Molecular Ecology (1999) **8**, 1105–1115

Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: evidence for male-biased gene flow

SILVESTER NYAKAANA*† and PETER ARCTANDER*
Zoological Institute, Department of Population Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark, †Institute of Environment and Natural Resources, Makerere University, PO Box 7298, Kampala, Uganda

Abstract

A drastic decline has occurred in the size of the Uganda elephant population in the last 40 years, exacerbated by two main factors; an increase in the size of the human population and poaching for ivory. One of the attendant consequences of such a decline is a reduction in the amount of genetic diversity in the surviving populations due to increased effects of random genetic drift. Information about the amount of genetic variation within and between the remaining populations is vital for their future conservation and management. The genetic structure of the African elephant in Uganda was examined using nucleotide variation of mitochondrial control region sequences and four nuclear microsatellite loci in 72 individuals from three localities. Eleven mitochondrial DNA (mtDNA) haplotypes were observed, nine of which were geographically localized. We found significant genetic differentiation between the three populations at the mitochondrial locus while three out of the four microsatellite loci differentiated KV and QE, one locus differentiated KV and MF and no loci differentiated MF and QE. Expected heterozygosity at the four loci varied between 0.51 and 0.84 while nucleotide diversity at the mitochondrial locus was 1.4%. Incongruent patterns of genetic variation within and between populations were revealed by the two genetic systems, and we have explained these in terms of the differences in the effective population sizes of the two genomes and male-biased gene flow between populations.

Keywords: African elephant, control region, genetic differentiation, male-biased gene flow, microsatellites, population structure

Received 28 September 1998; revision received 28 January 1999; accepted 28 January 1999

Introduction

The number of elephants in Uganda has drastically declined in the last four decades from an estimated 60 000 (Laws *et al*. 1970) to the current estimated population size of 1200–2000 individuals (Olivier 1991b). Out of the approximately 9000 elephants counted in Murchison falls national park in 1973, only 28 remained by 1991 (Olivier 1991b). By 1980, only 150 elephants were counted in Queen Elizabeth national park (Eltringham & Malpas 1980) out of the 3700 elephants censused in 1972 (Eltringham 1977). On the other hand, 397 elephants still remained in Kidepo Valley national park by 1982; a figure that did not differ significantly from the population size of 470 in 1971. However, recent population censuses suggest that elephant populations are recovering, although slowly (e.g. Olivier *et al*. 1989).

A significant reduction in the size of their home range as a consequence of an increase in the size of the human population (Brooks & Buss 1962) and poaching for ivory (Parker & Douglas-Hamilton 1976) have been the major contributing factors to this decline. In 1929, the elephant range constituted approximately 70% of all the land in Uganda but, by 1959, only about 17% of the land was occupied (Brooks & Buss 1962). Elephant home ranges vary in size between 15 and 52 km^2, especially when food and water are abundant (Hamilton 1973), or they can

Correspondence: S. Nyakaana. Fax: + 256 41 530134; E-mail: muienr@imul.com

© 1999 Blackwell Science Ltd

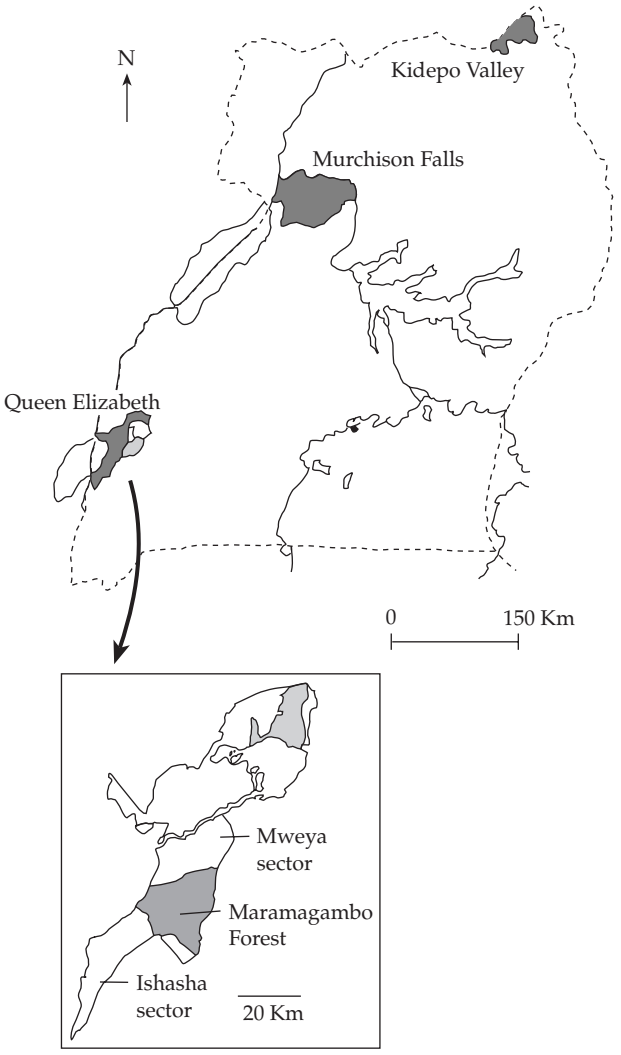

Fig. 1 Map of Uganda showing the three National Parks from which samples were obtained.

extend up to 1580 km² during the dry season (Leuthold & Sale 1973). Currently, the range of the surviving populations is restricted to the boundaries of mainly three protected areas, with very little or no gene flow between them. These areas are Kidepo Valley national park (1340 km²), Murchison Falls national park (3840 km²) and Queen Elizabeth national park (1978 km²) (see Fig. 1). However, the intensity of poaching has varied from one national park to another, with Murchison Falls and Queen Elizabeth national parks experiencing severe decline in numbers while numbers have remained relatively stable in Kidepo Valley national park. The latest population censuses in all the three national parks suggest that elephant populations are recovering, although slowly (Olivier *et al*. 1989; Olivier 1991b).

Since 1986, efforts have been made to restore the protected areas and curb the incidence of poaching. Effective conservation and management of the remaining elephant populations entails an assessment of the amount of genetic diversity present within and between the surviving populations.

We have used both nucleotide sequence variation within the control region (D-loop) of the mitochondrial genome and variation at four nuclear microsatellite loci to assess the population structure of the African elephant in Uganda.

This is because in animals with relatively complex social behaviour, sex-biased dispersal or philopatry, patterns of population structure based on uniparentally inherited markers, e.g. mitochondrial DNA (mtDNA) may differ from those based on biparentally inherited nuclear markers (Karl & Avise 1992). The high mutation rate of these markers makes them ideal for analysis of genetic differentiation between populations (e.g. Bruford & Wayne 1993; Paetkau & Strobeck 1994; 1996b; Arctander *et al*. 1996a; Stanley *et al*. 1996; Palsbøll *et al*. 1997; Clegg *et al*. 1998; Simonsen *et al*. 1998). This study was designed to address the following questions:

(i) How much genetic variation is still remaining in the three main populations of elephants in Uganda?

(ii) To what extent are the populations in Uganda genetically subdivided?

(iii) Do mitochondrial and nuclear markers recapitulate the same genetic scenario?

Materials and methods

Sampling

Skin biopsy samples were collected from the three main national parks in Uganda, namely, Queen Elizabeth National Park (QE; 11 individuals from Ishasha sector and 31 individuals from Mweya sector), Kidepo Valley National Park (KV; $n = 21$) and Murchison Falls National Park ($n = 9$, all solitary males) (Fig. 1). One sample obtained from Ghana (West Africa) was used as an outgroup comparison. Elephant social groups are matrilineal; to avoid sampling closely related individuals, individuals were randomly sampled from different family groups. Tissue samples were obtained by biopsy darting free-ranging elephants with a crossbow (Karesh *et al*. 1987), and sometimes from dry skin or scrapings obtained from carcass remains. DNA for three individuals from KV and two individuals from QE was extracted from fresh dung collected immediately after it was seen being produced by the individuals in question.

The outer layer of fresh faecal boluses contained sufficient DNA for subsequent analysis. All samples were preserved in 25% dimethylsulphoxide (DMSO) saturated with sodium chloride (Amos & Hoelzel 1991) and stored

at room temperature in the field and at −80 °C in the laboratory.

DNA extraction

DNA was extracted from both skin biopsies and faecal samples according to standard methods of digestion with proteinase K in extraction buffer (0.1 M NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and in the presence of 10% sodium dodecyl sulphate (SDS), followed by phenol–chloroform purification and ethanol precipitation (Sambrook *et al*. 1989).

mtDNA amplification and sequencing

Approximately 600 bp of the 5′ end segment of the control region was PCR amplified using primers *Laf CR1* (5′-GTATAAGACATTACAATGGTC-3′) located in the $tRNA^{Pro}$ gene, and *Laf CR2* (5′-AGATGTCTTATTTAAGAGGA-3′) located in the first conserved sequence block (CSB1) of the control region (Saccone *et al*. 1991). Amplifications were carried out in 20 µL reaction volumes containing 2–5 ng of genomic DNA, 20 pmol of each of the primers, 2 µL of 10× PCR reaction buffer (Boehringer Mannheim GmbH), 200 µM dNTPs and 0.4 units of *Taq* polymerase (Boehringer Mannheim GmbH). The following cycling parameters were used: one cycle of denaturation at 94 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 46 °C for 2 min and extension at 72 °C for 3 min. Single-strand DNA template for sequencing was generated by asymmetric amplification (Gyllensten & Erlich 1988) in 50 µL reaction volumes containing 2–5 ng of genomic DNA, 0.5 pmol of the sequencing primer, 50 pmol of the second primer, 5 µL of 10× PCR reaction buffer (Boehringer Mannheim GmbH), 500 µM dNTPs and 1 unit of *Taq* polymerase (Boehringer Mannheim GmbH).

All amplifications were performed on a Hybaid thermocycler and comprised 37 cycles of denaturing at 94 °C for 1 min, annealing at 48 °C for 2 min and extension at 72 °C for 1 min. Single-stranded DNA was cleaned using the QIAquick PCR purification kit 250 (Qiagen) and sequenced by the dideoxy chain-termination method (Sanger *et al*. 1977) using the sequenase kit version 2.0 (United States Biochemical), [α^{35}S]-dATP (Amersham) and primers complementary to the template. The sequencing reaction products were electrophoresed in a 6% polyacrylamide/7 M urea gel. The gel was fixed, dried and visualized by autoradiography on a Kodak film for 24–48 h.

Sequence analyses

Population structure. Sequences were aligned by eye using the program SeqApp version 1.9 (Gilbert 1993). Sequence divergences between haplotypes, nucleotide diversity within populations (Nei 1987; eqn 10.5) and net sequence divergence between populations (per cent sequence divergence between populations corrected for intrapopulation genetic diversity) were estimated with the program POPSTR version 1.2 (H. R. Siegismund, unpublished). Population subdivision was quantified and tested with the sequence statistic, K_{ST} (Hudson *et al*. 1992) and the F_{ST} statistic (Wright 1951) computed from the haplotype frequencies using the program ARLEQUIN version 1.1 (Schneider *et al*. 1997). K_{ST} is an index of population subdivision representing the proportion of the total nucleotide diversity attributable to genetic differences between different populations. K_{ST} was computed as $K_{ST} = 1 - (K_S/K_T)$ where K_S = weighted average number of differences between the sequences in the subpopulations and K_T = average number of differences between two sequences in the total sample. The statistical significance of K_{ST} estimates was found with Monte Carlo simulations (Hudson *et al*. 1992). A consensus phylogeny of the mtDNA haplotypes was estimated with neighbour joining (based on the HKY85 model) and maximum parsimony algorithms implemented in PAUP* version 4.0d64 (PPC, test) (Swofford 1998). A haplotype of an elephant from Ghana (West Africa) was used as an outgroup. The statistical support of each node was estimated using 100 bootstrap replicates of the data set.

Microsatellite loci analysis

Four polymorphic loci (*Laf MS01*, *Laf MS02*, *Laf MS03*, and *Laf MS04*) isolated from the African elephant genome were used (Nyakaana & Arctander 1998). The loci were amplified in 10 µL reaction volumes containing 25–50 ng of DNA, 1.5 mM $MgCl_2$, 10 mM Tris, 50 mM KCl, 200 µM of each dNTP, 2 pmol of each primer and 0.4 units of *Taq* DNA polymerase. Depending on each locus, 32–35 cycles of denaturation at 94 °C for 20 s, annealing at 48–54 °C for 20 s and extension at 72 °C for 20 s were used. The PCR products were separated on a 4% polyacrylamide gel on an ABI model 377 DNA sequencer.

Fisher's exact tests for genotypic disequilibrium were carried out using a markov chain method (Guo & Thompson 1992) implemented in GENEPOP version 3.1b (Raymond & Rousset 1997). Genotype distributions were tested for departure from Hardy–Weinberg proportions with the complete enumeration algorithm of Louis & Dempster (1987) for loci with up to four alleles or the Markov chain method (Guo & Thompson 1992) for loci with more than four alleles. The significance of the observed *P*-values was determined by comparing them with Levene's probability distribution for multiple alleles and large sample sizes (Levene 1949). One thousand

		10	20	30		Populations	
		11112	2222222222	2222233333	3		
		123438890	1223333334	5589901355	5		
		1019018956	6131234567	0490571101	8	KV	MF	QE
1	KV2	AAA-TTCTGT	TAGCTAACCG	CCCTGATCTG	C	7	–	–
2	KV4	....	......G...	....	.	7	3	1
3	KV7	....	...TCGGT..	...CA.C...	T	4	–	–
4	KV8	.G.ACC.CAC	..A...GTTA	TT..AG.TCA	.	1	1	–
5	KV17	....	......G...	.....G....	.	1	–	–
6	KV28	....	......G...	.TT.......	.	1	–	–
7	MF1	.GGACC.CAC	..A...GTTA	TT..AG.TCA	.	–	3	–
8	MF5	GGGACC.CAC	..A...GTTA	TT..AG.TCA	.	–	2	–
9	QE4	.G.ACCTCAC	.GA...GTTA	TT..AG.TCA	.	–	–	2
10	QE12	....	......G.T.	....A.....	.	–	–	38
11	QE13	...ACCTC..	C.....GTTA	TT..A..TCA	.	–	–	1
Number of bases		2221222222	2222222222	2222222222	2 Sum	21	9	42

Fig. 2 Polymorphic sites observed in a 400 bp segment of the D-loop region from 69 elephants. The vertical numbers indicate the positions relative to the reference sequence KV2. A hypen (-) represents a deletion introduced to optimise alignment. The haplotype sequences have been submitted to GenBank under Accession nos AF106203–AF106213.

dememorizations, 100 batches and 1000 iterations were used for each population. Sequential Bonferroni corrections were used to compute the critical significance levels for all simultaneous statistical tests (Rice 1989). Allele frequencies at each locus were also determined using the same program. Genetic diversity in each population was quantified as the total number of alleles at each locus and over all loci and as observed heterozygosity (H_O) and expected heterozygosity (H_E) at each locus. The unbiased expected heterozygosity at each locus in every population was estimated as $H_E = 2n(1 - \sum pi^2)/(2n - 1)$, where pi is the frequency of each of the alleles at a locus and n is the number of individuals sampled (Nei 1987; p. 178, eqn 84). We tested for population genetic differentiation using three different approaches.

(i) The unbiased multilocus R_{ST} statistic was computed using the program RSTCALC (Goodman 1997). It is based on the stepwise mutation model (Ohta & Kimura 1973; Shriver *et al*. 1993). According to this model, mutations at microsatellite loci are considered to occur at relatively high rates and to involve addition or subtraction of a single or a small number of repeat units for a given allele with equal probability. UR_{ST} values represent the fraction of the total variance of allele sizes which is due to genetic differences between populations. The statistical significance of UR_{ST} values was determined by comparing the observed UR_{ST} with the distribution of UR_{ST} over 2000 Monte Carlo simulations. UR_{ST} values were considered to be significant when the probability of obtaining the observed or a more extreme value was less than 0.05 in 2000 Monte Carlo simulations. UR_{ST} values range from 0 (signifying no differentiation) to 1 (signifying complete differentiation).

(ii) Differences in microsatellite allele frequencies between populations at each locus were evaluated using Fisher's exact test (Raymond & Rousset 1995) available in GENEPOP version 3.1b (Raymond & Rousset 1997).

(iii) Estimates of Wright's fixation index, F_{ST} (Wright 1951; Weir & Cockerham 1984) were computed using ARLEQUIN version 1.1 (Schneider *et al*. 1997).

Nei's unbiased genetic distances, D (Nei 1978), were computed for all pairwise comparisons using the program POPGENE version 1.21 (Yeh *et al*. 1997). The program NEIGHBOUR in the PHYLIP version 3.5c package (Felsenstein 1993) was used to estimate a population tree from these distances with the neighbour-joining algorithm.

Results

mtDNA sequence variation

A 400-bp segment of the 5′ end of the control region was sequenced for 72 individuals obtained from three localities in Uganda. Thirty-one sites (7.8%) were polymorphic, defining 11 haplotypes (Fig. 2). Apart from one insertion/deletion event at nucleotide position 39, all polymorphic sites were transition substitutions, 12 of which were purine and 18 were pyrimidine transitions. The largest differences were found in comparisons of haplotypes from Murchison Falls (MF) and those from other localities except KV8, QE4 and QE13 (Table 1). These three haplotypes were more similar to Murchison falls haplotypes compared to the other haplotypes in the populations where they were found. The distribution of the haplotypes suggested genetic differentiation between the populations as six out of the eight haplotypes that were observed in more than one individual were locality specific (see Fig. 2). Six of the haplotypes were found in KV, four in MF and four in QE. Haplotype KV4 was found in all the three localities while KV8 was shared

Table 1 Kimura's transition distances between the haplotypes are below the diagonal and number of transitions between the haplotypes are above the diagonal

	KV2	KV4	KV7	KV8	KV17	KV28	MF1	MF5	QE4	QE12	QE13
KV2		1	9	18	2	3	19	20	20	3	15
KV4	0.0025		8	17	1	2	18	19	19	2	14
KV7	0.0230	0.0204		21	9	10	22	23	23	8	18
KV8	0.0472	0.0444	0.0555		16	17	1	2	2	15	7
KV17	0.0050	0.0025	0.0230	0.0417		3	17	18	18	3	15
KV28	0.0076	0.0050	0.0256	0.0444	0.0076		18	19	19	4	14
MF1	0.0499	0.0472	0.0583	0.0025	0.0444	0.0472		1	3	16	8
MF5	0.0527	0.0499	0.0611	0.0050	0.0472	0.0499	0.0025		4	17	9
QE4	0.0527	0.0499	0.0611	0.0050	0.0472	0.0499	0.0076	0.0101		17	7
QE12	0.0076	0.0050	0.0204	0.0390	0.0076	0.0101	0.0417	0.0444	0.0444		12
QE13	0.0390	0.0363	0.0472	0.0178	0.0390	0.0363	0.0204	0.0230	0.0178	0.0309	

Table 2 K_{ST}, F_{ST}, and per cent sequence divergence (D) between populations based on mtDNA analysis

Comparison	K_{ST}	F_{ST}	D
KV and MF	0.30***	0.24***	1.5
KV and QE	0.19***	0.67***	0.4
MF and QE	0.32***	0.75***	1.4
Total	0.32***	—	—

*** Significant at 0.1% level.

between KV and MF. Haplotype QE1 was the most frequent, occurring in 53% of all the individuals analysed but nevertheless restricted to QE.

Nucleotide diversity (π, Nei 1987 eqn 10.5) in the total sample was $1.4 \pm 0.1\%$ but it varied significantly between populations, ranging from $0.5 \pm 0.1\%$ in QE to $2.4 \pm 0.6\%$ in MF, representing an approximate fivefold difference. The genetic diversity within KV was intermediate, at $1.2 \pm 0.3\%$. Sequence divergence between haplotypes varied from 0.25% to 6.1% (equivalent to one and 23 substitutions between haplotypes, respectively) as shown in Table 1. Per cent sequence divergence between populations was highest between MF and KV (1.5%) while KV and QE showed the least divergence (0.4%) despite the geographical distance separating them being greater than that between KV and MF. Sequence divergence between QE and MF was 1.4% (Table 2). A population cladogram showing the genetic relationship between the three populations is shown in Fig. 3.

The K_{ST} value for the total sample was 0.32, indicating significant genetic differentiation between the three populations ($P < 0.001$). Pairwise homogeneity tests for population subdivision based on K_{ST} and F_{ST} statistics showed that a significantly large proportion of the nucleotide sequence diversity (19–32% based on K_{ST} and 24–75% based on F_{ST}) was found between different populations rather than within populations, indicating significant subdivision between the populations (Table 2).

A neighbour-joining consensus tree showing evolutionary relationships between the haplotypes is shown in Fig. 4. A unique haplotype of an elephant from Ghana (WA3), genetically more distant from all Uganda haplotypes than any two of them compared, was used as an outgroup to root the tree. The tree generated using the maximum parsimony algorithm (not shown) also had the same topology. Two major clades, A and B, were recognized with bootstrap supports of 100% and 81%, respectively. The phylogenetic relationships of the haplotypes do not coincide with their geographical distribution; haplotypes from each of the populations are represented in each of the two clades. Nevertheless, the distribution of haplotypes in the two clades is not completely random. Clade A includes 20 out of the 21 individuals from KV, 39 out of 42 from QE and 3 out of 9 from MF. Within this clade, the predominantly KV/MF subclade is separated from the QE12 genotype by a weak bootstrap support of 60%. The few mutational events separating most of the haplotypes in this subclade (see Table 1) are responsible for the weak bootstrap support for this subclade plus the lack of resolution of the haplotypes within the subclade. Clade B contains mainly haplotypes from MF, apart from two haplotypes from QE and one from KV. All haplotypes in this clade are distinguished from the haplotypes in clade A by a 1-bp insertion at base pair position 39.

Microsatellite variation

Genotypic disequilibrium. No test for pairwise genotypic disequilibrium between the four loci was significant ($P > 0.05$) in any population (data not shown). Independent segregation of alleles at the different loci was therefore assumed in subsequent analyses.

Single-locus genotypic distribution. Observed genotype distributions were only significantly different from Hardy–Weinberg expectations for the Queen Elizabeth population at *Laf MS01* ($P = 0.01$) and *Laf MS03* ($P = 0.03$) and for the Kidepo valley population at *Laf MS03* ($P = 0.02$).

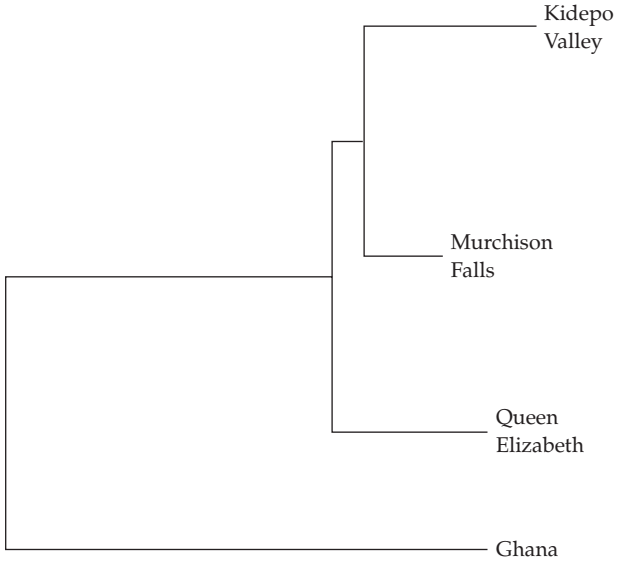

Fig. 3 Neighbour-joining population tree based on Nei's (1978) genetic distances.

The significant deviations were due to an excess of heterozygotes as indicated by negative F_{IS} values in Table 3. However, when deviations from Hardy–Weinberg proportions were considered across loci in each population, only QE deviated significantly ($P = 0.02$) after a Bonferroni correction (Rice 1989). The F_{IS} values were negative except in QE and KV at *Laf MS04* where the F_{IS} values were slightly positive (Table 3), suggesting that the extent of nonrandom mating was minimal.

Microsatellite allelic variation. All the four loci analysed were highly polymorphic: the most variable locus was *Laf MS02* with 10 alleles and the least variable was *Laf MS03* with six alleles. Considering all locus/population combinations, a total of 31 alleles was scored in the three populations. Total number of alleles per locus in each population ranged from 3 to 10 (Table 4). Observed heterozygosity values ranged from 0.46 to 1.00 while unbiased expected heterozygosities ranged from 0.51 to 0.84 (Table 3). The total number of alleles scored in each

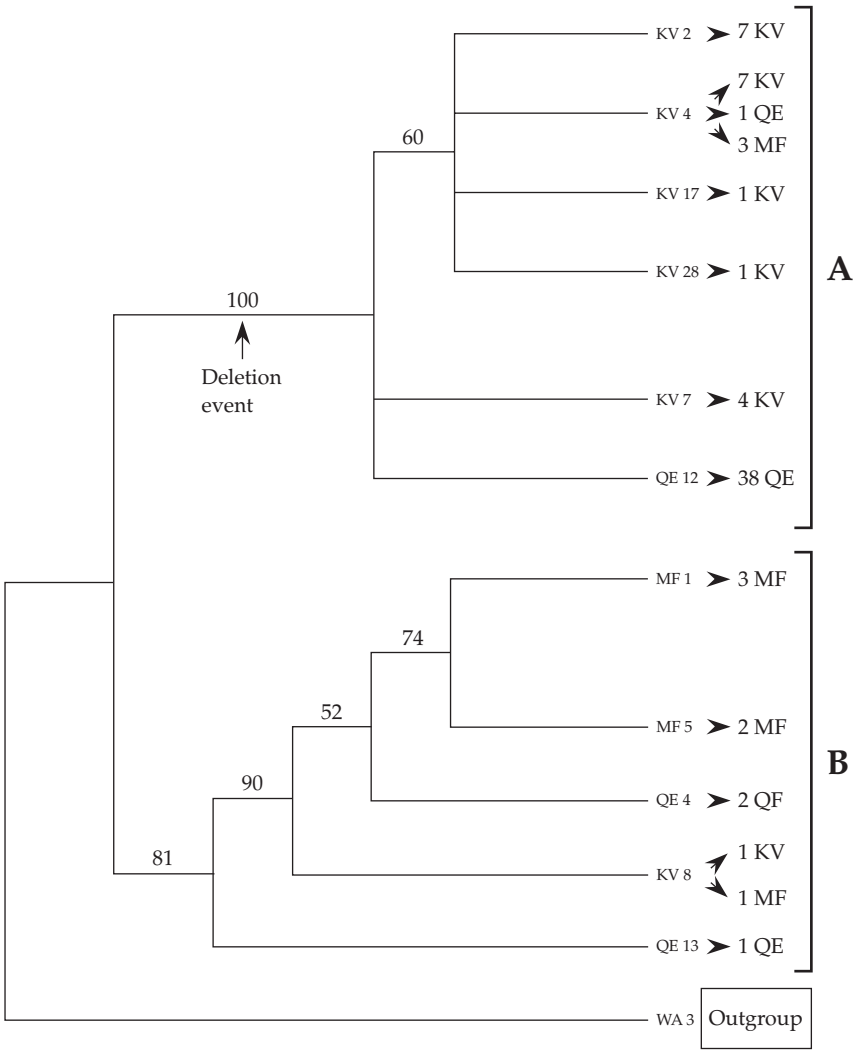

Fig. 4 Neighbour-joining consensus tree of the haplotypes based on the Hasegawa–Kishino–Yano (HKY 1985) model. The numbers above the branches represent percentage bootstrap values. The number of individuals from each population represented by each haplotype is also indicated.

Table 3 Observed heterozygosity (H_O), expected heterozygosity (H_E) inbreeding coefficients (F_{IS}) and exact probabilities for Hardy–Weinberg proportions at four microsatellite loci

	Laf MS01			Laf MS02			Laf MS03			Laf MS04		
	KV	MF	QE	KV	MF	QE	KV	MF	QE	KV	MF	QE
N	21	9	42	21	9	42	21	9	42	21	9	42
H_O	0.91	1.00	0.82	0.91	1.00	0.87	0.71	0.56	0.87	0.52	0.67	0.46
H_E	0.71	0.84	0.81	0.82	0.80	0.84	0.65	0.54	0.68	0.56	0.62	0.51
F_{IS}	−0.29	−0.20	−0.01	−0.11	−0.27	−0.04	−0.11	−0.03	−0.28	0.06	−0.08	0.09
$P_{(HW)}$	0.13	0.11	0.01	0.81	1.00	0.90	0.02	1.0	0.03	0.87	1.00	0.49

N, number of individuals.

Table 4 Allele frequencies at the microsatellite loci

	Laf MS01				Laf MS02				Laf MS03				Laf MS04		
	KV (N = 21)	MF (N = 9)	QE (N = 42)		KV (N = 21)	MF (N = 9)	QE (N = 42)		KV (N = 21)	MF (N = 9)	QE (N = 42)		KV (N = 21)	MF (N = 9)	QE (N = 42)
188	0.048	0.056	0.141	**134**			0.026	**140**	0.071	0.056	0.038	**143**			0.013
190	0.095	0.111	0.192	**138**	0.024		0.026	**142**	0.119		0.013	**149**			0.115
192	0.071	0.056	0.051	**140**	0.167		0.077	**144**	0.548	0.611	0.423	**151**	0.024	0.111	0.038
196	0.095	0.056	0.013	**142**		0.056	0.064	**146**	0.214	0.333	0.359	**153**	0.024		
198	0.190	0.278	0.128	**144**	0.095		0.026	**148**	0.048		0.115	**155**	0.357	0.389	0.154
200	0.500	0.278	0.333	**146**	0.143	0.111	0.051	**150**			0.051	**157**	0.571	0.500	0.679
202		0.167	0.115	**148**	0.333	0.222	0.269					**159**	0.024		
204			0.026	**150**	0.167	0.389	0.231								
				152	0.071	0.056	0.154								
				154		0.167	0.077								
A	6	7	8		7	6	10		5	3	6		5	3	5

N, number of individuals analysed; A, number of alleles; allele sizes in base pairs are represented in bold.

Table 5 UR_{ST} and F_{ST} statistics for the microsatellite loci

Comparison	UR_{ST}	F_{ST}
KV and MF	0.09*	0.01 ns
KV and QE	0.05*	0.02*
MF and QE	0.03 ns	0.01 ns

* Significant at 5% level; ns, not significant at 5% level.

population was highest in Queen Elizabeth national park (29) and lowest in Murchison Falls national park (19). Fewer alleles were recorded in MF at all loci except *Laf MS01* compared to KV and QE, most probably due to the smaller number of individuals examined in this population. The most frequent allele was 200 bp for *Laf MS01* in all populations while for *Laf MS02* allele 148 was the most frequent in QE and KV and allele 150 was the most common in MF. Alleles 144 and 157 were the most frequent at *Laf MS03* and *Laf MS04*, respectively. The frequencies of the most common alleles for each of the four loci ranged from 0.27 to 0.68 (Table 4).

Private alleles. Private alleles, whose occurrence is restricted to only one population, constituted 22.6% of all alleles. Five private alleles were observed in the Queen Elizabeth population while two alleles were observed in the Kidepo sample at locus *Laf MS04* (Table 4). None were observed in the Murchison population. All private alleles occurred with low frequencies ($P \leq 0.115$).

Population structure. Three out of the four loci (*Laf MS01*, *Laf MS03* and *Laf MS04*) had significant differences in allele frequencies between KV and QE (Fisher's exact probabilities, $P = 0.01$, 0.02 and 0.004, respectively) while only one locus (*Laf MS02*) differentiated KV from MF ($P = 0.02$). None of the loci differentiated MF from QE.

The proportion of the genetic variation attributable to genetic differences between the populations quantified with the F_{ST} statistic (Weir & Cockerham 1984) and unbiased estimates of the R_{ST} statistic (Slatkin 1995a) is shown in Table 5. UR_{ST} values were consistently higher than F_{ST} values in all pairwise comparisons. The unbiased R_{ST} values showed significant genetic differentiation at the 5% level between KV/MF and KV/QE but not between QE and MF. On the other hand, F_{ST} values were only significant

at the 5% level for the KV/QE comparison. In the QE population, sampling was done in two sectors, Mweya and Ishasha, separated by the Maramagambo forest. We suspected that this forest could have a significant influence on the movements of elephants between these two sectors resulting in genetic substructuring between these two areas. However, tests for genetic differentiation between the two areas were not significant ($K_{ST} = 0.001$, $P = 0.32$; $UR_{ST} = -0.0016$, $P = 0.41$).

Discussion

This study presents the first combined investigation of variation in mtDNA control region sequences and nuclear microsatellite loci in the African elephant. The results showed both concordance and discordance between the two marker sets. Lack of congruence between mitochondrial and nuclear-based variation has also been previously reported, e.g. in the brook charr (Ferguson *et al*. 1991), the *Ensatina eschscholtzii* complex (Moritz *et al*. 1992) and the trout (Bernatchez & Osinov 1995), where it has been explained by the differences in the modes of inheritance plus the different rates and modes of mutation in the two genetic marker systems.

While the mtDNA data showed significant genetic subdivision between all populations, suggesting limited female dispersal between populations, results from microsatellite analysis showed weak subdivision between KV and MF (only one locus out of four detected significant deviations in allele frequency distributions) and no subdivision between QE and MF. This discrepancy between mtDNA and microsatellite data suggests male-biased dispersal between the populations. Male-biased gene flow would lead to homogenization of nuclear alleles while the maternally inherited mitochondrial markers would remain restricted to individual localities. Our results are consistent with such a pattern of genetic exchange. The organization of the social structure of elephants enables matriarchal lineages to be maintained within stable family groups with home ranges which can be as small as 15–52 km^2 (Douglas-Hamilton 1973) leading to the observed mtDNA-based structuring of the elephant populations.

We also found incongruent patterns of genetic variation within and between the three populations. In contrast to the geographical distances separating them, the mtDNA data indicated greater genetic similarity between QE and KV compared to the KV/MF and QE/MF comparisons where the geographical separation between the populations is smaller. However, based on microsatellite data, the extent of genetic differentiation between populations was related to the geographical distances between them; KV and QE were the most distantly related populations and MF and QE could not be differentiated from each other using the four loci examined in this study (see Fig. 1). This relationship is depicted in the population tree based on Nei's genetic distance (Nei 1978) (Fig. 3).

The lack of complete phylogeographic partitioning of the mitochondrial haplotypes depicted in Fig. 4 suggests limited gene flow between the three populations. Haplotypes QE4 and QE13 (both females), which are more evolutionarily related to the MF population than the QE population (see Table 2) are an indirect evidence of such gene flow between the two populations. These haplotypes probably represent historical migrants from MF into the QE population. Similarly, KV8 is a possible migrant from MF into the KV population. Alternatively, this may mean that the three populations were founded from the same ancestral population but sufficient time has not elapsed for the haplotypes in each currently isolated population to attain reciprocal monophyly, a process that takes place in a period of approximately four N_e generations following isolation, where N_e is the effective population size (Avise *et al*. 1984; Nigel & Avise 1986).

High levels of heterozygosity in this species (as evidenced by negative F_{IS} values in Table 3) can be explained by the manner in which family groups are organized; family units are matrilineal, consisting of sisters, their female offspring and juvenile males which get expelled from the groups as soon as they attain puberty.

With such a social organization, most of the matings involve females and unrelated males, thereby resulting in levels of heterozygosity in excess of Hardy–Weinberg expectations. Negative F_{IS} values have also been reported in several monkey species with a similar social structure (e.g. Aoki & Nozawa 1984; Pope 1992).

The apparently contradictory levels of genetic variation shown by the two genomes in QE suggest a possible recent bottleneck. Because the effective population size of mtDNA is only 25% that of nuclear DNA, a recent bottleneck could easily have reduced mitochondrial variability without appreciably affecting nuclear variability. This observation should serve as a reminder that estimates of population structure based on one genetic system must be interpreted cautiously. The higher sensitivity of the mitochondrial genome to genetic drift relative to the nuclear genome was further evidenced by the higher values of K_{ST} compared to those of UR_{ST} in all the pairwise estimates of genetic subdivision (Table 2 and Table 5). Because of the smaller effective population size of the mitochondrial genome, the K_{ST} values are theoretically expected to be at least four times the UR_{ST} values in the absence of gender-biased gene flow and higher if it does exist.

Our UR_{ST} estimates indicated stronger subdivision between populations than the F_{ST} estimates (see Table 5). F_{ST} estimates of population subdivision are more realistic when dealing with alleles whose evolution is compatible

Table 6 Comparison of diversity indices of the elephant with those of other large mammals studied in the African region based on the nucleotide sequence of the 5′ end of the D-loop

	Base pairs analysed	% Variable sites	% Nucleotide diversity (K_T)	K_{ST}	Source/Reference
Elephant	400	7.8	1.4	0.32	This study
Buffalo	348	36	5.0	0.10	Simonsen *et al*. (1998)
Waterbuck	402	25	4.1	0.17	Simonsen (1997)
Impala	427	13	2.8	0.17	Arctander *et al*. (1996b)
Grant's gazelle	371	34	10.9	0.66	Arctander *et al*. (1996a)

with the infinite allele mutation model (Kimura & Crow 1964). This model assumes that mutations always lead to the production of new alleles. On the contrary, mutation patterns at most microsatellite loci have been shown to deviate from the infinite allele model (e.g. Weber & Wong 1993). Instead, most mutations at microsatellite loci produce allelic states that are phenotypically indistinguishable from those that already exist. Therefore an F_{ST}-based estimate of population subdivision using microsatellite data will greatly underestimate the extent of genetic differentiation between populations. UR_{ST} estimates of population subdivision, on the other hand, are more realistic because they are based on the stepwise mutation model (Ohta & Kimura 1973) which takes into account the mutational properties of microsatellite loci. Given the extent of population isolation and fragmentation caused by the extensive human habitation of the former elephant range during the last 100 years or so, current gene flow between the three populations may have even dropped to zero, thereby making the contemporary genetic structure even more subdivided than what is revealed by our present analysis.

A brief comparison of our results with those obtained in similar studies on other large mammals in Eastern and Southern Africa (Table 6) indicates that the Uganda elephant population has a relatively much lower mtDNA diversity.

Despite this, the elephant (apart from the Grant's gazelle) shows the strongest degree of subdivision amongst its populations. Georgiadis *et al*. (1994), using the sample-reuse algorithm of Davis *et al*. (1991), found significant subdivision between elephants from eastern and southern Africa but not between populations within each of the regions. Re-analysis of the same data by Siegismund & Arctander (1995) using the procedure of Hudson *et al*. (1992), which has a higher resolution power than the test of Davis *et al*. (1991), revealed significant subdivision between populations both at the regional and continental levels. The level of divergence we have detected between elephant populations in Uganda as quantified by the F_{ST} statistic, an analogue of the H_{ST} statistic of Hudson *et al*. (1992) ($F_{ST} = 0.24 - 0.74$, Table 2) is greater than what was detected both within regions ($F_{ST} = 0.23 - 0.27$) and between regions ($F_{ST} = 0.39$) by Georgiadis *et al*. (1994) and also by Siegismund & Arctander (1995) ($H_{ST} = 0.16 - 0.3$). A basic difference between our data set and that of Georgiadis *et al*. (1994) is that the latter was based on presence or absence of restriction sites while our analyses were based on specific nucleotide differences between haplotypes. Analysis of nucleotide sequence differences between haplotypes gives a higher resolution power capable of detecting fine-scale population subdivision where restriction site data would otherwise indicate homogeneity between populations.

Conclusions

Our results have clearly revealed significant genetic subdivision between the three elephant populations in Uganda, with stronger subdivision of the mitochondrial genome compared to the nuclear genome. Contrasting population genetic structure displayed by the two genomes in some of the population comparisons has been explained in terms of male-biased gene flow between those populations. In addition to strong differentiation among the populations, our results have revealed relatively low levels of mtDNA diversity compared to other large mammals in the Eastern and Southern African region.

Acknowledgements

This work is part of the PhD research of Silvester Nyakaana on the population genetic structure of the African elephant. Laboratory work was performed in the DNA laboratory in the Department of Population Biology, Copenhagen University. This research was funded by the Danish International Development Agency (DANIDA) under the Wildlife Genetics Project, a collaborative project between the University of Copenhagen (Denmark) and Makerere University (Uganda). We are greatly indebted to the Management of Uganda Wildlife authority, the chief wardens of Queen Elizabeth, Murchison Falls and Kidepo Valley national parks for granting us permission to conduct research in the national parks. We are also very grateful to Dr Eve Lawino Abe for assistance in the collection of samples from Queen Elizabeth national park. Our thanks also go to H. R. Siegismund, B. T. Simonsen and M. Bruford for comments that improved the quality of the manuscript.

References

Amos W, Hoelzel AR (1991) Long term preservation of whale skin for DNA analysis. In: *Genetic Ecology of Whales and Dolphins*. (ed. Hoelzel AR), pp. 99–104. Report of the International Whaling Commission (Special Issue, 13).

Aoki K, Nozawa K (1984) Average coefficient of relationship within troops of the Japanese monkey and other primate species with reference to the possibility of group selection. *Primates*, **25**, 171–184.

Arctander P, Kat PW, Aman RA, Siegismund HR (1996a) Extreme genetic differences among populations of Grant's gazelle (*Gazella granti*) in Kenya. *Heredity*, **76**, 465–475.

Arctander P, Kat PW, Simonsen BT, Siegismund HR (1996b) Population genetics of Kenyan Impalas: Consequences for conservation. In: *Molecular Genetic Approaches in Conservation* (eds Smith BT, Wayne RK), pp. 399–412. Oxford University Press, New York.

Avise JC, Nigel JE, Arnold J (1984) Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *Journal of Molecular Evolution*, **20**, 99–105.

Bernatchez L, Osinov A (1995) Genetic diversity of trout (Genus *Salmo*) from its most eastern nature range based upon mitochondrial DNA and nuclear gene variation. *Molecular Ecology*, **4**, 285–297.

Brooks AC, Buss IO (1962) Past and present status of the elephant in Uganda. *Journal of Wildlife Management*, **26** (1), 38–50.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Current Opinions in Genetics and Development*, **3**, 939–943.

Clegg SM, Hale P, Moritz C (1998) Molecular population genetics of the red Kangaroo (*Macropus rufus*): Mitochondrial DNA variation. *Molecular Ecology*, **7**, 679–686.

Davis SK, Strassmann JE, Hughes C, Pletcher LS, Templeton AR (1991) Population structure and kinship in Polistes (Hymenoptera, Vespidae): an analysis using ribosomal DNA and protein electrophoresis. *Evolution*, **44**, 1242–1253.

Douglas-Hamilton I (1973) On the ecology and behaviour of the Lake Manyara elephants. *East African Wildlife Journal*, **11**, 401–403.

Eltringham SK (1977) The numbers and distribution of Elephant *Loxodonta africana* in the Rwenzori National Park and Chambura Game Reserve, Uganda. *East African Wildlife Journal*, **15**, 19–39.

Eltringham SK, Malpas RC (1980) The decline in elephant numbers in Rwenzori and Kabalega Falls National Parks, Uganda. *African Journal of Ecology*, **18**, 73–86.

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package, Version 3.5c). Distributed by the author. Department of Genetics, University of Washington, Seattle.

Ferguson MM, Danzmann RG, Hutchings JA (1991) Incongruent estimates of population differentiation among brook charr. *Salvelinus fontinalis* from Capen Race, Newfoundland, Canada based upon allozyme and mitochondrial DNA variation. *Journal of Fisheries Biology*, **39** (Suppl. A), 79–85.

Georgiadis N, Bischof L, Templeton A, Patton J, Karesh W, Western D (1994) Structure and history of African elephant populations: I Eastern and Southern Africa. *Heredity*, **85**, 100–104.

Gilbert D (1993) SeqApp 1.9. Biocomputing office, Biology dept., Indiana University of Bloomington.

Goodman SJ (1997) RSTCALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and gene flow from microsatellite data and determining their significance. *Molecular Ecology*, **6**, 881–886.

Guo SW, Thompson EA (1992) Performing the exact tests of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Gyllensten UB, Erlich HA (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of *the HLA-DQA* locus. *Proceedings of the National Academy of Sciences of the USA*, **85**, 7652–7656.

Hamilton D (1973) On the ecology and behaviour of the Lake Manyara elephants. *East African Wildlife Journal*, **11**, 401–403.

Hudson RR, Boos DD, Kaplan N (1992) A statistical test for detecting geographical subdivision. *Molecular Biology and Evolution*, **9**, 138–151.

Karesh W, Smith F, Frazier-Taylor H (1987) A remote method for obtaining skin biopsy samples. *Conservation Biology*, **1**, 261–262.

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications for nuclear RFLPs. *Science*, **256**, 100–102.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725–738.

Laws RM, Parker ISC, Johnstone RCB (1970) Elephants and their habitats in north Bunyoro, Uganda. *East African Wildlife Journal*, **8**, 163–180.

Leuthold W, Sale JB (1973) Movements and patterns of habitat utilisation of elephants in Tsavo National Park, Kenya. *East African Wildlife Journal*, **11**, 369–384.

Levene H (1949) On a matching problem arising in genetics. *Annals of Mathematics and Statistics*, **20**, 91–94.

Louis EJ, Dempster ER (1987) An exact test for Hardy–Weinberg and multiple alleles. *Biometrics*, **43**, 805–811.

Moritz C, Schneider CJ, Wake DB (1992) Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology*, **41** (273), 291.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nigel J, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: *Evolutionary Processes and Theory* (eds Nevo E, Karlin S), pp. 515–534. Academic Press, New York.

Nyakaana S, Arctander P (1998) Isolation and characterisation of microsatellite loci in the African elephant (*Loxodonta africana* Blumenbach 1797). *Molecular Ecology*, **7**, 1436–1437.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetic Research*, **22**, 201–204.

Olivier RCD (1991b) Elephant Conservation plan for Uganda. Unpublished report to the Uganda National Parks.

Olivier RCD, Edroma EL, Campbell KLI (1989) Aerial monitoring of large mammal populations in the Queen Elizabeth national park, Uganda. Unpublished report.

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, **3**, 489–495.

Palsbøll PJ, Heide-Jørgensen MP, Dietz R (1997) Population structure and seasonal movements of narwhals, *Monodon monoceros*, determined from mtDNA analysis. *Heredity*, **78**, 284–292.

Parker ISC, Douglas-Hamilton I (1976) The decline of elephants in Uganda's Rwenzori and Kabalega Falls national parks.

Unpublished report to the minister of tourism and wildlife, Uganda.

Pope TR (1992) The influence of dispersal patterns and mating system on genetic differentiation within and between populations of the red Howler monkey (*Alouatta seniculis*). *Evolution*, **46**, 1112–1128.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49** (6), 1280–1283.

Raymond M, Rousset F (1997) GENEPOP, Version 3.1b: an updated version of GENEPOP (Version 1.2), Population genetics software for exact tests and ecumenism. *Journal of Heredity*, **86**, 248–249.

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

Saccone C, Pesole G, Sbisa E (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *Journal of Molecular Evolution*, **33**, 83–91.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Sanger F, Nicklen S, Coulsen AR (1977) DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences of the USA*, **74** (5463), 5467.

Schneider S, Jean-Marc K, Roessli D, Excoffier L (1997) *ARLEQUIN*, Version 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva.

Shriver MD, Jin L, Chakraborty R, Boerwinkle R (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. *Genetics*, **134**, 983–993.

Siegismund HR, Arctander P (1995) Structure of African elephant populations. *Journal of Heredity*, **86**, 467–469.

Simonsen BT (1997) Population structure and history of African bovids. PhD Thesis. Department of Population Biology, University of Copenhagen, Denmark.

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

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Stanley HF, Casey S, Carnahan JM, Goodman S, Harwood J, Wayne RK (1996) Worldwide patterns of mitochondrial DNA differentiation in the Harbor seal (*Phoca vitulina*). *Molecular Biology and Evolution*, **13**, 368–382.

Swofford DL (1998) *PAUP** (Phylogenetic Analysis Using Parsimony), Version 4.0d64 (*PPC, test*). Washington DC, Smithsonian Institution.

Weber JL, Wong C (1993) Mutation of human short tandem repeats. *Human Molecular Genetics*, **2**, 1123–1128.

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

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

Yeh FC, Yang RC, Boyle T (1997) *POPGENE*, Version 1.21: A microsoft window-based freeware for population genetic analysis. Department of Renewable Resources, University of Alberta, Edmonton, Canada.

This research is part of a broader research programme on the genetics and evolution of populations of large African mammals —The Wildlife Genetics Project. Silvester Nyakaana is working on population and family genetic structure of the African elephant and Peter Arctander's prime research interests are speciation patterns and population genetics of birds and mammals and evolution of noncoding DNA.
