Genetic structure of captive and free-ranging okapi (*Okapia johnstoni*) with implications for management

David W. G. Stanton[1], Philippe Helsen[2,9], Jill Shephard[2,3], Kristin Leus[4], Linda Penfold[5], John Hart[6], Noëlle F. Kümpel[7], John G. Ewen[8], Jinliang Wang[8], Peter Galbusera[2], Michael W. Bruford[1]

[1]School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK

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

[3]Present Address: School of Veterinary and Life Sciences, Murdoch University, Perth 6150, Australia

[4]Conservation Breeding Specialist Group, European Regional Network, Copenhagen Zoo, p/a Annuntiatenstraat 6, 2170 Merksem, Belgium

[5]South-East Zoo Alliance for Reproduction and Conservation (SEZARC), Yulee, FL, USA

[6]Lukuru Foundation, Projet Tshuapa-Lomami-Lualaba (TL2), 1235 Ave Poids Lourds, Kinshasa, Democratic Republic of the Congo

[7]Conservation Programmes, Zoological Society of London, London NW1 4RY, UK

[8]Institute of Zoology, Zoological Society of London, London NW1 4RY, UK

[9]Evolutionary Ecology Group, University of Antwerp, 2020 Antwerp, Belgium

Abstract

Breeding programs for endangered species increasingly use molecular genetics to inform their management strategies. Molecular approaches can be useful for investigating relatedness, resolving pedigree uncertainties, and for estimating genetic diversity in captive and wild populations. Genetic data can also be used to evaluate the representation of wild population genomes within captive population gene-pools. Maintaining a captive population that is genetically representative of its wild counterpart offers a means of conserving the original evolutionary potential of a species. Okapi, an even-toed ungulate, endemic to the Democratic Republic of Congo, have recently been reclassified as Endangered by the IUCN. We carried out a genetic assessment of the ex-situ okapi (*Okapia johnstoni*) population, alongside an investigation into the genetic structure of wild populations across their geographic range. We found that while levels of nuclear (12 microsatellite loci) genetic variation in the wild, founder and captive okapi populations were similar, mitochondrial (833 bp of Cyt *b*, CR, tRNA-Thr and tRNA-Pro) variation within captive okapi was considerably reduced compared to the wild, with 16 % lower haplotype diversity. Further, both nuclear and mitochondrial alleles present in captivity provided only partial representation of those present in the wild. Thirty mitochondrial haplotypes found in the wild were not found in captivity, and two haplotypes found in captivity were not found in the wild, and the patterns of genetic variation at microsatellite loci in our captive samples were considerably different to those of the wild samples. Our study highlights the importance of genetic characterisation of captive populations, even for well-managed ex-situ breeding programs with detailed studbooks. We recommend that the captive US population should be further genetically characterised to guide management of translocations between European and US captive populations.

Keywords: Captive; Wild; Ex-situ; In-situ; Founder

Introduction

A major objective of many modern captive breeding programs is to maintain sustainable populations and, if required, provide a source of individuals for eventual reintroduction into the wild. To achieve this objective, the effective maintenance of genetic diversity is considered to be a key goal (Kozfkay et al. 2008; Lacy 2013). Many captive breeding programs use pedigrees to guide genetic diversity management, reducing genetic drift and inbreeding by minimising kinship among captive individuals (MK strategies; Rudnick and Lacy 2007). A complete pedigree is more informative, in terms of estimating relatedness, than moderate numbers of molecular markers (Baumung and Sölkner 2003; Fernández et al. 2005), and pedigree management is the most accurate way to prevent inbreeding (Santure et al. 2010; Townsend and Jamieson 2013). However, pedigree management does not provide a comprehensive evaluation of genetic diversity when founder relatedness, levels of genetic diversity and the wild origin of captive founders are unknown (Henkel et al. 2012). Molecular methods are therefore increasingly being combined with pedigrees to optimise, assess and monitor captive conservation programs and to augment relatedness estimates among founders (Fienieg and Galbusera 2013). Such analyses usually use DNA profiling with microsatellite markers to estimate standard measures 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 and Jamieson 2013).

Ex-situ breeding programs 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 founders to the wild population (Miller et al. 2010). However, founders are often captured opportunistically or are already present in captivity when a regulated breeding program is established (Hedrick et al. 1997; Russello et al. 2007; Hedrick and Fredrickson 2008; Ivy et al. 2009). Given that it is becoming less acceptable or feasible to establish or augment captive populations with wild individuals (Williams and 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). Also, even if the diversity of the wild population is adequately represented in the founding population, genetic drift and selection may cause

allele frequencies to differ from the wild. This is partly because equalising founder contributions in captive breeding is challenging to implement with some founders becoming underrepresented or even absent (Frankham 2010). Even when founder contributions are equalized, genetic drift still occurs (albeit at a lower rate) because of the randomness of Mendelian segregation. Consequently, captive breeding usually leads to a loss of genetic variation, differentiation from the wild population, and an increase in the frequency of alleles that are deleterious and/or partially recessive in the wild (Frankham 2008).

One approach that may allow this problem to be tackled is via a genetic evaluation of both the wild, captive and/or founder populations (Witzenberger and Hochkirch 2011). Genetic markers can be used to assess the genetic diversity, identify the population structure (Pritchard et al. 2000), and evaluate levels of genetic differentiation (i.e. FST statistics) among wild, founder and captive populations. These analyses are usually carried out using microsatellites, predominantly due to them being relatively easy to develop (Schoebel et al. 2013), having high statistical power per locus (Witzenberger and Hochkirch 2011), and usually being selectively neutral. In contrast, mitochondrial DNA (mtDNA) has been used less frequently than microsatellite markers for informing captive breeding programs (Russello et al. 2007; Benavides et al. 2012). Also, analysis has usually been limited to measures of genetic diversity, rather than investigating geographic origin of founders or representation of wild alleles in captivity (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). This bias towards the use of microsatellites is likely to be due to the fact that, for some species, mtDNA has been shown to feature low diversity in the wild and/or founding individuals, limiting its utility (Hedrick et al. 1997). However, due to the unique inheritance 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 easier to interpret than when using nuclear markers alone (Avise et al. 1984, 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 used both nuclear and mitochondrial DNA to inform breeding management for the okapi (*Okapia johnstoni*), an even-toed ungulate endemic to the Democratic Republic of Congo that is under threat from habitat fragmentation, human encroachment and poaching. It has recently been reclassified from Near Threatened to Endangered by the IUCN (Mallon et al. 2013). Okapi have an ex-situ conservation program that is managed using a well-documented studbook (Leus and Hofman 2012), which lists 173 living okapi in captivity (as of the 1st January 2012, including 91, 59, 14, 7 and 2 individuals in the US, European, Epulu, Japanese and South African breeding programs respectively). 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 and Hofman 2012). Moreover it is currently unknown how genetically representative captive okapi are of wild population genetic diversity and evolutionary history, information that is particularly important in light of its recent reclassification (Mallon et al. 2013). Okapi predominantly occur across central, eastern and northern Democratic Republic of Congo, but also occur at lower density southwest of the Congo River (Stanton et al. 2014a). However no reliable estimates exist for current population size (Mallon et al. 2013). Wild okapi populations are thought to be declining rapidly (Okapi Conservation Workshop 2013; Quinn et al. 2013), which may be incurring a concomitant loss of genetic diversity. Further, 14 okapi at the conservation and research 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, and were the only captive individuals in the Democratic Republic of Congo, highlighting the importance of characterising the distribution of genetic diversity in- and ex-situ for verifying whether the diversity in captivity is as representative as possible of the wild population.

We used 12 microsatellite and 4 mtDNA markers (total of 833 bp) to characterise and compare genetic diversity and structure of wild, founder and captive okapi, including the individuals from the okapi conservation and research station in Epulu. We aimed to investigate whether the founder and captive okapi populations are genetically representative of the wild by comparing genetic structure

using microsatellite loci, and by evaluating the frequency of mitochondrial DNA haplotypes. We evaluate the processes that may have led to the observed pattern, and discuss the in- and ex-situ conservation implications.

Methods

Samples

We analysed a total of 363 samples, comprising 305 samples from the wild (247 dung samples, 44 museum skins and 14 confiscated skin samples), 32 modern day captive individuals (24 blood or fresh muscle tissue and 8 dung samples), and 26 "founder" samples [15 museum tissue (14) and hoof (1) samples, 6 dung samples, 3 bone and 2 tooth samples]. The founder samples are a combination of individuals that were wild-caught and intended for captive breeding (F0), and recent descendants of those F0 founders (mean generation number 0.61, SD 1.02). Museum specimens were sampled, with permission, from the National Center of Scientific Research, Paris, Natural History Museum of Denmark, Copenhagen, the Field Museum of Natural History, Chicago, the Royal Museum for Central Africa, Tervuren and the okapi conservation and research station, Epulu. The captive samples were selected by choosing individuals that represented as many of the founding lineages as possible (the 33 captives used in the present study can trace ancestry back to all but three of the 36 individuals that founded the living captive population on 1st January 2012) using pedigree data to ensure that the genetic diversity of the captive okapi population was widely represented. The captive sample dataset comprised, (i) all of the individuals from the okapi conservation and research station in Epulu, DRC that were not in the founder dataset [n = 8 from a total of 14 individuals in Epulu (6 founders)], (ii) a representative sample of the European captive population (n = 21; current population size of 65 individuals, total living and dead 335), and (iii) four individuals that were from, but not representative of, the captive US population (n = 93).

Lab methodology

DNA was extracted from blood, tissue, bone and teeth samples using a Qiagen Blood and Tissue Extraction Kit [Qiagen GMBH, Germany, following manufacturers instructions (http://www.qiagen.com/resources/resourcedetail?id=6b09dfb8-6319-464d-996c-79e8c7045a50&lang=en)] and from fecal samples using a DNA Stool Mini Kit [Qiagen GMBH, Germany, following manufacturers instructions (http://www.qiagen.com/resources/resourcedetail?id=c8fe97e7-78cc-4275-bbac-72c9b7c3de38&lang=en)]. For all museum samples, DNA was extracted in a dedicated ancient DNA (aDNA; File S1) room. For these samples, some minor modifications were made to the extraction protocol to increase DNA yield (File S1). PCR amplification of 13 microsatellite loci was carried out for all samples following Stanton et al. (2010), but with number of PCR cycles increased to 60 for museum samples (Table S1). Also, locus Oka-11 was ultimately excluded due to low PCR amplification rates. MtDNA PCR amplification used primers (OJ1-5) and conditions described in Stanton et al. (2014b) that amplify a 833 bp fragment of mtDNA comprised of 370 bp of the Cytochrome *b* (Cyt *b*) and 328 bp of the control region (CR), and the complete tRNA-Thr (69 bp) and tRNA-Pro (66 bp) genes. Due to the high number of wild samples from sampling region two (Fig. 1), a subset of 35 samples was randomly chosen from this region as representatives for mtDNA analysis. In total, the study used 69 wild, 26 captive and 12 founder individuals to generate the mtDNA dataset.

Data quality control and measurements

GIMLET v1.3.3 (Valière 2002) was used to calculate allelic dropout (ADO) and false allele (FA) rate for a preliminary subset of 14 okapi fecal samples. GEMINI v1.3.0 (Valière et al. 2002) was subsequently used to estimate the number of PCR repeats required to accurately create a consensus genotype for the full dataset. This analysis indicated that three repetitions 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 fecal samples. We recorded samples as being heterozygous for a given locus if both alleles

appeared at least twice among the four replicates and as homozygous if all the replicates showed identical homozygous profiles. 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. Heterozygote and homozygote classifications were carried out in the same way as for fecal samples, but adjusting proportionally for the number of repeats carried out. 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 ADO and FA rate for (i) dung samples and confiscated skin samples, and (ii) museum samples. There were five samples where only two repeats could be carried out due to insufficient DNA. GIMLET was used to calculate ADO 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 estimate relatedness using two different estimators (Lynch and Ritland 1999; Wang 2002) within each of the sample groupings (below) within the dataset.

Microsatellite loci were tested for Hardy–Weinberg and linkage equilibria using GENEPOP v4.2 (Raymond and Rousset 1995; Rousset 2008), using the founder and three of the wild sample sets separately [sampling regions 1, 2 and 4, excluding sampling region 3 due to low sample size (n = 3)].

Genetic diversity

We calculated a number of different measures of genetic diversity. Observed and expected heterozygosity for microsatellites were calculated using GENALEX (Peakall and Smouse 2012), and significance assessed using *t* tests in R (R Development Core Team 2011). Unbiased allelic richness was calculated using FSTAT v2.9.3 (Goudet 1995). MtDNA haplotype sequences were aligned in Sequencher 4.9 (GeneCodes) and haplotype diversity was calculated using DNASP v5 (Librado and

Rozas 2009). TEMPNET (Prost and Anderson 2011) was used to create a multi-layered spanning network of the complete (833 bp) mtDNA sequence, with the, captive and wild individuals separated into three different layers.

Genetic structure

To describe genetic differentiation between sample groupings, pairwise sample F_{ST} values were calculated and an analysis of molecular variance (AMOVA) carried out using the microsatellite markers for the following sets of samples: (1) Wild (sampling region 1; Fig. 1), (2) Wild (sampling region 2), (3) Wild (sampling region 3) and (4) Wild (sampling region 4), (5) Captive (United States), (6) Captive (Epulu, DRC), (7) Captive (Europe), (8) "Founders", using ARLEQUIN. Wild sampling regions were based on, (i) the results of the Bayesian clustering analysis and, (ii) a previous genetic study on okapi (Stanton et al. 2014b). Overall sample groupings were chosen to jointly investigate genetic differentiation within and among four wild sampling regions, the founders, and the three captive populations.

For the AMOVA, "populations" were designated as the eight previously described sets of samples. The population grouping that best partitioned the genetic variance observed was investigated by testing a total of 13 different hypotheses. These hypotheses were grouped into four different "sets" and the most statistically significant hypothesis was then compared to other hypotheses in the next set. The approach is shown visually in Figs. S1–S4. Set one tested two hypotheses: In the first, samples were grouped into founder, captive and wild samples, and the second was the same as hypothesis one, except captive samples were 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. 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). 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 combining them into a single group.

In order to identify the genetic structure among founder, captive and wild okapi populations mainly from genetic data (without sampling source or geographic information), Bayesian clustering was performed on microsatellite data using STRUCTURE 2.3.4 (Pritchard et al. 2000) and GENELAND (Guillot et al. 2005). STRUCTURE was run with 500,000 Markov chain Monte Carlo iterations, a burn-in of 50,000, correlated allele frequencies and K set at 1–8. Six independent runs were carried out for each K value. The number of clusters was estimated using the method of Evanno et al. (2005). This analysis was also repeated using only F0 founders to investigate the effect of excluding individuals with a generation number of greater than zero on the founder sample grouping. Relatives within each sample grouping ($r > 0.5$, using both the Lynch and Ritland 1999; Wang 2002 estimators) were not included in the STRUCTURE analysis. The analysis was also repeated excluding one individual from each dyad with a relatedness value greater than 0.25 to further investigate the effect of relatives in this analysis (Rodriguez-Ramilo and Wang 2012). STRUCTURE was also run using only the wild and founder datasets with $K = 4$, using sampling information (the four sampling regions), and updating allele frequencies using only the wild samples. This was done to further investigate genetic structure in the wild and to attempt to assign the founder individuals to a part of the wild okapi distribution. Individuals were assigned to a given population if they had greater than or equal to 0.5 probability of assignment to that population (Hobbs et al. 2011). GENELAND was run with $K = 1–10$, 500,000 iterations, uncorrelated allele frequencies and six independent runs. STRUCTURE and GENELAND were also run using only the wild individuals (utilising spatial information) to attempt to detect genetic structuring in the wild okapi population.

The relative contributions of the sampled founders into each of the three captive breeding programs were investigated using the okapi studbook. This was also investigated with respect to the STRUCTURE cluster that founder and captive individuals were assigned to, in order to determine if the pattern of genetic structure observed in captivity is related to the uneven distribution of founders into breeding programs. Significance was assessed using $\chi 2$ tests in R.

Results

Samples

After removal of duplicated samples and those that failed the genotyping quality control check described above, the final microsatellite and mtDNA datasets contained 143 and 107 samples respectively. Sample details are given in Table S1, along with the information about individuals that were excluded from the STRUCTURE analyses due to high relatedness values.

Genotype validation

Genotyping error rates on the full dataset were lower than those found in the preliminary error rate study (methods), indicating that the number of repeats carried out was sufficient. With GIMLET (Valière 2002) the mean ADO 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 ADO 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, only one had an error rate above zero (founding individual, ADO 0.111). Analyses were carried out excluding this individual, but this did not appreciably alter any of the results (data not shown), and the individual was therefore included in the results. Following Bonferroni correction ($p < 4.17 \times 10^{-3}$), three loci (Oka-02, Oka-07 and Oka-09) were found to be out of HWE. However, this result was not consistent between sample sets, with each of those three loci only being out of HWE in at most one of the four sample sets. These minor deviations from HWE are therefore likely due to random variation within sample sets, rather than being related to the loci themselves. Following Bonferroni correction ($p < 1.56 \times 10^{-3}$–7.58×10^{-4}) one pair of loci was found to be in linkage disequilibrium (Oka-09 vs. Oka-12), but again, only within one of the four sample groupings.

Genetic diversity

For microsatellites, genetic diversity was comparable between captive, founder and wild samples, whereas for mtDNA, there was a large number of wild haplotypes that were absent in the founder and

captive samples. 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; *t* test $p > 0.05$ 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 heterozygosity for these groups is shown graphically in Fig. S5. For mtDNA, haplotype diversity was 0.932, 0.894 and 0.783 for wild samples, founder samples and captive samples respectively. TEMPNET was used to visualise the mitochondrial haplotypes in wild, founder and captive okapi samples (Fig. 2). There were 39 haplotypes present in the wild dataset, ten in the founder dataset (of which six were not detected in the wild) and nine in the captive dataset (of which two were not detected in the wild sample).
Figure 2 highlights the haplotypes that have recently been lost from the dataset as a result of the death of the okapi at Epulu (denoted with an "X"; founders n = 5, captive n = 1).

Genetic structure

The molecular variation in our data was best explained when wild sampling region 2, captive Epulu and founder samples were grouped, and the captive US and captive European sample sets were each in a group of their own. Analysis of the thirteen sample group hypotheses (arranged into four "sets", Figs. S1–S4) revealed that the molecular variance in Set One was best explained by considering the captive populations as three separate groups rather than a single unit [among group variation 0 % ($p = 0.887$) vs. 2.24 % ($p = 0.190$)]. Set Two showed that of the sampling regions, molecular variance was most effectively partitioned when the founders were grouped with sampling region two (among group variation 3.11 %, $p = 0.030$). Set Three best explained the molecular variance when the okapi conservation and research station at Epulu were also grouped with the founders and the samples from sampling region two (among group variation 3.43 %, $p = 0.004$). Set four partitioned the most molecular variance when the okapi from sampling regions three and four were grouped together, but separate from sampling region one (among group variation 4.63 %, $p = 0.002$).

Pairwise F_{ST} values between sample sets 1–8 (described above) are given in Table 1. Significant F_{ST} values ranged between 0.016 [$p = 0.005$; between the founding individuals and wild (sampling

region 2) samples] and 0.116 [$p < 0.001$; between captive (Epulu) and captive (Europe) samples; not including the captive (US) samples, due to low sample size]. Other significant F_{ST} values include between wild (sampling region 1) and all captive sample groupings (F_{ST} = 0.093–0.106), and between the founders and wild (sampling region 1) (F_{ST} = 0.070).

When using the full wild-founder-captive dataset, Bayesian clustering showed clear genetic structure among captive sample groupings, and between captive versus wild sample groupings, in particular for the captive European sample set. All individuals were assigned to the same Bayesian clusters across all six independent STRUCTURE runs. The Evanno et al. (2005) method estimated the most appropriate number of genetic clusters in the complete microsatellite dataset to be four (Fig. S6). Excluding all except F0 founders from the founder dataset had little effect on the results (Fig. S7), as did excluding relatives from the founder and captive datasets (Fig. S8). The MK's of our captive samples was 0.104 (max 0.125) and 0.080 (max 0.289) for the US and European sample sets respectively. The STRUCTURE plot for K = 4 is shown in Fig. 3, with samples grouped into the eight sample sets described above. The clustering results for the full wild-founder-captive dataset in GENELAND were broadly the same as the STRUCTURE analysis (Fig. S9), and K was also estimated to be four. The results for the inferred number of clusters for the wild-only dataset in GENELAND (K = 5) is shown in Fig. 1, along with the delineations of the sampling regions. The majority of wild samples belonged to one of three geographically restricted genetic clusters (northwest, northeast and southwest), with another two genetic clusters spread across the range. Using STRUCTURE on the founder and wild samples only, the majority of founder samples were assigned to the same genetic cluster as the individuals from sampling region two (Fig. S10; probability of membership $\geq$0.5). The three individuals from sampling region three were also assigned to this cluster. When using STRUCTURE on the wild samples only, the majority of individuals could not be assigned to a genetic cluster (probability of membership <0.5) for K = 4, but the structure plot for K = 3 (Fig. S11) resembled the relevant section of the structure plot on the full sample set (Fig. 3).

The founder and captive individuals used in the present study for which there was STRUCTURE clustering information are shown in Fig. 4. The European captive sample set is shown to contain

founder lineages predominantly belonging to one cluster (cluster B, from Fig. 3; 83.3 %). The Epulu captive sample entirely contained founder lineages belonging to a second cluster (cluster A, 100 %) and the majority of the US captive samples contained founder lineages from the same cluster (57.1 %). It should be noted however that these US captive samples were only a subset of the entire US captive population, have above-average relatedness (compared to the entire US captive breeding program), and should not be considered representative of the population as a whole.

The okapi studbook (Leus and Hofman 2012) indicates that there is unequal founder representation in the different captive breeding programs [for example individual 381 (Fig. 4) contributed 93.1 % of it's descendants to either the US or Epulu captive populations, and none of its descendants to the European captive population]. Our results demonstrate that there also appears to be unequal representation of founders from each of the population clusters in the different captive populations ($\chi^2 = 6.520$, $p = 0.0384$). Of the eight founders belonging to cluster A, 78.6 % of their descendants (sampled in the present study) also belonged to that cluster, 14.3 % to cluster B and 7.1 % to a third cluster (cluster C, Fig. 3). Of the five founders belonging to cluster B, 33.3 % of their descendants (sampled in the present study) belonged to cluster A, 60.0 % to the cluster B and 6.7 % to cluster C. These results also show that any captive individual is most likely to share the same genetic cluster as its ancestor ($\chi^2 = 9$, $p = 0.0111$), which would be expected, as these Bayesian clustering algorithms are known to group more closely related individuals (Pritchard et al. 2000; Rodriguez-Ramilo and Wang 2012).

Discussion

Our analysis shows that although levels of nuclear genetic diversity are very similar between wild, founder and captive okapi, the number of mitochondrial haplotypes differs greatly, and microsatellite allele frequencies in the captive population are not representative of the wild. In particular, genetic differentiation based on FST values, was much higher comparing the European captive population versus the small subset of US captive individuals used in this study, and comparing the wild sampling

region one versus the founders or any of the captive populations, than between any of the four wild sampling regions. Many other captive populations have also been shown to demonstrate considerable genetic structure (e.g. Guan et al. 2009; Henry et al. 2009; Armstrong et al. 2011; Witzenberger and Hochkirch 2013), but the management implications of this are not always clear. Captive management generally aims to maintain genetic variation both within and among sub-populations, however the appropriate proportions of each are difficult to determine (Wang 2004).

Okapi captive populations were also genetically differentiated from the wild based on Bayesian clustering analysis (STRUCTURE and GENELAND) of microsatellite data. Significant genetic differentiation at nuclear loci between captive and wild populations has been observed in giant pandas (F_{ST} wild vs. wild = 0.075, captive vs. wild = 0.071–0.121; Shen et al. 2009), however the relative difference in genetic differentiation between captive vs. wild and wild vs. wild was much greater in the present study (Table 1). Mitochondrial genetic diversity was greatly reduced in the founders and their captive descendants, with only a small proportion of the haplotypes present in the wild represented in captivity. Also, the number of haplotypes in our founder dataset was reduced by 50 % after removing the founder individuals recently killed in Epulu. Mitochondrial genetic diversity has previously been shown to be considerably reduced in captive populations compared to their wild counterparts (Muñoz-fuentes et al. 2008; McGreevy et al. 2009).

The significant microsatellite differentiation detected between wild okapi sampling regions (Table 1) gives confirmation (using a different set of nuclear markers) to the results of Stanton et al. (2014b), who also detected genetic differentiation between present-day wild okapi samples based on mtDNA and nuclear DNA sequences. The present study also contextualises the okapi captive breeding program with respect to the genetic structure of these wild okapi. Genetic differentiation between pairwise comparisons on the same, as well as opposite, sides of the Congo River (Table 1) and the spatial patterning of genetic clusters (Fig. 1) indicate that factors other than this river [such as the biogeographic history of the Congo Basin, Stanton et al. (2014b)] have played an important role in shaping the present-day genetic diversity of the okapi in the wild. This information can be of use in in-situ conservation when determining priority areas, identifying barriers to movement, and planning

translocations (Pennock and Dimmick 1997; Crandall et al. 2000; Fraser and Bernatchez 2001). It could also be of potential use for identifying the wild origins of founders, useful when planning future translocations from the wild into captivity (Gautschi et al. 2003; Russello et al. 2007).

STRUCTURE (Fig. S10) and AMOVA (Figs. S1–S4) found strong support for a genetic grouping that included the founders of the captive population with the individuals sampled from sampling region two. These results are in agreement with anecdotal reports that most of the founding individuals originated from the okapi faunal reserve, within this sampling region (J. Lukas Pers. Comm). AMOVA also confirmed the results from the Bayesian clustering analysis (Figs. 2, S7) 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 range.

Levels of nuclear genetic diversity were similar between founder, captive and wild okapi population samples (Fig. S5). Other studies of captive populations have found 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. It is important to note that the maintenance of high genetic diversity in this ex-situ population is significant considering the considerable challenges that zoos have faced in breeding okapi in captivity (Gijzen and Smet 1974; Rabb 1978; Bodmer and Rabb 1992), a challenge that is not unique to okapi (Snyder et al. 1996). 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 its well-managed studbook, which minimises inbreeding using MK strategies (Rudnick and Lacy 2007; Leus and Hofman 2012).

Genetic diversity based on mtDNA sequences was, however, considerably reduced in captive samples. Loss of mtDNA diversity is likely to be much more rapid than for nuclear DNA due to the mitochondrial genome having an effective population size approximating to one-quarter that of the

nuclear genome, and is therefore more susceptible to founder effects and genetic drift (Hartl and Clark 1997). The importance of mtDNA diversity in captive populations is less well recognised than nuclear genetic diversity (Hedrick et al. 1997). In principle, more weight should be given to female effective size than male effective size. 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 (Stanton et al. 2014b). If the remaining mtDNA haplotypes are not actively managed in captivity in the future, their rapid loss is likely to continue.

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 hence not be representative of the wild (Henry et al. 2009; Nsubuga et al. 2010; McGreevy et al. 2011). This issue is clearly highlighted in the present study: 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 at microsatellite loci, inferred from STRUCTURE, AMOVA and F_{ST}values, and (ii) a considerable reduction of mtDNA haplotypes in captivity, relative to wild okapi populations. By 'not representative' we imply that although levels of nuclear genetic diversity are comparable, allele frequencies differ markedly between populations. This is important because an increase of alleles in captivity that are rare in the wild (i.e. due to genetic drift) may be detrimental in wild populations (Frankham 2008) and may consequently affect the success of any future reintroductions (Griffith et al. 1989; Fischer and Lindenmayer 2000; Wolf et al. 2002; Jule et al. 2008). It may also indicate an accumulation of genetic load due to relaxed selection, which may ultimately result in inbreeding depression (Boakes et al. 2006). Studies have shown that genetic structure can be introduced rapidly in captive populations due to founder effect (Hu et al. 2007; Armstrong et al. 2011; Witzenberger and Hochkirch 2013), although relatively few use a dataset as comprehensive as the present study, with representative samples from the wild, founders and captives (McGreevy et al. 2009; Shen et al. 2009; Gonçalves da Silva et al. 2010).

The extent to which captive sub-populations should be allowed to become genetically structured has been debated (Wang 2004; Frankham 2008; Leberg and Firmin 2008). Both theory (Kimura and Crow 1963; Robertson 1964) and empirical studies (Frankham 2008) indicate that genetic structure in captivity helps to increase the fitness of a population when it is reintroduced to the wild. The balance between genetic differentiation and gene-flow in the ex-situ okapi population has been carefully managed, and the unequal representation of certain founders in the different captive breeding programs is well known (Leus and Hofman 2012). Figures 3 and S7 appear to show that considerable genetic differentiation exists across parts of the ex-situ population, however the individuals labelled as "Captive US" are not representative of the genetic diversity present in the entire US breeding program. This is evident when the genetic differentiation between captive (US) and captive (Europe) sample sets, is compared between the molecular estimates in the present study (F_{ST} = 0.152) and those based on the studbook alone (F_{ST} = 0.025; current living European and US captive populations). The captive US samples used here are in fact close relatives of the sampled Epulu individuals, and Bayesian clustering algorithms are known to group more closely related individuals (Pritchard et al. 2000; Rodriguez-Ramilo and Wang 2012), explaining their membership to the same genetic cluster (Figs. 3, S7). These findings instead demonstrate the importance of historic gene-flow between Epulu and the US captive population, in light of the death of the 14 okapi individuals at Epulu, and highlight the need for full genetic characterisation of the captive populations outside of Europe.

Our results (Figs. 3, 4, S9) illustrate how captive populations can be subject to founder effect and genetic drift, exacerbated by unequal representation and limited exchange between regions (Leus and Hofman 2012), and leading to differentiated ex-situ subpopulations. 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. To some extent, this appears to be the case for okapi, with the present study showing that genetic differentiation appears to be considerably higher between parts of the captive and wild populations than between geographically distant parts of the wild population. Future studies should sample the US captive population more extensively, to investigate if the genetic differentiation identified in the samples analysed in this study are representative of the US

captive population as a whole. Genetic differentiation may help maintain the total genetic variation of the entire species when the genetic structuring is stable and lasting (Kimura and Crow 1963; Robertson 1964; Wang 2004; Frankham 2008; Leberg and Firmin 2008), but could lead to inbreeding depression, which can be mitigated by gene-flow (Rudnick and Lacy 2007). Maintaining a structured population in conservation can also be important for safekeeping of genetic diversity against disasters (e.g. the situation in Epulu). Determining and maintaining a suitable level of genetic structure is one of the main challenges of ex-situ management (Theodorou and Couvet 2015). The present study provides an important case study to help understand how the interaction between these population genetic processes can affect small, artificially bred populations.

Acknowledgments

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 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 the Frankfurt Zoological Society, the Wildlife Conservation Society, the Lukuru Foundation/TL2 project, the Zoological Society of London (ZSL), Marwell Zoo 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, 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 GIC (the Okapi Conservation Project), the US Fish and Wildlife Service, the UK's Darwin Initiative and the Mohamed bin Zayed Species Conservation Fund. Finally, the CRC gratefully acknowledges the structural support of the Flemish Government.

References

Armstrong E, Leizagoyen C, Martínez AM, González S, Delgado JV, Postiglioni A (2011) Genetic structure analysis of a highly inbred captive population of the African antelope *Addax nasomaculatus*. Conservation and management implications. Zoo Biol 30:399–411

Avise JC, Neigel JE, Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. J Mol Evol 20:99–105

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst 18:489–522

Baumung R, Sölkner J (2003) Pedigree and marker information requirements to monitor genetic variability. Genet Sel Evol 35:369–383

Benavides E, Russello M, Boyer D, Wiese RJ, Kajdacsi B, Marquez L, Garrick R, Caccone A (2012) Lineage identification and genealogical relationships among captive Galápagos tortoises. Zoo Biol 31:107–120

Boakes EH, Wang J, Amos W (2006) An investigation of inbreeding depression and purging in captive pedigreed populations. Heredity 98:172–182

Bodmer RE, Rabb GB (1992) *Okapia johnstoni*. Mamm Species 422:1–8

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15:290–295

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Fernández J, Villanueva B, Pong-Wong R, Toro MA (2005) Efficiency of the use of pedigree and molecular marker information in conservation programs. Genetics 170:1313–1321

Fienieg ES, Galbusera P (2013) The use and integration of molecular DNA information in conservation breeding programmes: a review. J Zoo Aquar Res 1:44–51

Fischer J, Lindenmayer DB (2000) An assessment of the published results of animal relocations. Biol Conserv 96:1–11

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). Mol Ecol 16:4039–4050

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. Mol Ecol 17:325–333

Frankham R (2010) Where are we in conservation genetics and where do we need to go? Conserv Genet 11:661–663

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

Fraser D, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol Ecol 10:2741–2752

Gautschi B, Jacob G, Negro JJ, Godoy JA, Müller JP, Schmid B (2003) Analysis of relatedness and determination of the source of founders in the captive bearded vulture, *Gypaetus barbatus*, population. Conserv Genet 4:479–490

Geyer CJ, Ryder OA, Chemnick LG, Thompson EA (1993) Analysis of relatedness in the California condors, from DNA fingerprints. Mol Biol Evol 10:571–589

Gijzen A, Smet S (1974) Seventy years of okapi, *Okapia johnstoni* (Sclater, 1901). Acta Zoologica et Pathologica Antverpiensia 59:1–90

Gonçalves da Silva A, Lalonde DR, Quse V, Shoemaker A, Russello MA (2010) Genetic approaches refine ex situ lowland tapir (*Tapirus terrestris*) conservation. J Hered 101:581–590

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Griffith B, Scott MJ, Carpenter JW, Reed C (1989) Translocation as a species conservation tool: status and strategy. Science (Washington) 245:477–480

Guan T-L, Zeng B, Peng Q-K, Yue B-S, Zou F-D (2009) Microsatellite analysis of the genetic structure of captive forest musk deer populations and its implication for conservation. Biochem Syst Ecol 37:166–173

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. Mol Ecol Notes 5:712–715

Haig SM, Ballou JD, Casna NJ (1992) Genetic identification of kin in Micronesian kingfishers. J Hered 86:423–431

Hartl DL, Clark AG (1997) Principles of population genetics. Sinauer associates, Sunderland

Hedrick PW, Fredrickson RJ (2008) Captive breeding and the reintroduction of Mexican and red wolves. Mol Ecol 17:344–350

Hedrick PW, Miller PS, Geffen E, Wayne RK (1997) Genetic evaluation of the three captive Mexican wolf lineages. Zoo Biol 16:47–69

Henkel JR, Jones KL, Hereford SG, Savoie ML, Leibo SP, Howard JJ (2012) Integrating microsatellite and pedigree analyses to facilitate the captive management of the endangered Mississippi sandhill crane (*Grus canadensis pulla*). Zoo Biol 31:322–335

Henry P, Miquelle D, Sugimoto T, McCullough DR, Caccone A, Russello MA (2009) In situ population structure and ex situ representation of the endangered Amur tiger. Mol Ecol 18:3173–3184

Hobbs GI, Chadwick EA, Bruford MW, Slater FM (2011) Bayesian clustering techniques and progressive partitioning to identify population structuring within a recovering otter population in the UK. J Appl Ecol 48:1206–1217

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 Rumin Res 67:252–256

Ivy JA, Miller A, Lacy RC, Dewoody JA (2009) Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (*Macropus parma*). J Hered 100:441–454

Jule KR, Leaver LA, Lea SEG (2008) The effects of captive experience on reintroduction survival in carnivores: a review and analysis. Biol Conserv 141:355–363

Khan HA, Arif IA, Shobrak M, Homaidan AAA, Farhan AHA, Sadoon MA (2011) Application of mitochondrial genes sequences for measuring the genetic diversity of Arabian oryx. Genes Genet Syst 86:67–72

Kimura M, Crow JF (1963) The measurement of effective population number. Evolution 17:279–288

Konovalov DA, Manning C, Henshaw MT (2004) KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. Mol Ecol Notes 4:779–789

Kozfkay CC, Campbell MR, Heindel JA, Baker DJ, Kline P, Powell MS, Flagg T (2008) A genetic evaluation of relatedness for broodstock management of captive, endangered Snake River sockeye salmon, *Oncorhynchus nerka*. Conserv Genet 9:1421–1430

Lacy RC (2013) Achieving true sustainability of zoo populations. Zoo Biol 32:19–26

Leberg PL, Firmin BD (2008) Role of inbreeding depression and purging in captive breeding and restoration programmes. Mol Ecol 17:334–343

Lesobre L, Lacroix F, Caizergues A, Hingrat Y, Chalah T, Jalme MS (2010) Conservation genetics of Houbara Bustard (*Chlamydotis undulata undulata*): population structure and its implications for the reinforcement of wild populations. Conserv Genet 11:1489–1497

Leus K, Hofman S (2012) The international studbook for the Okapi (*Okapia johnstoni*)—12 Oct. The Royal Zoological Society of Antwerp, Antwerp

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152:1753–1766

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. Accessed 26 Nov 2013

McGreevy TJ, Dabek L, Gomez-Chiarri M, Husband TP (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 Biol 28:183–196

McGreevy TJ, Dabek L, Husband TP (2011) Genetic evaluation of the association of zoos and aquariums Matschie's tree kangaroo (*Dendrolagus matschiei*) captive breeding program. Zoo Biol 30:636–646

Miller W, Wright SJ, Zhang Y, Schuster SC, Hayes VM (2010) Optimization methods for selecting founder individuals for captive breeding or reintroduction of endangered species. Pacific Symposium on Biocomputing

Muñoz-fuentes V, Green AJ, Sorenson MD (2008) Comparing the genetics of wild and captive populations of White-headed Ducks *Oxyura leucocephala*: consequences for recovery programmes. Ibis 150:807–815

Nsubuga AM, Holzman J, Chemnick LG, Ryder OA (2010) The cryptic genetic structure of the North American captive gorilla population. Conserv Genet 11:161–172

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

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539

Pennock DS, Dimmick WW (1997) Critique of the evolutionarily significant unit as a definition for "Distinct Population Segments" under the U.S. Endangered Species Act. Conserv Biol 11:611–619

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

Prost S, Anderson CNK (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. Methods Ecol Evol 2:663–667

Quinn A, Queslin E, Kümpel N (2013) Okapi (*Okapia johnstoni*): conservation status review. Zoological Society of London, 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. http://www.R-project.org/

Rabb GB (1978) Birth, early behavior and clinical data on the okapi. Acta zoologica et pathologica antverpiensia 71:93–105

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Robertson A (1964) The effect of non-random mating within inbred lines on the rate of inbreeding. Genet Res 5:164–167

Rodriguez-Ramilo ST, Wang J (2012) The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. Mol Ecol Res 12:873–884

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Res 8:103–106

Rudnick JA, Lacy RC (2007) The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. Conserv Genet 9:1439–1450

Russello MA, Hyseni C, Gibbs JP, Cruz S, Marquez C, Tapia W, Velensky P, Powell JR, Caccone A (2007) Lineage identification of Galápagos tortoises in captivity worldwide. Anim Conserv 10:304–311

Santure AW, Stapley J, Ball AD, Birkhead TR, 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. Mol Ecol 19:1439–1451

Schoebel CN, Brodbeck S, Buehler D, Cornejo C, Gajurel J, Hartikainen H, Keller D, Leys M, Říčanová S, Segelbacher G, Werth S, Csencsics D (2013) Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. J Evol Biol 26:600–611

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. Mol Ecol 18:1061–1070

Snyder NFR, Derrickson SR, Beissinger SR, Wiley JW, Smith TB, Toone WD, Miller B (1996) Limitations of captive breeding in endangered species recovery. Conserv Biol 10:338–348

Stanton DWG, Penfold LM, Zhan X, Bruford MW (2010) Microsatellite loci for the okapi (*Okapia johnstoni*). Conserv Genet Res 2:337–339

Stanton DWG, Hart J, Vosper A, Kumpel NF, Wang J, Ewen JG, Bruford MW (2014a) Non-invasive genetic identification confirms the presence of the endangered okapi *Okapia johnstoni* south-west of the Congo River. Oryx. doi:10.1017/S0030605314000593

Stanton DWG, Hart J, Galbusera P, Helsen P, Shephard J, Kümpel NF, Wang J, Ewen JG, Bruford MW (2014b) Distinct and diverse: range-wide phylogeography reveals ancient lineages and high genetic variation in the endangered okapi (*Okapia johnstoni*). PLoS One 9:e101081

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

Theodorou K, Couvet D (2015) The efficiency of close inbreeding to reduce genetic adaptation to captivity. Heredity 114:38–47

Townsend SM, Jamieson IG (2013) Molecular and pedigree measures of relatedness provide similar estimates of inbreeding depression in a bottlenecked population. J Evol Biol 26:889–899

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol 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. Mol Ecol Notes 2:83–86

Wang J (2002) An estimator for pairwise relatedness using molecular markers. Genetics 160:1203–1215

Wang J (2004) Application of the one-migrant-per-generation rule to conservation and management. Conserv Biol 18:332–343

Williams SE, Hoffman EA (2009) Minimizing genetic adaptation in captive breeding programs: a review. Biol Conserv 142:2388–2400

Witzenberger KA, Hochkirch A (2011) Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animals. Biodivers Conserv 20:1843–1861

Witzenberger KA, Hochkirch A (2013) Evaluating ex situ conservation projects: genetic structure of the captive population of the Arabian sand cat. Mamm Biol 78:379–382

Wolf MC, Griffith B, Reed C, Temple SA (2002) Avian and mammalian translocations: update and reanalysis of 1987 survey data. Conserv Biol 10:1142–1154

Fig. 1

Population genetic cluster assignment of samples using GENELAND. Sampling regions are for description in the text and for arrangement of samples for analyses, and were based on the GENELAND population assignment delineations. The *X* and *Y-axes* state the longitude and latitude values

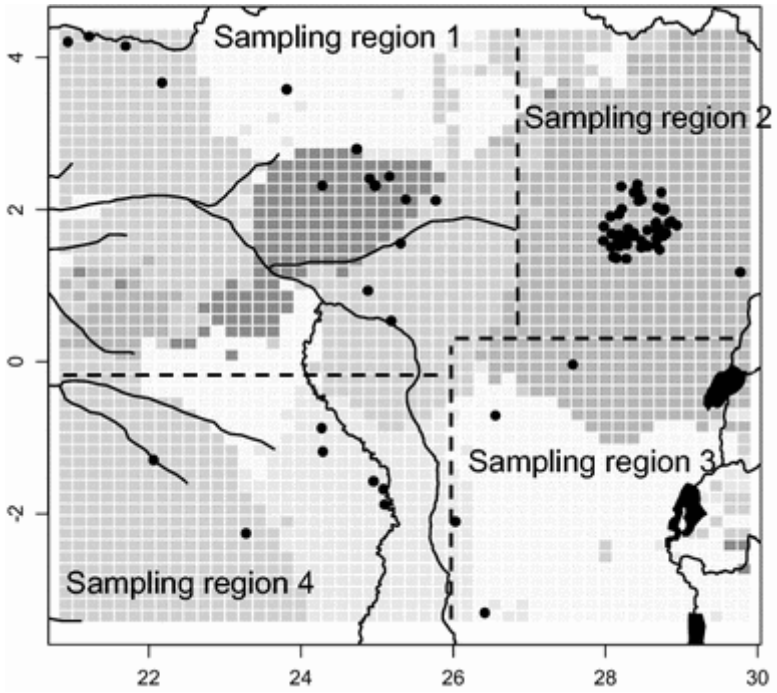

Fig. 2

3D network drawn in TempNet (Prost and Anderson 2011) of the mitochondrial DNA haplotype changes between wild, founder and captive okapi. Elipse size is proportional to the number of samples with that haplotype. *Small open circles* represent unsampled haplotypes for that particular layer and *small black circles* represent unsampled haplotypes across all layers. *Vertical lines* join haplotypes shared among all three layers. *Dotted lines* are used when a connection includes a haplotype unsampled in that particular layer. Haplotypes found only in okapi that were among the individuals killed in Epulu are denoted by an “X”

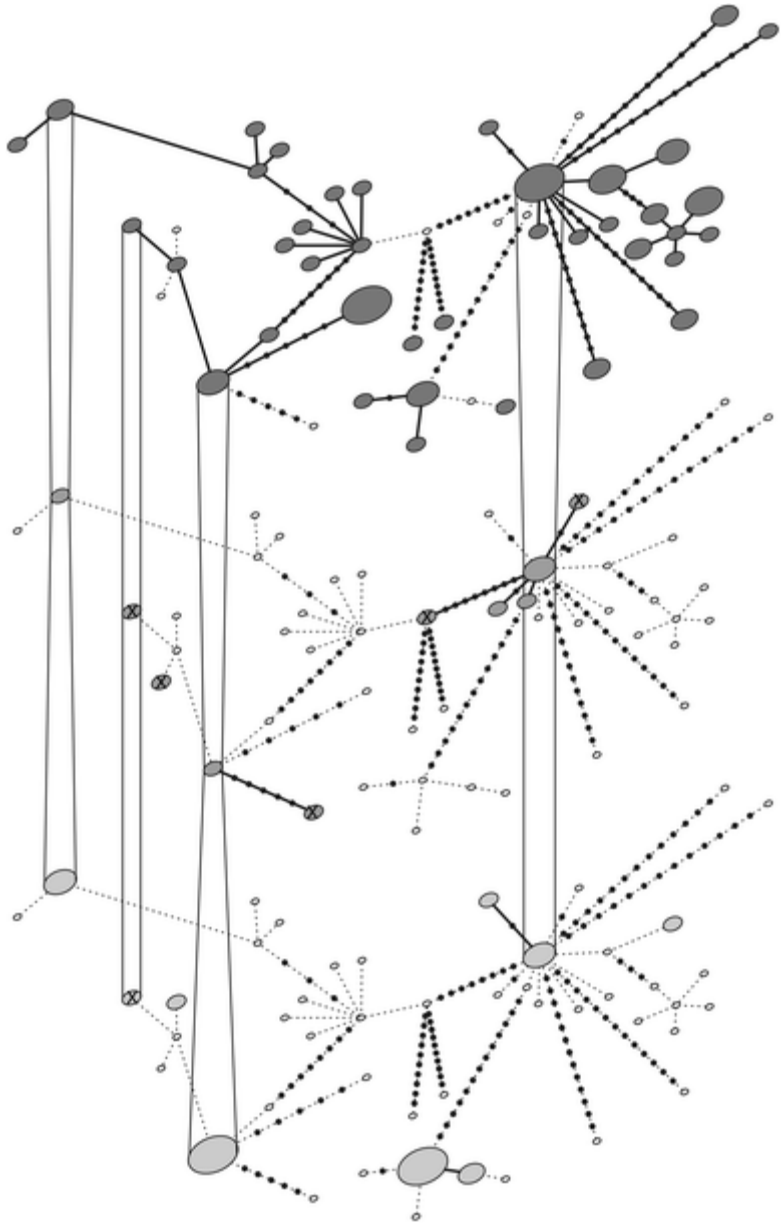

Fig. 3

Structure plot of founder, captive and wild okapi samples for K = 4. Numbers correspond to: (5) Captive (United States; N = 4), (6) Captive (Epulu, DRC; N = 8), (7) Captive Europe (N = 18), (8) “Founders” (N = 16), (1) Wild (sampling region 1; Fig. 1; N = 17), (2) Wild (sampling region 2; N = 54), (3) Wild (sampling region 3; N = 3) and (4) Wild (sampling region 4; N = 6)

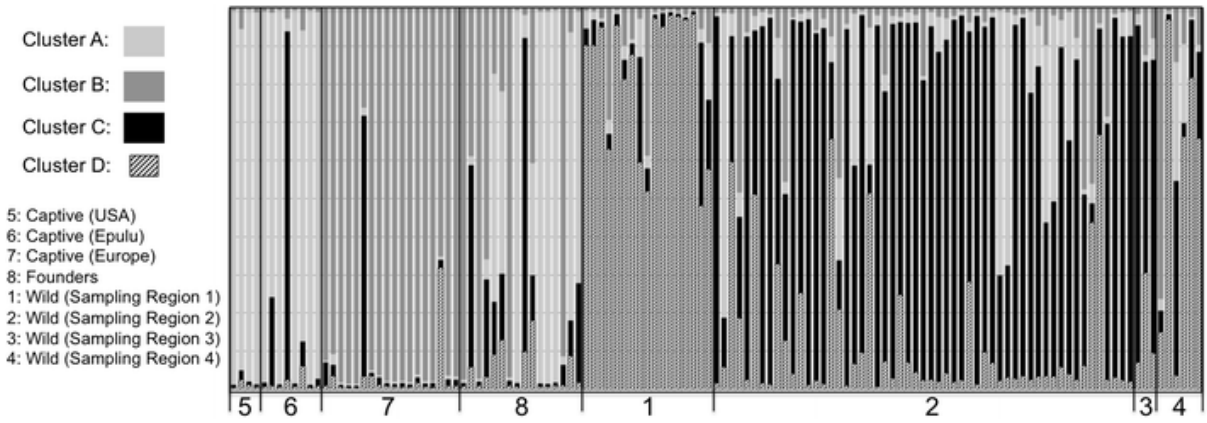

Fig. 4

Pedigree showing okapi founders and captives with STRUCTURE groups. 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 Fig. 3. *Straight lines* connect ancestors/descendants, and are not necessarily first order relatives. Only the data is shown for individuals that could be assigned to a population using STRUCTURE

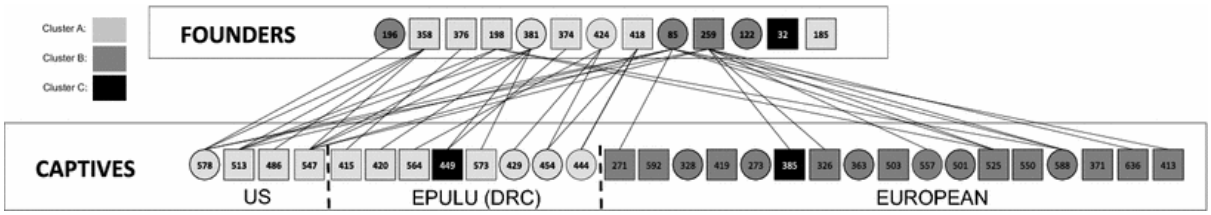

Table 1

Table of pairwise F_{ST} values based on microsatellite data

	Captive (US)	**Captive (Epulu)**	**Captive (Europe)**	**Founders**	**Wild (sampling region 1)**	**Wild (sampling region 2)**	**Wild (sampling region 3)**	**Wild (sampling region 4)**
Captive (US)	0.00000							
Captive (Epulu)	0.04138^{NS}	0.00000						
Captive (Europe)	0.15160***	0.11597***	0.00000					
Founders	0.03595 ^{NS}	0.01404	0.03603**	0.00000				
Wild (sampling region 1)	0.10639**	0.10490***	0.09282***	0.07035***	0.00000			
Wild (sampling region 2)	0.06870***	0.02741**	0.05679***	0.01637**	0.04628***	0.00000		
Wild (sampling region 3)	0.07303 ^{NS}	0.04522 ^{NS}	0.00465 ^{NS}	0.00606 ^{NS}	0.04522 ^{NS}	−0.02363 ^{NS}	0.00000	
Wild (sampling region 4)	0.10677*	0.08104**	0.05832**	0.03753**	0.05576**	0.04324***	−0.01489 ^{NS}	0.00000

The captive US sample set was composed of four individuals that are not representative of this captive population as a whole

NS not significant

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$