African Zoology 2017, 52(3): 155–162
Printed in South Africa — All rights reserved

AFRICAN ZOOLOGY
ISSN 1562-7020 EISSN 2224-073X
http://dx.doi.org/10.1080/15627020.2017.1379358

Genetic structure of wild and domesticated grasscutters (*Thryonomys swinderianus*) from south-western Nigeria

Oluwakayode M Coker[1], Abosede O Omonona[1], Olusegun A Fagbohun[2], Cortney Pylant[3] and James D Austin[3]*

[1] *Department of Wildlife and Ecotourism Management, University of Ibadan, Ibadan, Nigeria*
[2] *Department of Veterinary Microbiology and Parasitology, University of Ibadan, Ibadan, Nigeria*
[3] *Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida, USA*
* *Corresponding author, email: austinj@ufl.edu*

Grasscutter (*Thryonomys swinderianus*) is a large rodent distributed across sub-Saharan Africa that is highly valued as a source of protein. There has been little effort to evaluate the genetic structure of grasscutters despite long-term harvesting pressure and over 40 years of grasscutter domestication in West Africa. Our objectives were to quantify the genetic structure of wild grasscutters, and to compare genetic variation from wild samples to those from various farmed samples within south-western Nigeria. We genotyped 145 wild and 88 domesticated individuals at 11 microsatellite loci and present results quantifying regional genetic structuring and the relative patterns of diversity among wild and domesticated grasscutter populations. Our data reflect high differentiation between wild and domesticated grasscutters, and significantly greater allelic richness and gene diversity in the former. Despite this, domesticated populations appear to have similar levels of observed heterozygosity and comparable levels of differentiation among domesticated samples relative to wild samples. This may be the result of high turnover within captive colonies, or frequent infusion of new animals. More detailed molecular and quantitative genetic studies are recommended on this species to be able to understand their natural variation, degree of connectivity and to improve strategies for domestication.

Keywords: domestication, farming, gene flow, genetic drift, genetic structure, microsatellites, *Thryonomys swinderianus*

Online supplementary material: Supplementary information for this article is available at http://dx.doi.org/10.1080/15627020.2017.1379358

Introduction

Quantification of genetic variation can provide valuable insight into genetic structure, evolutionary history and, in cases where species are heavily managed or manipulated, the potential impact of captive rearing of wildlife. A case in point is the grasscutter (*Thryonomys swinderianus*), a species that is highly valued as a protein source in West and Central Africa (Asibey 1974; Asibey and Addo 2000). The grasscutter is endemic to the sub-Saharan humid regions and are found in many forests and savannas. It is estimated that 80 million grasscutters weighing up to 300 t are being harvested per year (Jori et al. 1995). It therefore contributes greatly to both local and export earnings of most West African countries (Baptist and Mensah 1986; NRC 1991). The popularity of grasscutter meat in the region has led to increased heavy pressure on wild populations and local habitats due to destructive sampling, including the setting of bush fires (NRC 1991; Adu et al. 1999). As a result of the high demand and concern of over-harvesting of wild stocks (Hanotte and Mensah 2002), there has been increasing efforts to domestically breed grasscutters for local consumption and export (Adu 2002). Despite long-term harvesting pressures and over 40 years of grasscutter farming in Africa there has been little effort to evaluate the genetic structure of wild populations, nor the potential genetic consequences of farming (Adenyo et al. 2013), despite the effort to develop species-specific markers (Adenyo et al. 2012, 2017).

Grasscutters are common along riverine and wetland habitats of the Guinean Forest–Savanna mosaic that extends into south-western Nigeria. Representative ecozones include Guinea Savanna (GS), Derived Savanna (DS), Rain Forest (RF) and Swamp and Mangrove Forest (SF) zones. These ecozones differ markedly by vegetation types, rainfall and humidity, although no physical biogeographic barriers are present along the north–south gradient in south-western Nigeria that would affect connectivity of wildlife. The objective of this study was to use nuclear microsatellite genotypes to determine the level of genetic structuring across the four main ecozones in south-western Nigeria. In addition, we evaluate the neutral genetic impact, if any, of existing grasscutter domestication by screening grasscutters from seven domesticated populations. Microsatellites are high-resolution genetic markers and ideal for the study of gene flow and connectivity. In addition, their high variability allows for the detection of genetic drift affecting small populations over short time periods. We ask: (1) What degree of genetic differentiation and structure is found in south-western Nigerian grasscutter populations? (2) Is there evidence of genetic structuring by

habitat (ecozone)? (3) Is genetic variation in domesticated populations similar to that from wild grasscutter samples? These questions will provide much-needed baseline information on an important Central and West African natural resource, and will inform ongoing development of grasscutter domestication.

Methods

Blood, tissue or hair samples were collected from 88 domesticated and 145 wild grasscutters across south-western Nigeria (Figure 1) between January 2014 and February 2015. Tissue samples were preserved in 70% ethanol, blood samples were air-dried on Whatman filter papers, and hair samples were air-dried and stored in small envelopes. Wild grasscutters were collected from Abugaga (AB; *n* = 13), Tede (TD; *n* = 13), Ado (AD; *n* = 6), Iganna (IG; *n* = 1), Ikere (IK; *n* = 3), Iseyin (IS; *n* = 15), Ode (OD; *n* = 1), Usi (US; *n* = 1), Asejire (AS; *n* = 19), Kila (KI; *n* = 9), Omi Adio (OA; *n* = 6), Olomore (OL; *n* = 10), Odo Ona (OO; *n* = 6), Owode (OW; *n* = 7), Epe (EP; *n* = 22), and Mowo (MW; *n* = 13). These correspond to Guinea Savanna (GS; *n* = 26), Derived Savanna (DS; *n* = 27), Rain Forest (RF; *n* = 57), and Swamp/Mangrove Forest (SF; *n* = 35)

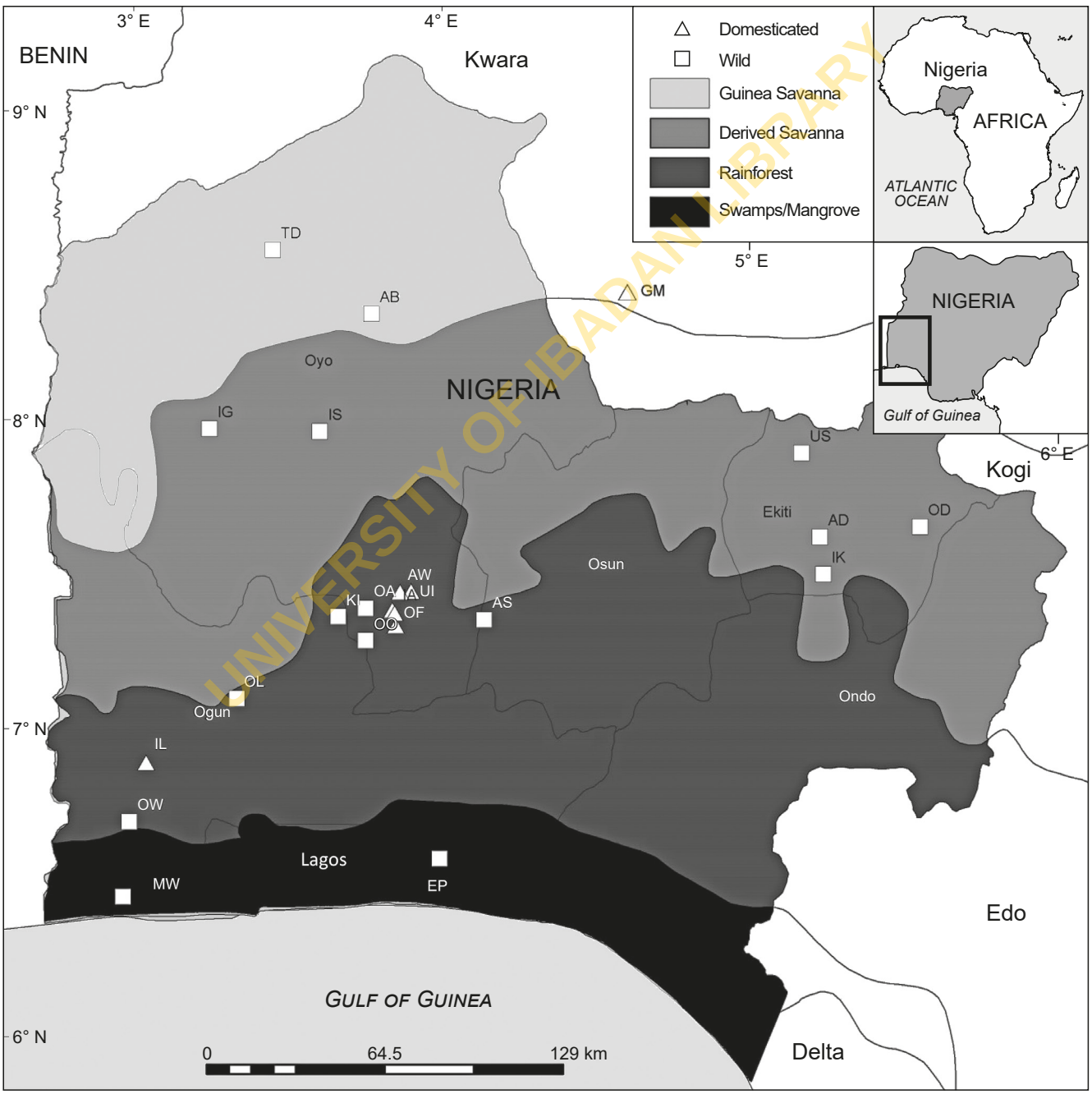

Figure 1: Map of the study area in south-western Nigeria

ecozones (Table 1). Samples from domesticated individuals were collected from the Federal College of Health and Animal Production Technology Farm (FC; n = 22), the Institute of Agricultural Research and Training (IA; n = 19), Domestication Unit, Department of Wildlife and Ecotourism, University of Ibadan (UI; n = 15), Odo Ona Farm (OF; n = 16), Zikbol Farms Awotan (AW; n = 6), Ilaro Farm (IL; n = 5) and Ganmo Farm (GM; n = 5) (Table 1).

Genomic DNA was extracted using the DNeasy® Blood and Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Extracted DNA was quantified using a Nanodrop Spectrophotometer ND1000 V3.7.1. Samples were genotyped at 11 grasscutter microsatellite loci (Adenyo et al. 2012), including seven dinucleotide (*Tsw 05*, *Tsw 06*, *Tsw 07*, *Tsw 08*, *Tsw 11*, *Tsw 12* and *Tsw 18*), three trinucleotide (*Tsw 31*, *Tsw 32* and *Tsw 33*), and one tetranucleotide (*Tsw 13*) markers. PCR amplifications were carried out in 15 µl reactions containing 5.66 µl H_2O, 7.5 µl Qiagen 2× Master Mix, 0.3 µl of 1.0 µM M13-tailed forward primer, 0.27 µl of 10 µM reverse primer, 0.27 µl of 10 µM M13 dye-labelled primer (hexachlorofluorescein and/or 5′-fluorescein [6-FAM]) and 1 µl of 20 ng/µl DNA sample. The thermocycling reactions included initial activation at 95 °C for 5 min, followed by 29–38 cycles of denaturation at 95 °C for 30 s, and annealing at 60 °C/55 °C for 90 s, and an extension at 72 °C for 60 s, followed by a final extension of 60 °C for 30 min. Number of cycles was 29 for tissue samples, 35 for hair samples and 38 for dried blood samples. The PCR products were electrophoresed on an ABI 3130xl using a ROX 400 size standard. Alleles were scored using GENEMARKER® 2.6 software and allele calls were manually confirmed. We used exact tests for departures from Hardy–Weinberg proportions (HWP) using Genepop 4.2 (Raymond and Rousset 1995). We limited our tests for departures from HWP, linkage disequilibrium (LD) and null alleles (below) to the six largest samples (≥12 genotypes) due to the known effect of sampling on measures of allele frequency, and to keep the number of pairwise tests to a reasonable number. For exact tests, we used 5 000 demarcations, 1 000 batches and 5 000 iterations per batch for the Markov chain mixing to ensure that standard errors were below 0.01, and increasing the accuracy of estimated p-values (Raymond and Rousset 1995). We examined patterns of significance at an adjusted $\alpha \leq 0.05$ applying the B-H method (Narum 2006). We took a similar approach to evaluating LD between loci. We used probability tests, applying 5 000 demarcations, 1 000 batches and 5 000 iterations per batch as above and examined significance using B-H tests. We used Microchecker to detect presence of null alleles, large allele dropout and scoring errors due to stuttering (van Oosterhout et al. 2004) for wild population samples having greater than nine genotypes, and again with all samples combined. We also confirmed genotyping consistency by repeating PCRs on all loci for ~18% of the samples. Genotypes were scored independently by two authors (OMC and JDA).

Average number of alleles (A), average number of effective alleles (n_e), allelic richness (set to the minimum samples size of 5; A_{r5}), observed (H_o) and expected (H_e), and unbiased gene diversity (uH_e) were calculated in FSTAT 2.9.3.2 (Goudet 2002). We evaluated the statistical power to detect genetic heterogeneity for our set of wild samples, using simulations based on the number of loci and

Table 1: Summary statistics for wild and domesticated grasscutters (*T. swinderianus*) genotypes for locations containing ≥5 genotypes. Multi-locus means and standard errors are reported. Approximate year of captive colony initiation is indicated for domesticated populations. n = sample size, A = average number of alleles, n_e = number of effective alleles, H_o = observed heterozygosity, H_e = expected heterozygosity, uH_e = unbiased expected heterozygosity, A_{r5} = allelic richness (standardized to N = 5)

Site	A	A_{r5}	n_e	H_o	H_e	uH_e
Guinea Savanna						
AB, n = 13	7.545 (0.867)	5.11 (0.404)	4.921 (0.709)	0.710 (0.052)	0.754 (0.031)	0.784 (0.033)
TD, n = 13	8.273 (0.964)	5.42 (0.403)	5.449 (0.585)	0.713 (0.055)	0.793 (0.023)	0.825 (0.023)
Derived Savanna						
AD, n = 6	5.182 (0.658)	4.86 (0.565)	4.272 (0.658)	0.655 (0.066)	0.699 (0.048)	0.766 (0.052)
IS, n = 15	8.364 (1.193)	5.30 (0.487)	5.821 (0.951)	0.738 (0.030)	0.787 (0.027)	0.814 (0.028)
Rain Forest						
AS, n = 19	8.909 (0.909)	5.12 (0.359)	5.044 (0.642)	0.742 (0.061)	0.768 (0.032)	0.789 (0.033)
KI, n = 9	5.909 (0.610)	4.80 (0.400)	4.252 (0.457)	0.602 (0.072)	0.719 (0.047)	0.762 (0.050)
RO, n = 6	5.727 (0.557)	5.26 (0.485)	4.493 (0.592)	0.652 (0.079)	0.726 (0.043)	0.792 (0.047)
OL, n = 10	6.909 (0.732)	5.23 (0.432)	5.069 (0.668)	0.709 (0.065)	0.765 (0.032)	0.806 (0.034)
OO, n = 6	5.818 (0.464)	5.37 (0.392)	4.593 (0.415)	0.773 (0.065)	0.760 (0.026)	0.829 (0.028)
OW, n = 7	5.455 (0.578)	4.80 (0.448)	4.112 (0.597)	0.626 (0.073)	0.699 (0.048)	0.756 (0.052)
Swamp/Mangrove Forest						
EP, n = 22	7.091 (0.995)	4.50 (0.371)	4.262 (0.592)	0.585 (0.040)	0.727 (0.031)	0.745 (0.031)
MW, n = 13	6.545 (0.666)	4.68 (0.363)	4.251 (0.492)	0.682 (0.053)	0.729 (0.033)	0.758 (0.034)
FC, n = 22 (2004)	6.182 (0.615)	4.15 (0.282)	3.691 (0.366)	0.665 (0.055)	0.690 (0.043)	0.707 (0.044)
IA, n = 19 (1997)	7.273 (0.854)	4.76 (0.415)	4.637 (0.682)	0.675 (0.044)	0.733 (0.039)	0.752 (0.041)
OF, n = 16 (2011)	6.000 (0.632)	4.34 (0.384)	4.056 (0.469)	0.705 (0.065)	0.672 (0.061)	0.719 (0.062)
UI, n = 15 (1998)	5.818 (0.519)	4.23 (0.342)	3.714 (0.438)	0.706 (0.065)	0.672 (0.055)	0.696 (0.057)
AW, n = 6 (2008)	4.091 (0.343)	3.95 (0.295)	3.399 (0.227)	0.712 (0.064)	0.692 (0.021)	0.755 (0.023)
GM, n = 5 (2007)	3.727 (0.304)	3.73 (0.304)	2.752 (0.221)	0.618 (0.063)	0.607 (0.038)	0.675 (0.042)
IL, n = 5 (1998)	4.273 (0.384)	4.27 (0.384)	3.457 (0.321)	0.691 (0.083)	0.684 (0.030)	0.760 (0.033)

observed allele frequencies using POWSIM 4.0 (Ryman and Palm 2006). POWSIM simulates sampling of genes from a specified number of populations that have diverged because of random drift (assumes no gene flow) into an expected predefined level of differentiation (measured as F_{ST}). Samples were drawn from simulations in numbers matching our sample sizes (12 populations of 6+ samples) and allele frequencies to test for genetic homogeneity using Fisher's exact test. Estimates of power were obtained as the proportion of significant outcomes when repeating the simulations 1 000 times for each level of F_{ST} using default iteration parameters. We evaluated divergence times of 25 and 50 generations and variaous effective populations sizes (ranging from 1 000 to 5 000). We also did an initial null simulation setting $t = 0$, to test that the initial number of false significances (α) was close to 0.05.

We calculated fixation indices (Θ_{ST}, an analogue of F_{ST}, Weir and Cockerham 1984) and Nei's genetic distance (D) in GENALEX 6.5 (Peakall and Smouse 2006, 2012) for both wild and domesticated populations. We tested for significantly greater diversity (A_{r5}, H_o, H_e, and Θ_{ST}) in wild vs domesticated samples using 1 000 permutation replicates in FSTAT (one-sided, P-values wild > domestic). We used Arlequin 3.5 to conduct an analysis of molecular variance (AMOVA) to describe variation partitioned among and within the wild populations, and also to quantify variation among wild and domesticated populations. We also ran an AMOVA nesting locations within ecoregions.

We examined population genetic structure among wild and domesticated grasscutters using principal coordinates analysis (PCoA), which is free from assumptions associated with HWE and linkage disequilibrium. PCoA was conducted using the covariance-standardised method implemented in GENALEX 6.5. We also applied a Bayesian genetic clustering analysis (Pritchard et al. 2000) to evaluate the degree and geographic pattern of population structuring. We examined posterior support for varying numbers of genetic clusters (K) using the program STRUCTURE 2.3.4 (Pritchard et al. 2000; Evanno et al. 2005). We examined LnP(K) in the range of 1–5 for wild populations and 1–15 for the combined wild and domesticated populations. We performed each K value exploration 15 times. Markov chain Monte Carlo consisted of 5.0×10^4 burn-in iterations followed by 2.0×10^5 iterations. Results from STRUCTURE runs were post-processed using STRUCTURE HARVESTER 0.6.94 (Earl and vonHoldt 2012). The Evanno Delta K transformation method (Evanno et al. 2005) was used to aid in assessing the best K. CLUMPP 1.1.2b (Jakobsson and Rosenberg 2007) and Distruct 1.1 (Rosenberg 2004) were used for cluster matching and graphical display of structure harvester output. The relationship between genetic and geographic distance was investigated through Mantel tests for the wild populations ($n = 12$) using 999 permutations to test for significance in Genalex.

Results

Of the 66 global tests for departures from HWP, the B-H p-value adjusted for 66 tests was $\alpha = 0.0061$, resulted in eight significant departures from HWP, including five population tests of *Tsw 33* (all but AB; $p = 0.0449$) and one test each for *Tsw 7* (AS), *Tsw 8* (TD) and *Tsw 11* (AS). MICROCHECKER results revealed evidence of scoring error due to stuttering at *Tsw 33*, and potential null alleles at three loci (*Tsw 13, Tsw 31* and *Tsw 33*). Due to the small sample sizes in some locations and likelihood of large historical effective population sizes in rodents, most of our departures from HWP can be dismissed as sampling variance. However, we omitted *Tsw33* from some analyses (pairwise differentiation and STRUCTURE) where departures from HWP may influence results.

Summaries of diversity indices are presented in Table 1 (see also Supplementary Table S1), and reflect a general trend toward higher values in wild vs domesticated samples. Mean population A_{r5} was significantly higher in wild relative to domesticated samples (wild mean = 5.04, domesticated mean = 4.20, $P = 0.002$). Among wild populations the average observed heterozygosity (H_o) was 0.682 for wild and 0.683 for domesticated ($P = 0.5$) and the average expected heterozygosity (H_e) was 0.789 and 0.723 for wild and domesticated, respectively ($P = 0.002$).

Significant genetic variance was limited among vegetation zone (1.1%) and among populations (1.6%), with the majority of variance being explained by among (12%) and within individuals (85%). For microsatellites, the within-individual and -population variance components usually comprise the majority of the total variance (Meirmans 2006). Even comparing the wild and domesticated populations similarly found most of the variation nested within individuals (84%) and 2.3% explained by the partition among wild and domesticated populations (Table 2).

Genetic differentiation among wild and among domesticated populations were similar ($F_{ST} = 0.075$ and 0.068, respectively), and both significantly greater than 0 ($p < 0.001$). Statistical power to detect differentiation due to drift alone showed that the sample sizes and microsatellite markers were adequate for detecting low levels of differentiation (e.g. $F_{ST} = 0.01$) with high power ($P > 0.8$) across most realistic t and n_e values (Supplementary Material S1). F_{ST} values across all simulated parameters were well below observed, suggesting that drift is insufficient in explaining observed levels of differentiation. Nei's genetic distances between wild locations ranged between 0.14 (between IS and AS) and 0.52 (between AD and MW) (Table 3), with an average genetic distance of 0.29 between each pair of population. Nei's genetic distance ranged between 0.13 (between IA and OO) and 0.58 (between UI and AW) with an average genetic distance of 0.35 between each pair of domesticated population. The mean pairwise Nei's D between domesticated and wild samples was 0.40 (Table 3).

Principal coordinates analysis reflected that the genetic distances between populations were primarily arranged by latitude (20.5% axis 1) among the wild populations (Figure 2a), and the first two axes described 56% of the distance. Examining the wild and domesticated samples together, the first two PCoA axes (48%) separated most of the wild populations from the domesticated populations (Figure 2b).

Results from Bayesian clustering analysis indicated $K = 3$ genetic clusters (mean LnP(3) = −6215.595, SD = 15.42) for all the wild populations and $K = 2$ genetic clusters (mean

Table 2: Result of AMOVA on wild samples (top) and wild vs domesticated samples (bottom)

Source of variation	df	Sum of squares	Variance components	Proportion of variation (%)	*P*-value
Wild only					
Among vegetation zones	3	29.74	0.047	1.09	0.0313
Among populations	12	69.87	0.070	1.60	<0.0001
Among individuals	129	613.73	0.527	12.12	<0.0001
Within individuals	145	537.0	3.703	85.19	<0.0001
Total	289	1 250.34	4.347		
Wild vs domesticated					
Among wild or domesticated	1	29.34	0.094	2.33	<0.0001
Among populations	21	150.56	0.156	3.88	<0.0001
Among individuals	210	870.37	0.378	9.41	<0.0001
Within individuals	233	789.5	3.388	84.37	<0.0001
Total	465	1 839.78	4.016		

Table 3: Pairwise matrix of Nei's genetic distance (*D*) (lower matrix) and Weir and Cockerham (1984) analog of F_{ST} (upper matrix) within and among (shaded) wild and domesticated populations. Significant F_{ST} ($\alpha = 0.05$) are in bold

Wild populations (WP)												Farm populations (FP)								
AB	TD	AD	IS	AS	KI	OA	OL	OO	OW	EP	MW	FC	IA	OF	UI	AW	GM	IL		
	0.03	0.05	0.03	**0.04**	**0.06**	0.04	**0.03**	0.04	**0.06**	**0.05**	**0.05**	**0.05**	**0.03**	**0.05**	**0.07**	**0.05**	**0.10**	0.04	AB	
0.20		0.04	0.02	**0.03**	**0.05**	0.04	0.02	0.03	0.04	**0.03**	**0.05**	**0.05**	**0.04**	**0.04**	**0.06**	**0.05**	**0.10**	**0.06**	TD	
0.32	0.27		0.04	0.05	**0.07**	0.05	0.06	0.05	0.05	**0.07**	**0.08**	**0.05**	0.04	0.04	**0.07**	0.07	**0.08**	0.07	AD	
0.19	0.16	0.23		0.02	0.04	0.04	0.03	0.03	0.04	**0.03**	**0.04**	**0.05**	**0.03**	**0.04**	**0.06**	**0.05**	**0.09**	0.04	IS	
0.27	0.23	0.28	0.14		0.03	0.03	0.02	0.02	0.04	**0.03**	**0.04**	**0.06**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.06**	AS	
0.46	0.38	0.38	0.26	0.16		0.06	0.04	0.05	0.05	**0.04**	**0.06**	**0.08**	**0.06**	**0.08**	**0.06**	**0.09**	**0.10**	**0.08**	KI	
0.29	0.29	0.29	0.28	0.23	0.33		0.05	0.04	0.06	**0.06**	**0.06**	**0.05**	0.04	**0.05**	**0.08**	0.06	0.07	0.07	OA	WP
0.24	0.15	0.37	0.21	0.19	0.28	0.32		0.03	0.05	0.02	**0.05**	**0.05**	**0.05**	**0.06**	**0.06**	**0.06**	**0.10**	**0.07**	OL	
0.28	0.23	0.33	0.21	0.17	0.32	0.28	0.22		0.05	0.04	0.05	**0.06**	0.04	0.05	**0.06**	**0.07**	**0.09**	0.07	OO	
0.36	0.25	0.31	0.25	0.24	0.30	0.34	0.32	0.30		**0.05**	**0.07**	**0.06**	**0.04**	**0.06**	**0.06**	**0.08**	**0.08**	**0.07**	OW	
0.33	0.24	0.47	0.25	0.18	0.25	0.39	0.16	0.26	0.34		**0.05**	**0.07**	**0.05**	**0.07**	**0.08**	**0.08**	**0.10**	**0.07**	EP	
0.40	0.37	0.52	0.27	0.32	0.43	0.46	0.33	0.42	0.43	0.37		**0.08**	**0.05**	**0.07**	**0.09**	**0.07**	**0.10**	**0.07**	MW	
0.28	0.28	0.27	0.29	0.36	0.52	0.29	0.33	0.41	0.32	0.41	0.50		**0.03**	**0.03**	**0.08**	0.04	**0.08**	**0.05**	FC	
0.21	0.29	0.21	0.18	0.24	0.35	0.27	0.33	0.32	0.26	0.36	0.30	0.20		**0.03**	**0.07**	**0.04**	**0.06**	0.03	IA	
0.26	0.26	0.23	0.21	0.27	0.47	0.30	0.36	0.32	0.30	0.44	0.42	0.16	0.13		**0.08**	**0.06**	**0.08**	0.04	OF	
0.45	0.34	0.40	0.35	0.38	0.40	0.47	0.34	0.38	0.37	0.46	0.63	0.45	0.44	0.41		**0.10**	**0.10**	**0.09**	UI	FP
0.33	0.34	0.39	0.33	0.47	0.63	0.34	0.42	0.56	0.48	0.50	0.46	0.21	0.26	0.36	0.58		**0.10**	0.07	AW	
0.63	0.64	0.43	0.53	0.49	0.61	0.38	0.61	0.56	0.39	0.63	0.56	0.42	0.28	0.39	0.50	0.57		**0.10**	GM	
0.27	0.41	0.41	0.28	0.42	0.56	0.46	0.47	0.47	0.42	0.47	0.46	0.27	0.15	0.20	0.49	0.41	0.53		IL	

LnP(2) = −9713.158, SD = 1.79) for the combined wild and domesticated populations, separating the wild populations from the domesticated populations (Figure 3). Genetic distance across samples was not significantly explained by geographic distance (Mantel test $R^2 = 0.05$, $P = 0.144$).

Genotype data is available from the corresponding author upon request.

Discussion

The evaluation of wildlife resources is necessary for its sustainable use and conservation (Allendorf et al. 2013; Mills 2013). We examined the neutral genetic structure of grasscutter samples collected from the wild spanning six states and four ecozones in south-western Nigeria. In addition, we quantified diversity metrics from seven domesticated populations, mostly from Ibadan (Oyo State), where the highest concentrations of state-run grasscutter farms are located.

Our analysis of wild samples revealed a high level of heterozygosity (> 0.75) and allelic diversity (average number of alleles > 5.18; Table 1). These values were comparable to those screened in unrelated grasscutters from Ghana during the development of these microsatellite markers (Adenyo et al. 2012), and are within the range of other widespread rodent species sampled at comparable scales (e.g. Peakall et al. 2003; Crawford et al. 2009). The only published genetic study on grasscutters (Adenyo et al. 2013) identified high mitochondrial diversity within and among Ghanaian populations sampled across the same ecozones.

The estimate of pairwise genetic differentiation among wild populations indicate a moderate genetic differentiation among grasscutter populations (Balloux and Lugon-Moulin 2002). However, only a small proportion of microsatellite variation could be attributed to ecozone or among populations relative to among or within individuals (Table 2), suggesting little structuring by ecological zones. The PCoA plot (Figure 2) hints at some geographic patterning

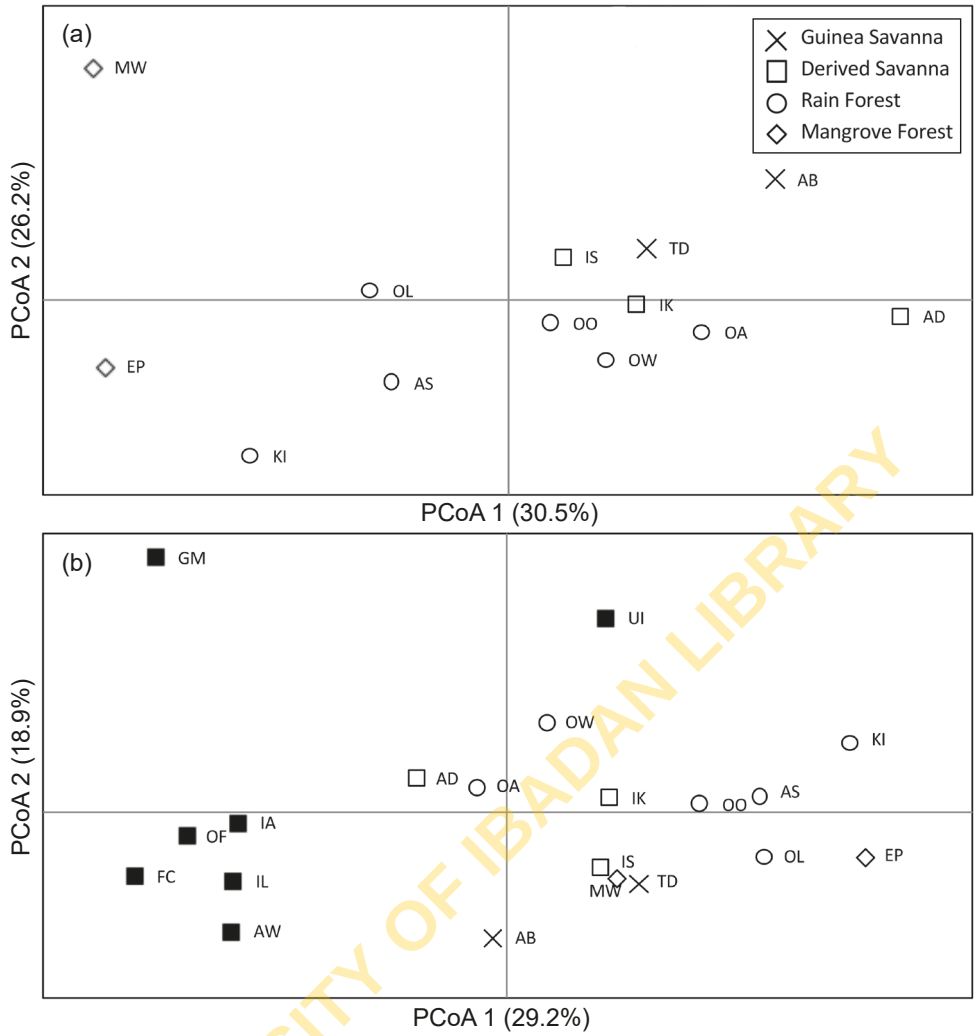

Figure 2: Graph of the first two axes from the principal coordinates analysis for all wild (a; samples > 5), and wild and domesticated samples (b; filled squares)

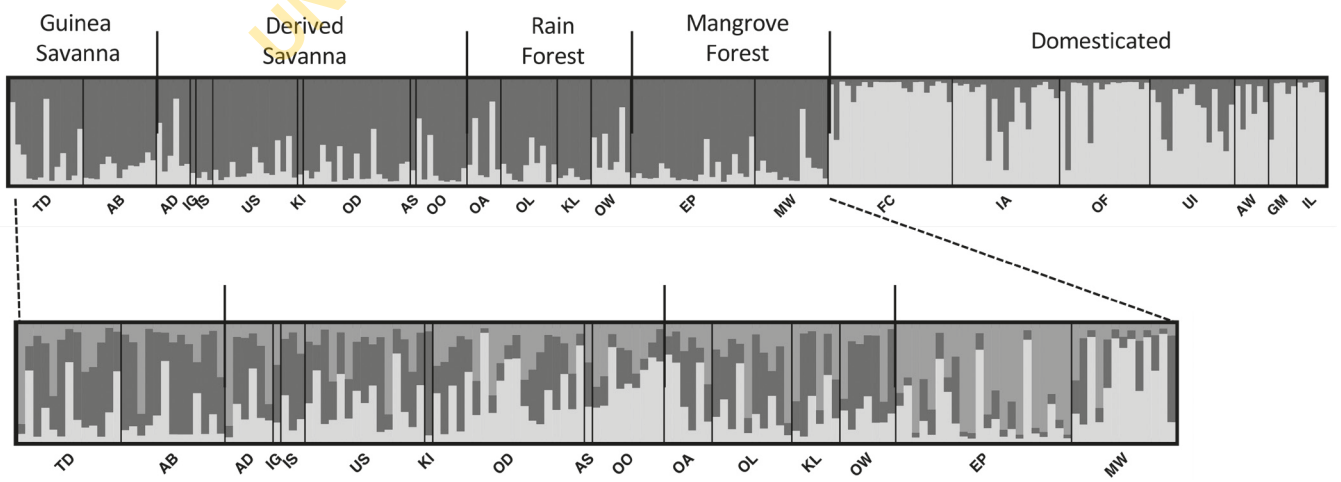

Figure 3: Structure results showing all samples (wild and domesticated; top), and wild only (bottom). For the total data set, $K = 2$ was chosen as the best representative model, $K = 3$ for wild considered separately, based on Delta K (see Supplementary Figure S1)

among samples from natural sites, with the first principal axis depicting a latitudinal pattern to variation across sites (however, this observation is tempered by small sample sizes for most ecozones). Ecozones, and ecozone boundaries in particular, have been examined as important drivers of speciation in some song birds (Smith et al. 1997), and behavioural evidence for dispersal preferences toward habitats similar to natal habitats is increasingly common (Davis and Stamps 2004; Fletcher et al. 2015). Where strong enough, this tendency might be expected to result in discrete genetic populations based on ecozone (Sacks et al. 2004). Alternatively, adaptation to differences among ecozones could drive genetic structuring (Rousett 2004). However, the limited genetic structuring among wild grasscutter samples and the clinal pattern observed in the PCoA suggests that the ecotone mosaic represented at this scale does not play an important role in limiting gene flow. One hypothesis for this may be the importance of riverine habitats and their distribution within the coastal region of south-western Nigeria.

Grasscutters' historical habitat was likely to be primarily restricted to riverine basins and other wet habitats (Feldhamer et al. 1999), forest edges and clearings where abundant food sources existed (Monadjem et al. 2015). The majority of Nigeria's hydrological resources exist in a north–south orientation, which would allow for continuous gene flow across forest ecozones. In addition, expansion of their distribution to human-altered habitats (clearing and irrigation in agricultural fields, and road edges) has likely positively influenced their dispersal ability.

There is very little ecological literature on this species relative to production-related research (Monadjem et al. 2015). Evaluation of standing genetic patterns can reveal information about processes that shape genetic variation, and provide a comparative framework with which to examine the genetic impacts of farming practices. Domestication of animals has the potential to decrease genetic variation compared with their wild ancestors because of founder effects, genetic drift and selection in populations with small effective sizes (Bruford et al. 2003; Taberlet et al. 2008), and capable of producing behavioural, morphological and physiological changes as domestication progresses (Price 1984). In our study, domestication has had a clear impact on neutral genetic variation. Our data reflect strong differentiation between wild and domesticated grasscutters, though only subtle reductions in diversity in the latter. For example, HE remains above 0.7 in most domesticated samples, though allelic variation is lower than wild counterparts, as expected as allelic variation is expected to be reduced at a quicker rate than heterozygosity (Nei et al. 1975). Assuming the domesticated animals are not sourced from outside the geographic area of this study (i.e. a different gene pool), it appears that the amount of genetic diversity has been minimally impacted. Domesticated gene pools are becoming differentiated, though not in a uniform manner (Figure 2), which may reflect differences in farming practices or sources of livestock, or genome sampling. For example, GM and IL are the main geographic outliers of the domesticated samples, and are outliers on the PCoA analysis as well. GM is only represented by five genotypes, which may be driving its divergence (i.e. under-sampled). However, these two locations may be sourcing animals from different locations than the majority of sites in southern Oyo, and thus reflecting subtly different gene pools. Low divergence among regions may also reflect sampling, which in some cases was low. However, our power analysis suggests that differentiation exceeded that expected under drift alone for these markers.

The ongoing domestication of grasscutter populations could entail a reduction of genetic variation and an increase in inbreeding. Several studies have demonstrated these effects in domesticated animals (Bruford et al. 2003). Our study illustrates that wild and domesticated grasscutter populations have comparable levels of neutral genetic diversity, but are highly differentiated, reflecting the impact of founder effects and genetic drift. Over the past two decades, grasscutter farmers in Nigeria had been counselled to start their farms with stocks from multiple sources to maximise initial diversity of domesticated populations. Maintaining diversity within domesticated populations has long been recognised to serve to maintain productivity in captive populations of mammals, where productivity can decline with the loss of diversity and inbreeding (Falconer 1981), and potentially to improve the domesticated population resiliency to disease (Ralls and Ballou 1983). Domestication of plants and animals has been shown to be a potential threat to their long-term genetic viability (Taberlet et al. 2008). As such it would be worthwhile to avoid some of the pitfalls associated with both domestication of grasscutters and the over-harvesting of their wild populations.

Acknowledgements — We thank the following people for their assistance with sampling: Dr JO Abiola (Faculty of Veterinary Medicine, University of Ibadan), Mr Olusola Philips and Mr Emmanuel D Musa (Department of Wildlife and Ecotourism Management, University of Ibadan). OMC was sponsored by a nine-month Fulbright Visiting Scholar Program at the University of Florida.

References

Adenyo C, Hayano A, Inoue E, Kayang BB, Inoue-Murayama M. 2012. Development of microsatellite markers for grasscutter (*Thryonomys swinderianus*, Rodentia) using next-generation sequencing technology. *Conservation Genetic Resources* 4: 1011–1014.

Adenyo C, Hayano A, Kayang BB, Owusu EH, Inoue-Murayama M. 2013. Mitchondrial D-loop diversity of grasscutter (*Thryonomys swinderianus* Rodentia: Hystricomorpha) in Ghana. *Open Journal of Animal Sciences* 3: 145–153.

Adenyo C, Ogden R, Kayang B, Onuma M, Nakajima N, Inoue-Murayama M. 2017. Genome-wide DNA markers to support genetic management for domestication and commercial production in a large rodent, the Ghanaian grasscutter (*Thryonomys swinderianus*). *Animal Genetics* 48: 113–115.

Adu EK. 2002. Research on grasscutter production in Ghana. In: Atta-Agyapong K, Weidinger R (eds), *Proceedings of workshop on promoting grasscutter production for poverty reduction in Ghana*. Accra: Qualitype Printing and Graphics. pp 27–33.

Adu EK, Alhassan WS, Nelson FS. 1999. Smallholder farming of the greater grasscutter, *Thryonomys swinderianus*, Temminck, in souther Ghana: a baseline survey of management practices. *Tropical Animal Health and Production* 31: 223–232.

Allendorf FW, Luikart G, Aitken SN. 2013. *Conservation and the genetics of populations*. Hoboken: John Wiley and Sons.

Asibey EOA. 1974. Wildlife as a source of protein in Africa south of the Sahara. *Biological Conservation* 6: 32–39.

Asibey EOA, Addo PG. 2000. The grasscutter, a promising animal for meat production. In: Turnham D (ed.), *African perspectives: practices and policies supporting sustainable development.* Harare: Weaver Press. p 120.

Balloux F, Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology* 11: 155–165.

Baptist R, Mensah GA. 1986. Benin and West Africa: grasscutter – farm animal of the future? *World Animal Review* 60: 2–6.

Bruford MW, Bradley DG, Luikart G. 2003. DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics* 4: 900–910.

Crawford JC, Zhiwei L, Nelson TA, Nielson CK, Bloomquist CK. 2009. Genetic population structure within and between beaver (*Castor canadensis*) populations in Illinois. *Journal of Mammalogy* 90: 373–379.

Davis JM, Stamps JA. 2004. The effect of natal experience on habitat preferences. *Trends in Ecology and Evolution* 19: 411–416.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

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

Falconer DS. 1981. *Introduction to quantitative genetics*. New York: The Ronald Press Company.

Feldhamer GA, Drickamer LC, Vessey SH, Merritt JF. 1999. *Mammalogy*. Boston: McGraw Hill.

Fletcher RJ Jr, Robertson EP, Wilcox RC, Reichert BE, Austin JD, Kitchens WM. 2015. Affinity for natal environments by dispersers impacts reproduction and explains geographic structure of a highly mobile bird. *Proceedings of the Royal Society B* 282: 20152845.

Goudet J. 2002. FSAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3.2. Available at www.unil.ch/izea/softwares/fstat.html [accessed 10 July 2010].

Hanotte O, Mensah GA. 2002. Biodiversity and domestication of ''non-conventional'' species: a worldwide perspective. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, 19–23 August 2002, Montpellier, France*. Castanet-Tolosan: INRA. pp 543–546.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

Jori F, Mensah GA, Adjanohoun E. 1995. Grasscutter production: an example of rational exploitation of wildlife. *Biodiversity and Conservation* 4: 257–265

Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60: 2399–2402.

Mills LS. 2013. *Conservation of wildlife populations: demography, genetics and management*. Hoboken: John Wiley and Sons.

Monadjem A, Taylor PJ, Denys C, Cotterill FPD. 2015. *Rodents of sub-Saharan Africa*. Berlin: DeGruyter.

Narum SR. 2006. Beyond Bonferroni: less conservative analyses for conservation genetics. *Conservation Genetics* 7: 783–787.

Nei M, Maruyama T, Chakroborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29: 1–10.

NRC (National Research Council). 1991. Grasscutter. In: *Microlivestock: little-known small animals with a promising economic future*. Washington, DC: National Academy Press. pp 233–239.

Peakall R, Ruibal M, Lindenmayer DB. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bushrat, *Rattus fuscipes. Evolution* 57: 1182–1195.

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

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.

Price EO. 1984. Behavioral aspects of animal domestication. *Quarterly Review of Biology* 59: 1–32.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population genetic structure using multilocus genotypic data. *Genetics* 155: 945–959.

Ralls K, Ballou J. 1983. Extinction: lessons from zoos. In: Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas WL (eds), *Genetics and conservation: a reference for managing wild animal and plant populations*. Menlo Park: Benjamin/Cummings. pp 164–184.

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

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.

Rousset F. 2004. *Genetic structure and selection in subdivided populations*. Princeton: Princeton University Press.

Ryman N, Palm S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes* 6: 600–602.

Sacks BN, Brown SK, Ernest HB. 2004. Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Molecular Ecology* 13: 1265–1275.

Smith TB, Wayne RK, Girman DJ, Bruford MW. 1997. A role for ecotones in generating rainforest biodiversity. *Science* 276: 1855–1857.

Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P. 2008. Are cattle, sheep, and goats endangered species? *Molecular Ecology* 17: 275–284.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources* 4: 535–538.

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

Received 6 March 2017, revised 26 August 2017, accepted 9 September 2017
Associate Editor: Victor Rambau