetic diversity in captivity is debatable (Helsen et al In Prep). However, if it is determined that ex-situ management should take mtDNA into account, once fully developed, MTfate

will provide a tool to make this possible. The version of MTfate used in the present study (version 1.0) is an over-simplified representation of the real-life situation. Version 1.1 (*not* used in the present study) also includes overlapping generations. Future versions will focus on including gene-flow between captive sub-populations, model changes in population size, and model variations in birth-death rate between individuals.

A1.6 Acknowledgements

Many thanks to Mario for streamlining (read: re-writing it and making it ~100X faster...) the code. I look forward to finishing this work off with you, and your tuition with learning how to code VBA like a pro!

A1.7 Supplementary

Figure SA1. Mean and SD of number of haplotypes remaining over 100 generations, as a function of number of iterations for the one population model. This model simulated okapi mitochondrial DNA lineage survival in captivity over a period of 100 generations. The model was run in VBA (Microsoft), and used a haploid Wright-Fisher model of genetic drift.

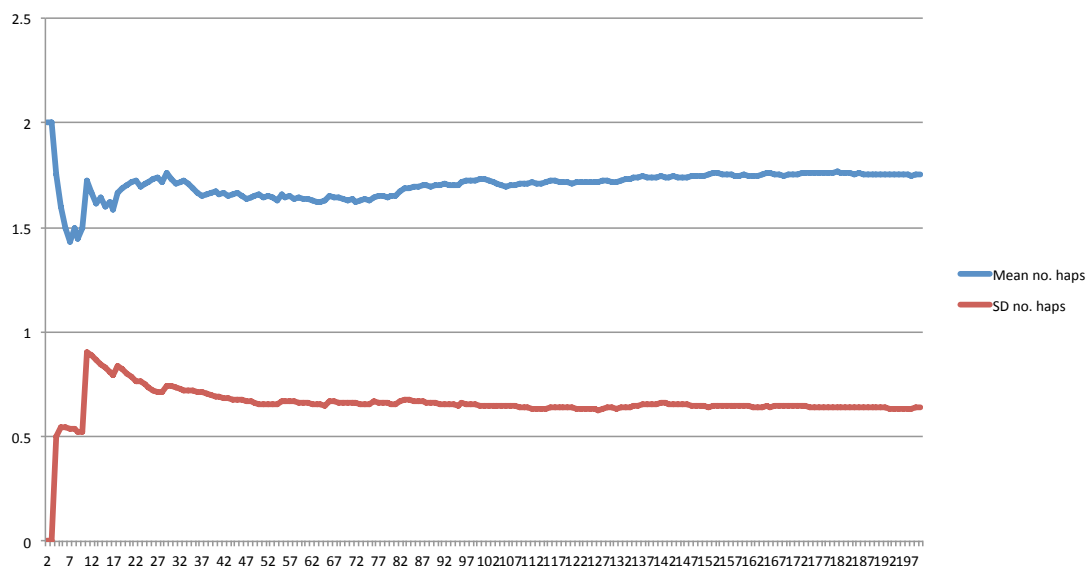


Figure SA2. Mean and SD of number of haplotypes remaining over 100 generations, as a function of number of iterations for the two population model. This model simulated okapi mitochondrial DNA lineage survival in captivity over a period of 100 generations. The model was run in VBA (Microsoft), and used a haploid Wright-Fisher model of genetic drift.

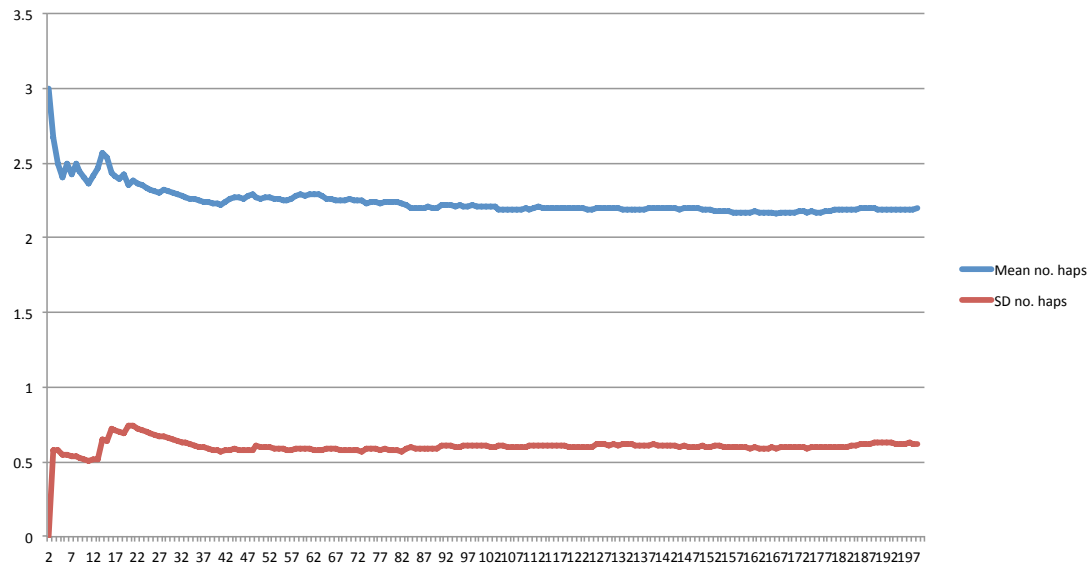
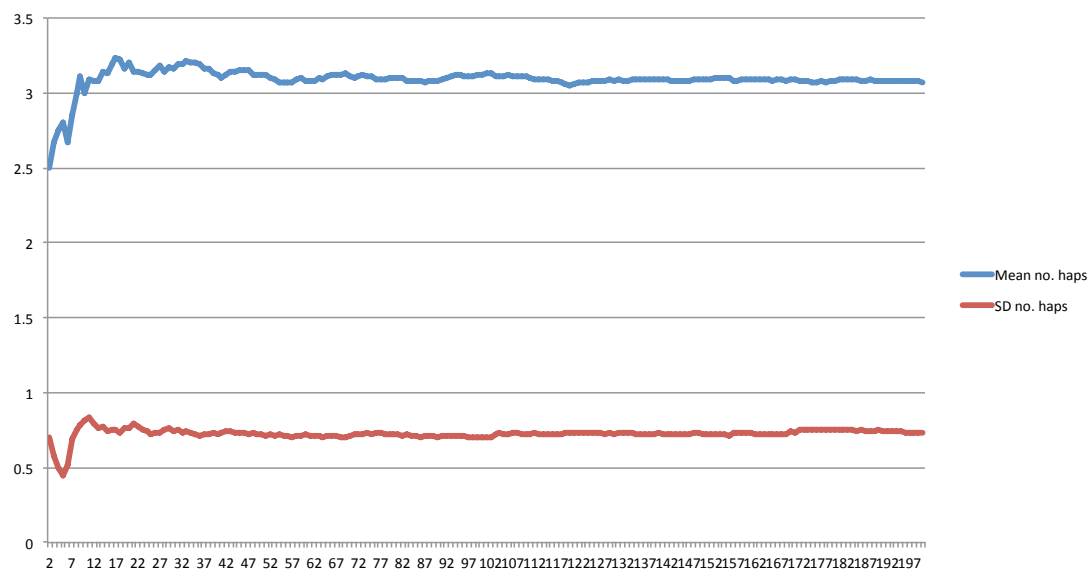


Figure SA3. Mean and SD of number of haplotypes remaining over 100 generations, as a function of number of iterations for the four population model. This model simulated okapi mitochondrial DNA lineage survival in captivity over a period of 100 generations. The model was run in VBA (Microsoft), and used a haploid Wright-Fisher model of genetic drift.



BIBLIOGRAPHY

- Aitken, N., Smith, S., Schwarz, C. & Morin, P.a. (2004) Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. *Molecular Ecology*, **13**, 1423-1431.
- Allendorf, F.W. & Luikart, G. (2007) *Conservation and the genetics of populations*, Blackwell Publishing, Main Street, Malden, MA.
- Alpers, D.L., Van Vuuren, B.J., Arctander, P. & Robinson, T.J. (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, **13**, 1771-1784.
- Anka, Z., Séranne, M., Lopez, M., Scheck-Wenderoth, M. & Savoye, B. (2009) The long-term evolution of the Congo deep-sea fan: A basin-wide view of the interaction between a giant submarine fan and a mature passive margin (ZaiAngo project). *Tectonophysics*, **470**, 42-56.
- Anuradha Reddy, P., Kumaraguru, A., Bhagavatula, J., Gour, D.S., Bhavanishankar, M., Shekhar Sarkar, M., Harika, K., Hussain, S.M. & Shivaji, S. (2012) Tiger presence in a hitherto unsurveyed jungle of India—the Sathyamangalam forests. *Conservation Genetics*, **13**, 779-787.
- Apio, A., Kabasa, J.D., Ketmaier, V., Schröder, C., Plath, M. & Tiedemann, R. (2010) Female philopatry and male dispersal in a cryptic, bush-dwelling antelope: a combined molecular and behavioural approach. *Journal of Zoology*, **280**, 213-220.
- Archie, E.A. & Chiyo, P.I. (2012) Elephant behaviour and conservation: social relationships, the effects of poaching, and genetic tools for management. *Molecular Ecology*, **21**, 765-778.
- Arctander, P., Johansen, C. & Coutellec-Vreto, M. (1999) Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution*, **16**, 1724-1739.
- Armstrong, E., Leizagoyen, C., Martínez, A.M., González, S., Delgado, J.V. & Postiglioni, A. (2011) Genetic structure analysis of a highly inbred captive population of the African antelope *Addax nasomaculatus*. Conservation and management implications. *Zoo Biology*, **30**, 399-411.
- Armstrong, P.R., Chan, K.M.A., Daily, G.C., Ehrlich, P.R., Kremen, C., Ricketts, T.H. & Sanjayan, M.A. (2007) Ecosystem-service science and the way forward for conservation. *Conservation Biology*, **21**, 1383-1384.
- Avise, J.C. (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371-379.
- Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.
- Avise, J.C., Neigel, J.E. & Arnold, J. (1984) Demographic Influences on Mitochondrial DNA Lineage Survivorship in Animal Populations. *Journal of Molecular Evolution*, **20**, 99-105.
- Baguette, M. & Van Dyck, H. (2007) Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. **22**, 1117-1129.
- Ballard, J.W.O. & Michael, C.W. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729-744.

- Barbosa, S., Pauperio, J., Searle, J.B. & Alves, P.C. (2013) Genetic identification of Iberian rodent species using both mitochondrial and nuclear loci: application to noninvasive sampling. *Molecular Ecology Resources*, **13**, 43-56.
- Baumung, R. & Sölkner, J. (2003) Pedigree and marker information requirements to monitor genetic variability. *Genetics Selection Evolution*, **35**, 369-383.
- Beaumont, M.A., Zhang, W. & Balding, D.J. (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025-2035.
- Beheregaray, L.B. (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754-3774.
- Benavides, E., Russello, M., Boyer, D., Wiese, R.J., Kajdacs, B., Marquez, L., Garrick, R. & Caccone, A. (2012) Lineage identification and genealogical relationships among captive Galápagos tortoises. *Zoo Biology*, **31**, 107-120.
- Bensasson, D., Zhang, D.-X., Hartl, D.L. & Hewitt, G.M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution*, **16**, 314-321.
- Bermingham, E. & Avise, J.C. (1986) Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics*, **113**, 939-965.
- Bermingham, E. & Moritz, C. (1998) Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367-369.
- Beyers, R. (2008) *Natural and anthropogenic influences on elephants and other ungulates in the Congo forest*. Unpublished PhD thesis, University of British Columbia.
- Blyton, M.D.J., Banks, S.C., Peakall, R. & Lindenmayer, D.B. (2012) Using probability modelling and genetic parentage assignment to test the role of local mate availability in mating system variation. *Molecular Ecology*, **21**, 572-586.
- Bodmer, R.E. & Gubista, K.R. (1988) A note on the social structure of free-ranging okapi. *Acta Zoologica et Pathologica Antverpiensia*, **80**, 11-18.
- Bodmer, R.E. & Rabb, G.B. (1992) Okapi johnstoni. *Mammalian Species*, **422**, 1-8.
- Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C. & Taberlet, P. (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261-3273.
- Bonis, L.d., Koufos, G. & Sen, S. (1997) A giraffid from the Middle Miocene of the island of Chios, Greece. *Palaeontology*, **40**, 121-133.
- Bowland, A.E. & Perrin, M.R. (2009) Temporal and spatial patterns in blue duikers *Philatomba monticola* and red duikers *Cephalophus natalensis*. *Journal of Zoology*, **237**, 487-498.
- Bowler, D.E. & Benton, T.G. (2005) Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological reviews of the Cambridge Philosophical Society*, **80**, 205-225.
- Braga, J. (1995) Skeletal variation and measure of divergence among chimpanzees. *Comptes Rendus de l'Academie des Sciences*, **320**, 1025-1030.
- Bromham, L. & Penny, D. (2003) The modern molecular clock. *Nature reviews. Genetics*, **4**, 216-224.
- Brown, D.M., Brenneman, R.a., Koepfli, K.-P., Pollinger, J.P., Milá, B., Georgiadis, N.J., Louis, E.E., Grether, G.F., Jacobs, D.K. & Wayne, R.K. (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57-70.
- Brown, S.M. & Houlden, B.A. (2000) Conservation genetics of the black rhinoceros (*Diceros bicornis*). *Conservation Genetics*, **1**, 365-370.

- Bruford, M.W. & Wayne, R.K. (1993) Microsatellites and their application to population genetic studies. *Current Opinion in Genetics & Development*, **3**, 939-943.
- Burzynski, A. (2003) Evidence for Recombination of mtDNA in the Marine Mussel *Mytilus trossulus* from the Baltic. *Molecular Biology and Evolution*, **20**, 388-392.
- Busby, G.B.J., Gottelli, D., Wachter, T., Belbachir, L.M.F., DeSmet, K., Fellous, A.B.-B.A., Belghoul, M. & Durant, S.M. (2009) Genetic analysis of scat reveals leopard *Panthera pardus* and cheetah *Acinonyx jubatus* in southern Algeria. *Oryx*, **43**, 412-415.
- Caballero, S., Jackson, J., Mignucci-Giannoni, A.a., Barrios-Garrido, H., Beltrán-Pedrerros, S., Montiel-Villalobos, M.a.G., Robertson, K.M. & Baker, C.S. (2008) Molecular systematics of South American dolphins Sotalia: sister taxa determination and phylogenetic relationships, with insights into a multi-locus phylogeny of the Delphinidae. *Molecular phylogenetics and evolution*, **46**, 252-268.
- Cameron, E.Z. (2004) Facultative adjustment of mammalian sex ratios in support of the Trivers-Willard hypothesis: evidence for a mechanism. *Proceedings of the Royal Society B*, **271**, 1723-1728.
- Caswell, J.L., Mallick, S., Richter, D.J., Neubauer, J., Schirmer, C., Gnerre, S. & Reich, D. (2008) Analysis of chimpanzee history based on genome sequence alignments. *PLoS genetics*, **4**, e1000057.
- Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1660.
- Clifford, S.L., Anthony, N.M., Bawe-Johnson, M., Abernethy, K.a., Tutin, C.E.G., White, L.J.T., Bermejo, M., Goldsmith, M.L., McFarland, K., Jeffery, K.J., Bruford, M.W. & Wickings, E.J. (2004) Mitochondrial DNA phylogeography of western lowland gorillas (*Gorilla gorilla gorilla*). *Molecular Ecology*, **13**, 1551-1565, 1567.
- Clutton-Brock, T.H. (1989a) Female transfer and inbreeding avoidance in social mammals. *Nature*, **337**, 70-72.
- Clutton-Brock, T.H. (1989b) Mammalian mating systems. *Proceedings of the Royal Society B Biological Sciences*, **236**, 339-372.
- Cockburn, A. (1988) *Social behaviour in fluctuating populations*, London: Croom Helm.
- Cowling, S.A., Cox, P.M., Jones, C.D., Maslin, M.A., Peros, M. & Spall, S.A. (2008) Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. *Global Change Biology*, **14**, 827-840.
- Creer, S. (2007) Choosing and using introns in molecular phylogenetics. *Evolutionary bioinformatics online*, **3**, 99-108.
- Csilléry, K., Blum, M.G.B., Gaggiotti, O.E. & François, O. (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, **25**, 410-418.
- Darwin, C. (1871) *The descent of man and selection in relation to sex*, New York: Modern Library, Random House.
- Davey, J.W., Hohenlohe, P.a., Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature reviews. Genetics*, **12**, 499-510.
- DeMenocal, P.B. (1995) Plio-Pleistocene African Climate. *Science*, **270**, 53-59.

- DeMenocal, P.B. (2004) African climate change and faunal evolution during the Pliocene–Pleistocene. *Earth and Planetary Science Letters*, **220**, 3–24.
- DiBattista, J.D., Feldheim, K.a., Thibert-Plante, X., Gruber, S.H. & Hendry, A.P. (2008) A genetic assessment of polyandry and breeding-site fidelity in lemon sharks. *Molecular Ecology*, **17**, 3337–3351.
- Dietrich, W.O. (1942) Ältestquartäre Säugetiere aus der Südlichen Serengeti, Deutsch-Ostafrika. *Palaeontographica, Abt. A*, **94**, 3–22.
- Dobson, F.S., Way, B.M. & Baudoin, C. (2010) Spatial dynamics and the evolution of social monogamy in mammals. *Behavioral Ecology*, **21**, 747–752.
- Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 214.
- EDGE (2014) *EDGE of existence*
http://www.edgeofexistence.org/about/media_enquiries.php. Accessed 7th Jan 2014. http://www.edgeofexistence.org/about/media_enquiries.php 2014]
- Eggert, L.S., Rasner, C.A. & Woodruff, D.S. (2002) The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proceedings of the Royal Society B*, **269**, 1471–2954.
- Eisenberg, J.F. & Lockhart, M. (1972) An ecological reconnaissance of Wilpattu National Park, Ceylon. . *Smithsonian Contributions to Zoology*, **101**, 205–220.
- Emlen, S.T. & Oring, L.W. (1977) Ecology, sexual selection, and the evolution of mating systems. *Science*, **197**, 215–223.
- Eriksson, J., Hohmann, G., Boesch, C. & Vigilant, L. (2004) Rivers influence the population structure of bonobos (*Pan paniscus*). *Molecular Ecology*, **13**, 3425–3435.
- Eriksson, J., Siedel, H., Lukas, D., Kayser, M., Erler, A., Hashimoto, C., Hohmann, G., Boesch, C. & Vigilant, L. (2006) Y-chromosome analysis confirms highly sex-biased dispersal and suggests a low male effective population size in bonobos (*Pan paniscus*). *Molecular Ecology*, **15**, 939–949.
- 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.
- Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
- Fabbri, E., Caniglia, R., Mucci, N., Thomsen, H.P., Krag, K., Pertoldi, C., Loeschcke, V. & Randi, E. (2012) Comparison of single nucleotide polymorphisms and microsatellites in non-invasive genetic monitoring of a wolf population. *Archives of Biological Sciences*, **64**, 321–335.
- Faircloth, B.C. (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, **8**, 92–94.
- Faria, P.J., Kavembe, G.D., Jung’a, J.O., Kimwele, C.N., Estes, L.D., Reillo, P.R., Mwangi, A.G. & Bruford, M.W. (2011) The use of non-invasive molecular techniques to confirm the presence of mountain bongo *Tragelaphus eurycerus isaaci* populations in Kenya and preliminary inference of their mitochondrial genetic variation. *Conservation Genetics*, **12**, 745–751.
- Fernández, J., Villanueva, B., Pong-Wong, R. & Toro, M.A. (2005) Efficiency of the use of pedigree and molecular marker information in conservation programs. *Genetics*, **170**, 1313–1321.

- Fernandez, M.H. & Vrba, E.S. (2005) A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. *Biological reviews of the Cambridge Philosophical Society*, **80**, 269-302.
- Fienieg, E.S. & Galbusera, P. (2013) The use and integration of molecular DNA information in conservation breeding programmes: a review. *Journal of Zoo and Aquarium Research*, **1**, 44-51.
- Fischer, A., Prüfer, K., Good, J.M., Halbwax, M., Wiebe, V., André, C., Atencia, R., Mugisha, L., Ptak, S.E. & Pääbo, S. (2011) Bonobos fall within the genomic variation of chimpanzees. *PloS one*, **6**, e21605.
- Fischer, J. & Lindenmayer, D.B. (2000) An assessment of the published results of animal relocations. *Biological Conservation*, **96**, 1-11.
- Fisher, R.A. (1930) *The genetical theory of natural selection*, Clarendon Press, Oxford.
- Forstmeier, W., Segelbacher, G., Mueller, J.C. & Kempenaers, B. (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology*, **16**, 4039-4050.
- Frankel, O.H. & Soulé, M.E. (1981) *Conservation and evolution*, Cambridge University Press.
- Frankham, R. (1995) Inbreeding and extinction: a threshold effect. *Conservation Biology*, **9**, 792-799.
- Frankham, R. (2002) Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, **10**, 1500-1508.
- Frankham, R. (2008) Genetic adaptation to captivity in species conservation programs. *Molecular Ecology*, **17**, 325-333.
- Frankham, R. (2010) Where are we in conservation genetics and where do we need to go? *Conservation Genetics*, **11**, 661-663.
- Frankham, R., Ballou, J.D. & Briscoe, D.A. (2002) *Introduction to Conservation Genetics*, Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, United Kingdom.
- Frankham, R., Hemmer, H., Ryder, O.A., Cothran, E.G., Soulé, M.E., Murray, N.D. & Snyder, M. (1986) Selection in captive populations. *Zoo Biology*, **5**, 127-138.
- Frankham, R., Lees, K., Montgomery, M.E., England, P.R., Lowe, E.H. & Briscoe, D.A. (1999) Do population size bottlenecks reduce evolutionary potential? *Animal Conservation*, **2**, 255-260.
- Frankham, R. & Ralls, K. (1998) Conservation biology: Inbreeding leads to extinction. *Nature*, **392**, 441-442.
- Fraser, D.J. & Bernatchez, L. (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, **10**, 2741-2752.
- Freeland, J.R. (2005) *Molecular Ecology*, John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.
- Freeman, A.R., Machugh, D.E., Mckeown, S., Walzer, C., Mcconnell, D.J. & Bradley, D.G. (2001) Sequence variation in the mitochondrial DNA control region of wild African cheetahs (*Acinonyx jubatus*). *Heredity*, **86**, 355-362.
- Funk, W.C., Blouin, M.S., Corn, P.S., Maxell, B.A., Pilliod, D.S., Amish, S. & Allendorf, F.W. (2005) Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, **14**, 483-496.
- Galbusera, P. (2013) Personal Communication.

- Garvin, M.R., Saitoh, K. & Gharrett, a.J. (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. *Molecular Ecology Resources*, **10**, 915-934.
- Gautschi, B., Jacob, G., Negro, J.J., Godoy, J.A., Müller, J.P. & Schmid, B. (2003) Analysis of relatedness and determination of the source of founders in the captive bearded vulture, *Gypaetus barbatus*, population. *Conservation Genetics*, **4**, 479-490.
- Geist, V. (1974) On the Relationship of Social Evolution and Ecology in Ungulates. *American Journal of Zoology*, **14**, 205-220.
- GeneCodes Sequencher. pp. Sequence analysis software. Gene Codes Corporation, Ann Arbor, MI USA <http://www.genecodes.com>. Accessed 12 December 2013.
- Gentry, A.W. (1990) Ruminant artiodactyls of Pasalar, Turkey. *Journal of Human Evolution*, **19**, 529-550.
- Gijzen, A. & Smet, S. (1974a) Seventy years of okapi, *Okapia johnstoni* (Sclater, 1901). *Acta Zoologica et Pathologica Antverpiensia*, **59**, 1-90.
- Gijzen, A. & Smet, S. (1974b) Seventy years okapi, *Okapia johnstoni* (Sclater, 1901). *Acta Zoologica et Pathologica Antverpiensia*, **59**, 1-90.
- Głazewska, I., Prusak, B. & Gralak, B. (2013) Pedigrees as a source of information in mtDNA studies of dogs and horses. *Animal Genetics*, **44**, 227-230.
- Glenn, T.C. & Schable, N.A. (2005) Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**, 202-222.
- Gonçalves da Silva, A., Lalonde, D.R., Quse, V., Shoemaker, A. & Russello, M.A. (2010) Genetic approaches refine ex situ lowland tapir (*Tapirus terrestris*) conservation. *The Journal of Heredity*, **101**, 581-590.
- Goossens, B., Chikhi, L., Jalil, M.F., Ancrenaz, M., Lackman-Ancrenaz, I., Mohamed, M., Andau, P. & Bruford, M.W. (2005) Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia. *Molecular Ecology*, **14**, 441-456.
- Goudet, J. (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485-486.
- Goudet, J., Perrin, N. & Waser, P. (2002) Tests for sex-biased dispersal using bi-parentally inherited. *Molecular Ecology*, **11**, 1103-1114.
- Graur, D. & Martin, W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in genetics*, **20**, 80-86.
- Greenwood, P.J. (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**, 1140-1162.
- Gren, G.A., Zhou, Q. & Zhong, W. (1989) Ecology and social behaviour of Mongolian gerbils, *Meriones unguiculatus*, at Xilinhote, Inner Mongolia, China. *Animal Behaviour*, **37**.
- Griffith, B., Scott, M.J., Carpenter, J.W. & Reed, C. (1989) Translocation as a species conservation tool: status and strategy. *Science(Washington)*, **245**, 477-480.
- Grivet, D., Sork, V.L., Westfall, R.D. & Davis, F.W. (2008) Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): a multivariate genetic approach to conservation planning. *Molecular Ecology*, **17**, 139-156.
- Guindon, S. & Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systems Biology*, 696-704.

- Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A.C. (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature*, **352**, 255-257.
- Hájková, P., Zemanová, B., Roche, K. & Hájek, B. (2008) An evaluation of field and noninvasive genetic methods for estimating Eurasian otter population size. *Conservation Genetics*, **10**, 1667-1681.
- Hamilton, A.C. (1975) A quantitative analysis of altitudinal zonation in Ugandan forests. *Plant Ecology*, **30**, 99-106.
- Harris, J.M. (1991) Family Giraffidae. In *Koobi Fora Research Project. Volume 3. The fossils ungulates: geology, fossil artiodactyls, and palaeoenvironments*. (ed J.M. Harris), pp. 93-138. Clarendon Press, Oxford.
- Harris, J.M., Solounias, N. & Geraads, D. (2010) Giraffoidea. In *In: Werdelin L, Sanders WJ, editors. Cenozoic Mammals of Africa* pp. 797-811. University of California Press, Berkeley, Los Angeles & London.
- Hart, J. (2013) *Okapia johnstoni*, Okapi. In *Mammals of Africa: Vol VI: Pigs, Hippopotamuses, Chevrotain, Giraffes, Deer and Bovids* (eds J. Kingdon & M. Hoffmann), Bloomsbury Publishing, London.
- Hart, J., Beyers, R., Grossmann, F., Carbo, M., Dino, S.o. & Kahindo, F. (2008) La Reserve de Faune a Okapis. La distribution et frequence de la grande faune et des activites humaines. Avec une evaluation de l'impact de 10 ans de conflict: 1996 - 2006. Kinshasa: IMU Rapport No 9.
- Hart, J., Detwiler, K.M., Gilbert, C.C., Burrell, A.S., Fuller, J.L., Emetshu, M., Hart, T.B., Vosper, A., Sargis, E.J. & Tosi, A.J. (2012) Lesula: A new species of cercopithecus monkey endemic to the Democratic Republic of Congo and implications for conservation of Congo's central basin. *PloS one*, **7**, e44271.
- Hart, J.A. & Hart, T.B. (1989) Ranging and feeding behaviour of okapi (*Okapia johnstoni*) in the Ituri Forest of Zaire: food limitation in a rain-forest herbivore? *Symposium of the Zoological Society of London*, **61**, 31-50.
- Hartl, D.L. & Clark, A.G. (1997) *Principles of population genetics*, Sinauer associates, Sunderland.
- Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., van Vuuren, B.J., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskou, V., Nguyen, T.T. & Couloux, A. (2012) Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *Comptes Rendus Biologies*, 32-50.
- Hassanin, A., Ropiquet, A., Gourmand, A.-L., Chardonnet, B. & Rigoulet, J. (2007) Mitochondrial DNA variability in *Giraffa camelopardalis*: consequences for taxonomy, phylogeography and conservation of giraffes in West and central Africa. *Comptes Rendus Biologies*, **330**, 265-274.
- Hasselquist, D. & Sherman, P.W. (2001) Social mating systems and extrapair fertilizations in passerine birds. *Behavioral Ecology*, **12**, 457-466.
- Hayes, K. & Burge, R. (2003) *Coltan Mining in the Democratic Republic of Congo: How tantalum-using industries can commit to the reconstruction of the DRC*. Fauna & Flora International. Cambridge, UK.
- Haynes, G.D. & Latch, E.K. (2012) Identification of novel single nucleotide polymorphisms (SNPs) in deer (*Odocoileus spp.*) using the BovineSNP50 BeadChip. *PloS one*, **7**, e36536.
- Hedrick, P.W. & Fredrickson, R.J. (2008) Captive breeding and the reintroduction of Mexican and red wolves. *Molecular Ecology*, **17**, 344-350.
- Hedrick, P.W., Miller, P.S., Geffen, E. & Wayne, R.K. (1997) Genetic evaluation of the three captive mexican wolf lineages. *Zoo Biology*, **16**, 47-69.

- Henkel, J.R., Jones, K.L., Hereford, S.G., Savoie, M.L., Leibo, S.P. & Howard, J.J. (2012) Integrating microsatellite and pedigree analyses to facilitate the captive management of the endangered Mississippi sandhill crane (*Grus canadensis pulla*). *Zoo Biology*, **31**, 322-335.
- Henry, P., Miquelle, D., Sugimoto, T., McCullough, D.R., Caccone, A. & Russello, M.A. (2009) *In situ* population structure and *ex situ* representation of the endangered Amur tiger. *Molecular Ecology*, **18**, 3173-3184.
- Hewison, A.J.M. & Gaillard, J.-m. (1999) Successful sons or advantaged investment in ungulates. *Trends in Ecology & Evolution*, **14**, 229-234.
- Hey, J. (2010) The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. *Molecular Biology and Evolution*, **27**, 921-933.
- Hilgartner, R., Fichtel, C., Kappeler, P.M. & Zinner, D. (2012) Determinants of Pair-Living in Red-Tailed Sportive Lemurs (*Lepilemur ruficaudatus*). *Ethology : formerly Zeitschrift fur Tierpsychologie*, **118**, 466-479.
- Ho, S.Y.W., Lanfear, R., Bromham, L., Phillips, M.J., Soubrier, J., Rodrigo, A.G. & Cooper, A. (2011) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**, 3087-3101.
- Ho, S.Y.W. & Larson, G. (2006) Molecular clocks: when times are a-changin'. *Trends in Genetics*, **22**, 79-83.
- Howard, W.E. (1960) Innate and environmental dispersal of individual vertebrates. *American Midland Naturalist*, **63**, 152-161.
- Hu, J., Pan, H.-J., Wan, Q.-H. & Fang, S.-G. (2007) Nuclear DNA microsatellite analysis of genetic diversity in captive populations of Chinese water deer. *Small Ruminant Research*, **67**, 252-256.
- IES (2008) [Institute for Environmental Security] *Mining, forest change and conflict in the Kivus, eastern Democratic Republic of Congo*. The Hague: IES.
- IUCN (2008) *IUCN SSC Antelope Specialist Group 2008. Okapia johnstoni*. In: *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1*. <http://www.iucnredlist.org> Downloaded on 02 October 2013. [accessed 1st August
- IUCN (2013) *The IUCN Red List of Threatened Species. Version 2013.1*. <http://www.iucnredlist.org>. Downloaded on 02 July 2013.
- Ivy, J.A., Miller, A., Lacy, R.C. & Dewoody, J.A. (2009) Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (*Macropus parma*). *The Journal of Heredity*, **100**, 441-454.
- Jenkins, C.N., Pimm, S.L. & Joppa, L.N. (2013) Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E2602-2610.
- Johnson, P.C.D. & Haydon, D.T. (2007) Maximum likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. *Genetics*, **175**, 827-842.
- Johnston, A.R. & Anthony, N.M. (2012) A multi-locus species phylogeny of African forest duikers in the subfamily Cephalophinae: evidence for a recent radiation in the Pleistocene. *BMC Evolutionary Biology*, **12**, 120.
- Johnston, H. (1900a) On a new horse. *Proceedings of the Zoological Society of London*, 774-774.
- Johnston, H. (1900b) On a new horse. *Proceedings of the Zoological Society of London*, 774-775.

- Jones, O. & Wang, J. (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, **10**, 551-555.
- Jule, K.R., Leaver, L.A. & Lea, S.E.G. (2008) The effects of captive experience on reintroduction survival in carnivores: A review and analysis. *Biological Conservation*, **141**, 355-363.
- Kappeler, P.M. (1999) Primate socioecology: New insights from males. *Naturwissenschaften*, **86**, 18-29.
- Kass, R.E. & Raftery, A.E. (1995) Bayes Factors. *Journal of the American Statistical Association*, **90**, 773-795.
- Kawamoto, Y., Takemoto, H., Higuchi, S., Sakamaki, T., Hart, J.a., Hart, T.B., Tokuyama, N., Reinartz, G.E., Guislain, P., Dupain, J., Cobden, A.K., Mulavwa, M.N., Yangozene, K., Darroze, S., Devos, C. & Furuichi, T. (2013) Genetic structure of wild bonobo populations: diversity of mitochondrial DNA and geographical distribution. *PloS one*, **8**, e59660.
- Kelly, C.D., Godin, J.-g.J. & Wright, J.M. (1999) Geographical variation in multiple paternity within natural populations of the guppy (*Poecilia reticulata*). *Proceedings of the Royal Society B Biological Sciences*, **266**, 2403-2408.
- Kennis, J., Nicolas, V., Hulselmans, J., Katuala, P.G.B., Wendelen, W., Verheyen, E., Dudu, A.M. & Leirs, H. (2011) The impact of the Congo River and its tributaries on the rodent genus *Praomys*: speciation origin or range expansion limit? *Zoological Journal of the Linnean Society*, **163**, 983-1002.
- Khan, H.A., Arif, I.A., Shobrak, M., Homaidan, A.A.A., Farhan, A.H.A. & Sadoon, M.A. (2011) Application of mitochondrial genes sequences for measuring the genetic diversity of Arabian oryx. *Genes & Genetic Systems*, **86**, 67-72.
- Kimura, M. & Crow, J.F. (1963) The measurement of effective population number. *Evolution*, **17**, 279-288.
- King, E.D.A., Banks, P.B. & Brooks, R.C. (2011) Sexual conflict in mammals: consequences for mating systems and life history. *Mammal Review*, **43**, 47-58.
- Kingdon, J. (1997) *The Kingdon fieldguide to African mammals*, Academic press Limited, 24-28 Oval Road, London NW1 7DX.
- Klaus-Hügi, C., Klaus, G. & Schmid, B. (2000) Movement Patterns and Home Range Of The Bongo (*Tragelaphus eurycerus*) In The Rain Forest Of The Dzanga National Park, Central African Republic. *African Journal of Ecology*, **38**, 53-61.
- Kleiman, D.G. (1977) Monogamy in mammals. *The Quarterly Review of Biology*, **52**, 39-69.
- Konovalov, D.A., Manning, C. & Henshaw, M.T. (2004) KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Molecular Ecology Notes*, **4**, 779-789.
- Kozfkay, C.C., Campbell, M.R., Heindel, J.A., Baker, D.J., Kline, P., Powell, M.S. & Flagg, T. (2008) A genetic evaluation of relatedness for broodstock management of captive, endangered Snake River sockeye salmon, *Oncorhynchus nerka*. *Conservation Genetics*, **9**, 1421-1430.
- Kvist, L. (2003) Paternal Leakage of Mitochondrial DNA in the Great Tit (*Parus major*). *Molecular Biology and Evolution*, **20**, 243-247.
- Lacy, R.C. (1989) Analysis of founder representation in pedigree: Founder equivalents and founder genome equivalents. *Zoo Biology*, **8**, 111-123.
- Lacy, R.C. (1993) VORTEX: A computer simulation model for population viability analysis. *Wildlife Research*, **20**, 45-65.

- Lacy, R.C. (2013) Achieving true sustainability of zoo populations. *Zoo Biology*, **32**, 19-26.
- Lacy, R.C., Ballou, J.D. & Pollak, J.P. (2012) PMx: software package for demographic and genetic analysis and management of pedigreed populations. *Methods in Ecology and Evolution*, **3**, 433-437.
- Ladoukakis, E.D. & Zouros, E. (2001) Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. *Molecular Biology and Evolution*, **18**, 1168-1175.
- Lankester, E.R. (1901) On *Okapia johnstoni*. *Proceedings of the Zoological Society of London*, **2**, 279-281.
- Lawson Handley, L.J. & Perrin, N. (2007) Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology*, **16**, 1559-1578.
- Leakey, R.E. & Lewin, R. (2000) *The Sixth Extinction: Patterns of Life and the Future of Humankind*, Library Services Branch, Province of British Columbia.
- Leberg, P.L. & Firmin, B.D. (2008) Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology*, **17**, 334-343.
- Lesobre, L., Lacroix, F., Caizergues, A., Hingrat, Y., Chalah, T. & Jalme, M.S. (2010) Conservation genetics of Houbara Bustard (*Chlamydotis undulata undulata*): population structure and its implications for the reinforcement of wild populations. *Conservation Genetics*, **11**, 1489-1497.
- Leus, K. & Hofman, S. (2012) *The international studbook for the Okapi (Okapia johnstoni) - 12 Oct. 2011*. The Royal Zoological Society of Antwerp: Antwerp.
- Li, H. & Durbin, R. (2011) Inference of human population history from individual whole-genome sequences. *Nature*, **475**, 493-496.
- Li, R., Fan, W., Tian, G., Zhu, H., He, L., Cai, J., Huang, Q., Cai, Q., Li, B., Bai, Y., Zhang, Z., Zhang, Y., Wang, W., Li, J., Wei, F., Li, H., Jian, M., Li, J., Zhang, Z., Nielsen, R., Li, D., Gu, W., Yang, Z., Xuan, Z., Ryder, O.A., Leung, F.C.-C., Zhou, Y., Cao, J., Sun, X., Fu, Y., Fang, X., Guo, X., Wang, B., Hou, R., Shen, F., Mu, B., Ni, P., Lin, R., Qian, W., Wang, G., Yu, C., Nie, W., Wang, J., Wu, Z., Liang, H., Min, J., Wu, Q., Cheng, S., Ruan, J., Wang, M., Shi, Z., Wen, M., Liu, B., Ren, X., Zheng, H., Dong, D., Cook, K., Shan, G., Zhang, H., Kosiol, C., Xie, X., Lu, Z., Zheng, H., Li, Y., Steiner, C.C., Lam, T.T.-Y., Lin, S., Zhang, Q., Li, G., Tian, J., Gong, T., Liu, H., Zhang, D., Fang, L., Ye, C., Zhang, J., Hu, W., Xu, A., Ren, Y., Zhang, G., Bruford, M.W., Li, Q., Ma, L., Guo, Y., An, N., Hu, Y., Zheng, Y., Shi, Y., Li, Z., Liu, Q., Chen, Y., Zhao, J., Qu, N., Zhao, S., Tian, F., Wang, X., Wang, H., Xu, L., Liu, X., Vinar, T., et al. (2010) The sequence and *de novo* assembly of the giant panda genome. *Nature*, **463**, 311-317.
- Librado, P. & Rozas, J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.
- Livingstone, D.A. (1975) Late Quaternary Climatic Change in Africa. *Annual Review of Ecology and Systematics*, **6**, 249-280.
- Lopes, J.S., Balding, D. & Beaumont, M. (2009) PopABC: a program to infer historical demographic parameters. *Bioinformatics*, **25**, 2747-2749.
- Lorenzen, E.D., Masembe, C., Arctander, P. & Siegismund, H.R. (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.

- Lu, Z., Johnson, W.E., Menotti-Raymond, M., Yuhki, N., Martenson, J.S., Mainka, S., Shi-Qiang, H., Zhihe, Z., Li, G., Pan, W., Mao, X. & O'Brien, S.J. (2001) Patterns of Genetic Diversity in Remaining Giant Panda Populations. *Conservation Biology*, **15**, 1596-1607.
- Lukas, D. & Clutton-Brock, T.H. (2013) The evolution of social monogamy in mammals. *Science*, **341**, 526-530.
- Maglio, V.J. (1973) Origin and evolution of the elephantidae. *Transactions of the American Philosophical Society*, **63**, 1-149.
- Magliocca, F., Querouil, S. & Gautier-Hion, A. (2002) Grouping patterns, reproduction, and dispersal in a population of sitatungas (*Tragelaphus spekei gratus*). *Canadian Journal of Zoology*, **80**, 245-250.
- Malele Mbala, S. (2010) *Analysis of forest financing in the democratic republic of congo*.
- Mallon, D., Kümpel, N., Quinn, A., Shurter, S., Lukas, J., Hart, J. & Mapilanga, J. (2013) *Okapia johnstoni*. In: *IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2*. <http://www.iucnredlist.org> Downloaded on 26 November 2013.
- Manel, S., Schwartz, M.K., Luikart, G. & Taberlet, P. (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189-197.
- Mapilanga, J.J. (2013) Realities on the terrain to conserve biodiversity in DRC: Post conflict country. A presentation by the ICCN Director of National Parks in charge, at the Zoological Society of London, 27th November.
- Mardis, E.R. (2008) The impact of next-generation sequencing technology on genetics. *Trends in genetics*, **24**, 133-141.
- Margan, S.H., Nurthen, R.K., Montgomery, M.E., Woodworth, L.M., Lowe, E.H., Briscoe, D.A. & Frankham, R. (1998) Single large or several small? Population fragmentation in the captive management of endangered species. *Zoo Biology*, **17**, 467-480.
- Martins, K., Chaves, L.J., Buso, G.S.C. & Kageyama, P.Y. (2006) Mating system and fine-scale spatial genetic structure of *Solanum lycocarpum* St.Hil. (Solanaceae) in the Brazilian Cerrado. *Conservation Genetics*, **7**, 957-969.
- Maslin, M.A. & Christensen, B. (2007) Tectonics, orbital forcing, global climate change, and human evolution in Africa: introduction to the African paleoclimate special volume. *Journal of Human Evolution*, **53**, 443-464.
- Mboumba, J.F., Deleporte, P., Colyn, M. & Nicolas, V. (2011) Phylogeography of *Mus (Nannomys) minutoides* (Rodentia, Muridae) in West Central African savannahs: singular vicariance in neighbouring populations. *Journal of Zoological Systematics and Evolutionary Research*, **49**, 77-85.
- McGreevy, T.J., Dabek, L., Gomez-Chiarri, M. & Husband, T.P. (2009) Genetic diversity in captive and wild Matschie's tree kangaroo (*Dendrolagus matschiei*) from Huon Peninsula, Papua New Guinea, based on mtDNA control region sequences. *Zoo Biology*, **28**, 183-196.
- McGreevy, T.J., Dabek, L. & Husband, T.P. (2011) Genetic evaluation of the Association of Zoos and Aquariums Matschie's tree kangaroo (*Dendrolagus matschiei*) captive breeding program. *Zoo Biology*, **30**, 636-646.
- Miller, W., Wright, S.J., Zhang, Y., Schuster, S.C. & Hayes, V.M. (2010) Optimization methods for selecting founder individuals for captive breeding or reintroduction of endangered species. *Pacific Symposium on Biocomputing*, 2010.

- Möller, L.M. (2012) Sociogenetic structure, kin associations and bonding in delphinids. *Molecular Ecology*, **21**, 745-764.
- Mondol, S., Bruford, M.W. & Ramakrishnan, U. (2013) Demographic loss, genetic structure and the conservation implications for Indian tigers. *Proceedings of the Royal Society B*, **280**, 20130496.
- Moodley, Y. & Bruford, M.W. (2007) Molecular biogeography: towards an integrated framework for conserving pan-African biodiversity. *PloS one*, **2**, e454.
- Morin, P.A. & McCarthy, M. (2007) Highly accurate SNP genotyping from historical and low-quality samples. *Molecular Ecology Notes*, **7**, 937-946.
- Moritz, C. (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, **9**, 373-375.
- Muñoz-fuentes, V., Green, A.J. & Sorenson, M.D. (2008) Comparing the genetics of wild and captive populations of White-headed Ducks *Oxyura leucocephala*: consequences for recovery programmes. *Ibis*, **150**, 807-815.
- Muwanika, V.B., Nyakaana, S., Siegismund, H.R. & Arctander, P. (2003) Phylogeography and population structure of the common warthog (*Phacochoerus africanus*) inferred from variation in mitochondrial DNA sequences and microsatellite loci. *Heredity*, **91**, 361-372.
- Navidi, W., Arnheim, N. & Waterman, M.S. (1992) A multiple-tube approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *American Journal of Human Genetics*, **50**, 347-359.
- Nersting, L.G. & Arctander, P. (2001) Phylogeography and conservation of impala and greater kudu. *Molecular Ecology*, **10**, 711-719.
- Nicolas, V., Missoup, A.D., Denys, C., Kerbis Peterhans, J., Katuala, P., Couloux, A. & Colyn, M. (2011) The roles of rivers and Pleistocene refugia in shaping genetic diversity in *Praomys misonnei* in tropical Africa. *Journal of Biogeography*, **38**, 191-207.
- Nielsen, R. & Wakeley, J. (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885-896.
- Nixon, S.C. & Lusenge, T. (2008) *Conservation status of okapi in Virunga National Park, Democratic Republic of Congo*, The Zoological Society of London, London.
- Nsubuga, A.M., Holzman, J., Chemnick, L.G. & Ryder, O.A. (2010) The cryptic genetic structure of the North American captive gorilla population. *Conservation Genetics*, **11**, 161-172.
- O'Neill, E.M., Schwartz, R., Bullock, C.T., Williams, J.S., Shaffer, H.B., Aguilar-Miguel, X., Parra-Olea, G. & Weisrock, D.W. (2013) Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. *Molecular Ecology*, **22**, 111-129.
- Okapi Conservation Project (2012) <http://www.okapiconservation.org/news/okapi-conservation-project-epulu-update-june-28-2012/>.
- Okapi Conservation Workshop (2013). In *Okapi Conservation Workshop*, Kisangani, Democratic Republic of Congo. June.
- Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. & Hofreiter, M. (2004) Genetic analyses from ancient DNA. *Annual review of genetics*, **38**, 645-679.
- Peakall, R. & Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

- Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537-2539.
- Pearson, H., Greed, G.R. & Wright, A.I. (1978) An account of the okapi herd owned by the Bristol, Clifton and West of England Zoological Society. *Acta Zoologica et Pathologica Antverpiensia*, **71**, 79-86.
- Pennock, D.S. & Dimmick, W.W. (1997) Critique of the Evolutionarily Significant Unit as a Definition for “Distinct Population Segments” under the U.S. Endangered Species Act. *Conservation Biology*, **11**, 611-619.
- Pope, L.C., Domingo-Roura, X., Erven, K. & Burke, T. (2006) Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology*, **15**, 371-386.
- Posada, D. (2008) jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, 1253-1256.
- Potapov, P.V., Turubanova, S.A., Hansen, M.C., Adusei, B., Broich, M., Altstatt, A., Mane, L. & Justice, C.O. (2012) Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sensing of Environment*, **122**.
- Potts, R. (2013) Hominin evolution in settings of strong environmental variability. *Quaternary Science Reviews*, **73**, 1-13.
- Primack, R.B. (2000) A Primer of Conservation Biology, Second Edition. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts U.S.A.
- Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959.
- Prost, S. & Anderson, C.N.K. (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution*, **2**, 663-667.
- Puritz, J.B., Addison, J.a. & Toonen, R.J. (2012) Next-generation phylogeography: a targeted approach for multilocus sequencing of non-model organisms. *PloS one*, **7**, e34241.
- Queller, D.C. & Goodnight, K.F. (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258-275.
- Quinn, A., Queslin, E. & Kümpel, N. ((In Prep)) Okapi conservation status review.
- Quinn, A., Queslin, E. & Kümpel, N. (2013) *Okapi (Okapia johnstoni): Conservation status review*. Zoological Society of London.
- R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rabb, G.B. (1978) Birth, early behavior and clinical data on the okapi. *Acta Zoologica et Pathologica Antverpiensia*, **71**, 93-105.
- Rahbek, C. (1993) Captive breeding—a useful tool in the preservation of biodiversity? *Biodiversity and Conservation*, **2**, 426-437.
- Rambaut, A. & Drummond, A.J. (2007) Tracer v1.5, Available from <http://tree.bio.ed.ac.uk/software/>. Accessed 11 June 2014.
- Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
- Reichard, U. (2010) Extra-pair Copulations in a Monogamous Gibbon (*Hylobates lar*). *Ethology : formerly Zeitschrift fur Tierpsychologie*, **100**, 99-112.

- Ribeiro, A.M., Lloyd, P., Feldheim, K.A. & Bowie, R.C.K. (2012) Microgeographic socio-genetic structure of an African cooperative breeding passerine revealed: integrating behavioural and genetic data. *Molecular Ecology*, **21**, 662-672.
- Robertson, A. (1964) The effect of non-random mating within inbred lines on the rate of inbreeding. *Genetical Research*, **5**, 164-167.
- Rohland, N., Pollack, J.L., Nagel, D., Beauval, C., Airvaux, J., Pääbo, S. & Hofreiter, M. (2005) The population history of extant and extinct hyenas. *Molecular Biology and Evolution*, **22**, 2435-2443.
- Rollins, L.A., Browning, L.E., Holleley, C.E., Savage, J.L., Russell, A.F. & Griffith, S.C. (2012) Building genetic networks using relatedness information: a novel approach for the estimation of dispersal and characterization of group structure in social animals. *Molecular Ecology*, **21**, 1727-1740.
- Ross, K.G. (2001) Molecular ecology of social behaviour: analyses of breeding systems and genetic structure. *Molecular Ecology*, **10**, 265-284.
- Rousset, F. (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.
- Rudnick, J.A. & Lacy, R.C. (2007) The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conservation Genetics*, **9**, 1439-1450.
- Russello, M.A., Hyseni, C., Gibbs, J.P., Cruz, S., Marquez, C., Tapia, W., Velensky, P., Powell, J.R. & Caccone, A. (2007) Lineage identification of Galápagos tortoises in captivity worldwide. *Animal Conservation*, **10**, 304-311.
- Santure, A.W., Stapley, J., Ball, A.D., Birkhead, T.R., Burke, T. & Slate, J. (2010) On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology*, **19**, 1439-1451.
- Schaller, G.B. (1993) The last panda. In *A Primers of Conservation Biology, Second Edition (Primack, R B)* pp. 135. Sinauer Associates, Inc., Sunderland, Massachusetts U.S.A.
- Schwacke, L., Schwacke, J. & Rosel, P. (2005) RE-RAT: relatedness estimation and rarefaction analysis tool. <http://people.musc.edu/~schwackh/>
- Schwartz, M. & Vissing, J. (2002) Paternal inheritance in mitochondrial DNA. *The New England Journal of Medicine*, **347**, 576-580.
- Sclater, P.L. (1901) On an apparently new Species of Zebra from the Semliki Forest. *Proceedings of the Zoological Society of London*, **1**, 50-52.
- Seddon, J.M., Parker, H.G., Ostrander, E.A. & Ellegren, H. (2005) SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology*, **14**, 503-511.
- Sheldon, B.C. (2004) Maternal dominance, maternal condition, and offspring sex ratio in ungulate mammals. *The American Naturalist*, **163**, 40-54.
- Shen, F., Zhang, Z., He, W., Yue, B., Zhang, A., Zhang, L., Hou, R., Wang, C. & Watanabe, T. (2009) Microsatellite variability reveals the necessity for genetic input from wild giant pandas (*Ailuropoda melanoleuca*) into the captive population. *Molecular Ecology*, **18**, 1061-1070.
- Snyder, N.F.R., Derrickson, S.R., Beissinger, S.R., Wiley, J.W., Smith, T.B., Toone, W.D. & Miller, B. (1996) Limitations of Captive Breeding in Endangered Species Recovery. *Conservation Biology*, **10**, 338-348.
- Soule, M., Gilpin, M., Conway, W. & Foote, T. (1986) The millenium ark: How long a voyage, how many staterooms, how many passengers? *Zoo Biology*, **5**, 101-113.

- Stanley, H.M. (1890) *In Darkest Africa*, Searle and Rivington, Ltd, London.
- Stanton, D.W.G., Hart, J., Galbusera, P., Helsen, P., Shephard, J., Kumpel, N.F., Wang, J., Ewen, J.G. & Bruford, M.W. (In Press) Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered okapi (*Okapia johnstoni*). *PloS one*.
- Stanton, D.W.G., Penfold, L.M., Zhan, X. & Bruford, M.W. (2010) Microsatellite loci for the okapi (*Okapia johnstoni*). *Conservation Genetics Resources*, **2**, 337-339.
- Starin, E.D. (2001) Notes on sitatunga in The Gambia. *African Journal of Ecology*, **38**, 339-342.
- Stephens, M., Smith, N.J. & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.
- Stockwell, C.A., Hendry, A.P. & Kinnison, M.T. (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94-101.
- Storz, J.F. (1999) Genetic Consequences of Mammalian Social Structure. *Journal of Mammalogy*, **80**, 553-569.
- Stuart, C. & Stuart, T. (1997) *Field guide to the larger mammals of Africa*, Struik Publishers, Cornelis Struik House, 80 McKenzie Street, Cape Town 8001.
- Sugg, D.W., Chesser, R.K., Dobson, F.S. & Hoogland, J.L. (1996) Population genetics meets behavioral ecology. *Trends in Ecology and Evolution*, **11**, 338-342.
- Taberlet, P., Camarra, J.-J., Griffin, S., Uhrès, E., Hanotte, O., Waits, L.P., Dubois-Paganon, C., Burke, T. & Bouvet, J. (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology*, **6**, 869-876.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G. & Cosson, J.-F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.
- Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P. & Bouvet, J. (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **24**.
- Taberlet, P. & Luikart, G. (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, **68**, 41-55.
- Taberlet, P., Waits, L.P. & Luikart, G. (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution*, **14**, 323-327.
- The Okapi Studbook (2013) <<http://theokapi.org/>> Downloaded on 26th November 2013. [accessed 1st August 2012]
- Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R., Marques, T.A. & Burnham, K.P. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, **47**, 5-14.
- Townsend, S.M. & Jamieson, I.G. (2013) Molecular and pedigree measures of relatedness provide similar estimates of inbreeding depression in a bottlenecked population. *Journal of Evolutionary Biology*, **26**, 889-899.
- Uchida, A. (1996) What we don't know about great ape variation. *Trends in Ecology and Evolution*, **16**, 163-168.
- United Nations (2011) <http://www.un.org> (accessed April 6, 2011).
- Valière, N. (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377-379.

- Valière, N., Berthier, P., Mouchiroud, D. & Pontier, D. (2002) GEMINI: software for testing the effects of genotyping errors and multitubes approach for the individual identification *Molecular Ecology Notes*, **2**, 83-86.
- Van der Made, J. & Morales, J. (2012) *Mitilanotherium inexpectatum* (Giraffidae, Mammalia) from Huélago (Lower Pleistocene; Guadix-Baza basin, Granada, Spain) - observations on a peculiar biogeographic pattern. *Estudios Geológicos*, **67**, 613-627.
- van Hooft, W.F., Groen, A.F. & Prins, H.H. (2002) Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Molecular Ecology*, **11**, 267-279.
- Veit, M.L., Robertson, R.J., Hamel, P.B. & Friesen, V.L. (2005) Population genetic structure and dispersal across a fragmented landscape in cerulean warblers (*Dendroica cerulea*). *Conservation Genetics*, **6**, 159-174.
- Vosper, A., Masselink, J. & Maisels, F. (2012) Great ape and human impact monitoring in Okapi Faunal Reserve, Democratic Republic of Congo.
- Wakeley, J. (2008) *Coalescent Theory: An Introduction*, Roberts & Company Publishers, Greenwood Village, Colorado.
- Walker, F.M., Sunnucks, P. & Taylor, A.C. (2008) Evidence for habitat fragmentation altering within-population processes in wombats. *Molecular Ecology*, **17**, 1674-1684.
- Wang, J. (2004) Application of the one - migrant - per - generation rule to conservation and management. *Conservation Biology*, **18**, 332-343.
- Wang, J. (2011) A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, **11**, 141-145.
- Webster, M.S., Tarvin, K.A., Tuttle, E.M. & Pruett-Jones, S. (2004) Reproductive promiscuity in the splendid fairy-wren: effects of group size and auxiliary reproduction. *Behavioral Ecology*, **15**, 907-915.
- Welsh, D. & Sederberg, J.S. (1990) Extra-pair copulations in black brant. *The Condor*, **92**, 242-244.
- Wiens, J. (2001) The landscape context of dispersal by Clobert, J; Danchin, E; Dhondt, A A and Nichols, J D. In *Dispersal* pp. 96-109. Oxford University Press, New York.
- Williams, S.E. & Hoffman, E.A. (2009) Minimizing genetic adaptation in captive breeding programs: A review. *Biological Conservation*, **142**, 2388-2400.
- Wilson, P.J. & White, B.N. (1998) Sex identification of Elk (*Cervus elaphus canadensis*), Moose (*Alces alces*), and White-Tailed Deer (*Odocoileus virginianus*) using the Polymerase Chain Reaction. *Journal of Forensic Science*, **43**, 477-482.
- Witzenberger, K.A. & Hochkirch, A. (2011) Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animals. *Biodiversity and Conservation*, **20**, 1843-1861.
- Witzenberger, K.A. & Hochkirch, A. (2013) Evaluating ex situ conservation projects: Genetic structure of the captive population of the Arabian sand cat. *Mammalian Biology*, **78**, 379-382.
- Wolf, M.C., Griffith, B., Reed, C. & Temple, S.A. (2002) Avian and Mammalian Translocations: Update and Reanalysis of 1987 Survey Data. *Conservation Biology*, **10**, 1142-1154.

- Won, Y.-J. & Hey, J. (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, **22**, 297-307.
- Woodworth, L.M., Montgomery, M.E., Briscoe, D.A. & Frankham, R. (2002) Rapid genetic deterioration in captive populations: Causes and conservation implications. *Conservation Genetics*, **3**, 277-288.
- Wright, S. (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97-159.
- Xia, X. & Xie, Z. (2001) DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity*, **92**, 371-373.
- Zhan, X.J., Zhang, Z.J., Wu, H., Goossens, B., Li, M., Jiang, S.W., Bruford, M.W. & Wei, F.W. (2007) Molecular analysis of dispersal in giant pandas. *Molecular Ecology*, **16**, 3792.
- Zhang, B., Li, M., Zhang, Z., Goossens, B., Zhu, L., Zhang, S., Hu, J., Bruford, M.W. & Wei, F. (2007) Genetic Viability and Population History of the Giant Panda, Putting an End to the “Evolutionary Dead End”? *Molecular Biology and Evolution*, **24**, 1801-1810.