TITLE:

Characterization of microsatellite loci in the straw-colored fruit bat, *Eidolon helvum* (Pteropodidae)

AUTHORS:

Alison J. Peel^{1, 2*}

Stephen J. Rossiter³

James L.N. Wood¹

Andrew A. Cunningham²

David R. Sargan¹

AUTHOR ADDRESSES:

¹ Department of Veterinary Medicine, University of Cambridge, Cambridge, UK.

² Institute of Zoology, Zoological Society of London, London, UK.

³ School of Biological and Chemical Sciences, Queen Mary London, UK.

KEYWORDS (4-6):

African bats, Chiroptera, multiplex, polymorphic loci, X-linked

CORRESPONDING AUTHOR:

Alison J. Peel

Department of Veterinary Medicine, University of Cambridge, Cambridge, UK.

Tel: +44 1223 764 958

Fax: +44 1223 764 667

Email: alisonpeel@gmail.com

RUNNING TITLE

Eidolon helvum fruit bat microsatellites

Word count (including abstract and captions): 1000

ABSTRACT

Twenty microsatellite loci developed for the straw-colored fruit bat, *Eidolon helvum*, are described. These markers were used in multiplex PCRs to amplify genomic DNA from 142 individuals sampled from nine populations across Africa. Nineteen loci were polymorphic, with a mean number of alleles per locus of 16.2 (3-46). Observed and expected heterozygosity values ranged from 0.01 to 0.90 and 0.03 to 0.97, respectively. These markers will be used in combination with mitochondrial DNA sequences to investigate gene flow and the genetic metapopulation structure of *E. helvum*.

TECHNICAL NOTE

The straw-colored fruit bat, *Eidolon helvum* (Order Chiroptera, Family Pteropodidae) is a gregarious, migratory, tree-roosting species, which forms very large seasonal colonies across sub-Saharan Africa (Funmilayo 1979; Mutere et al. 1980; Richter & Cumming 2006), often near to large human populations (Accra, Kampala) (Kingdon 1974; DeFrees & Wilson 1988).

To date, the drivers, routes and patterns of migration in this species are unknown. Indeed, it is unclear whether *E. helvum* has a panmictic population structure across a continuous distribution that shifts according to seasonal resource availability, or whether a finer substructure and specific migration routes exist. As this species is a suspected reservoir for potentially-zoonotic viruses, data on its gene flow and metapopulation structure are urgently required for investigations into viral transmission dynamics within and between populations.

Here we describe a set of microsatellite markers that will be used in combination with mitochondrial DNA sequences to investigate the genetic metapopulation structure of *E. helvum*.

For library development, *E. helvum* muscle tissue samples were collected from dead bats in Ghana and stored in 70 % ethanol. Genomic DNA was extracted using a standard phenol–chloroform method. Four genomic libraries, enriched for two dinucleotide motifs (CA, GA) and two tetranucleotide motifs (AAAC, TAGA), were prepared from pooled genomic DNA by Genetic Identification Services (GIS, Chatsworth, CA, USA). Libraries were constructed following Jones et al. (2002). Primers were designed for 89 microsatellite loci using DesignerPCR 1.03 (Research Genetics, Inc.), and a subset of these (n = 33) was selected for screening.

For genotyping, genomic DNA was extracted from wing membrane biopsies (4-mm) or liver samples from 142 adult bats from nine locations across continental Africa using

DNeasy Blood and Tissue Kit (QIAGEN Ltd., Crawley, West Sussex, UK). Marker amplification was assessed in a subset of samples ($n = 8$) in 10 μ l PCRs, containing 4 ng template DNA, 0.2 μ M of each primer, and 5 μ L 2X Type-it Multiplex PCR Master Mix (QIAGEN Ltd.). Amplification was performed using the following conditions: 5 min at 94°C; 30 cycles of 30 sec at 95 °C, 90 sec at 57 °C, and 30 sec at 72 °C; then 30 sec at 60 °C. PCR products were initially checked by gel electrophoresis on 2% Agarose gels, run at 160 V for 60 min to ensure a single locus was amplified. Loci that did not consistently amplify, or amplified more than one band, were discarded.

Primers that produced a clean product were selected and forward primers extended with an appropriate 5' 19mer to allow indirect end labeling using an M13, T7 or SP6 fluorescent labeled primer. Initially, 30 markers were assessed for polymorphism in a subset of samples ($n = 24$). Genotyping was performed by capillary electrophoresis using a Beckman CEQ 8000. Allele sizes were scored automatically prior to manual verification. Ten loci were discarded due to scoring difficulties (stuttering and ambiguous peaks). One locus (M) was monomorphic and was retained as an internal control. Remaining samples were run as multiplex PCRs and genotyped at 20 loci (Table 1). Genotyping error rates (calculated at the allelic level) were assessed by replicate extractions (1%), PCRs (20%), and allele scorings (10%), replicates by different users, and inclusion of positive and negative controls (Bonin et al. 2004).

Observed and expected heterozygosity values were calculated and each locus was tested for evidence of departure from Hardy-Weinberg equilibrium (HWE) and for genotypic disequilibrium using FSTAT 2.9.3 (Goudet 1995), with appropriate Bonferroni corrections for multiple testing. All loci were analyzed in MICRO-CHECKER 2.2.3 (van Oosterhout et al. 2004) to test for null alleles, stuttering and large allelic dropout as a cause of departure from HWE.

All 142 samples were scored at 20 microsatellite loci. The mean number of alleles per locus was 16.2 (3-46). Observed and expected heterozygosity values ranged from 0.01 to 0.90 and 0.03 to 0.97, respectively (Table 1). Six loci showed heterozygote deficiencies

resulting in significant deviation from HWE, four of which (F, S, Ac and Ae) showed deviation in one to two geographically separate populations, and two (B and Ad) in five to six populations. B was homozygous in 100 % of known males ($n = 41$), and 16 % ($n = 6$) of females, suggesting it was X-linked, whereas deviation in Ad appeared to be due to the presence of null alleles. Eight other loci (T, S, F, K, Ac, Ai, Ah and Ag) were flagged as having potential null alleles. Locus M was essentially invariant, with 97% of individuals homozygote for a single allele. No loci showed significant genotypic disequilibrium.

The overall genotyping error rate was 1.8 %, comprising 0 % extraction error, 2.5 % amplification and scoring error, and 1.2 % human error. Two loci (E and Ag) contributed substantially to the error rates, the former due to scoring difficulties (alleles 1 base pair apart) and the latter due to allelic dropout. The remaining 18 loci gave a mean amplification and scoring error rate of 1.8 %.

Overall, 17 loci were identified which are in HWE and sufficiently polymorphic for ongoing studies to investigate gene flow and the genetic metapopulation structure of *E. helvum*.

ACKNOWLEDGEMENTS

The authors thank David Hayman for his initial work in Ghana and for providing tissue samples. We greatly appreciate assistance from the Kasanka Trust, Zambian Wildlife Authority, National Research Council of Malawi, Museums of Malawi, Malawi Department of Animal Health, Tanzanian Commission for Science and Technology, Tanzanian Wildlife Research Institute, Ugandan National Council for Science and Technology, and Makerere University. Additional samples were kindly provided by Javier Juste, Ivan Kuzmin, Guy Gembu Tungaluna and the Californian Academy of Sciences. We thank Jesús Aguirre for advice and technical assistance, and Melissa Nollet for processing Zambian samples. This work was funded by the Cambridge Infectious Disease Consortium, Charles Slater Trust, Jowett Fund and the ZSL Daisy Balogh Fund.

REFERENCES

- Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261.
- DeFrees S, Wilson D (1988) *Eidolon helvum*. *Mammalian Species*, **312**, 1-5.
- Funmilayo O (1979) Ecology of the Straw-coloured Fruit Bat in Nigeria. *Revue de zoologie africaine*, **93**, 589-600.
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. In: *Journal of Heredity*.
- Kingdon J (1974) *East African mammals: an atlas of evolution in Africa*.
- Mutere F, Wilson D, Gardner A (1980) *Eidolon helvum* revisited. In: *Proceedings of the 5th International Bat Research Conference*, pp. 145-150.
- Richter H, Cumming G (2006) Food availability and annual migration of the straw-colored fruit bat (*Eidolon helvum*). *Journal of Zoology*, **268**, 35-44.
- van Oosterhout C, Hutchinson W, Wills D, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. In: *Mol Ecol Notes*, pp. 535-538.

Table 1 Characterization of 20 microsatellite loci for the straw-colored fruit bat, *Eidolon helvum*.

Multiplex mix	Locus name	Genbank Accession no	Primer sequence (5'-3') ^a F: Forward, R: Reverse	No of alleles	Allele size range (bp)	H _o	H _e	n	Repeat motif
TSY	T (EhD7)		F: SP6-TGCCTATCTGCTTGTCTGTC R: AGGACCTCAAGTTTCATCAAAG	10	113-142	0.72	0.82	141	(TAGA) _n
	S ^b (EhD3)		F: M13-CGCTATGAGACCTTCAGTAAG R: TGTCCATATTCTGCATGTC	21	167-250	0.76	0.90	141	(TAGA) _n
	Y (EhD113)		F: T7-CCATAAACAGTGCTATACCACA R: AAACCGATGAAGCCATTC	15	259-284	0.80	0.86	141	(TAGA) _n
FWB	F ^b (EhA6)		F: SP6-GCTTGTCAAAATGGTCAAAAC R: GATGGTGAGGCAGTAGATGTG	18	157-189	0.86	0.91	141	(CA) _n
	W (EhD108)		F: M13-ACCTTCTGCTTTTCACCACCAC R: TCCCATAGACCCTCCTTGAG	20	167-244	0.90	0.90	142	(TAGA) _n
	B ^c (EhA4)		F: T7-CAACAGGGAAGGATCTGTGTC R: TTCCAATTTCATGTTTAGGTC	12	284-312	0.35	0.85	142	(CA) _n
MNQX	M (EhC3)		F: SP6-TAAGGAGTTGATACCCAACATG R: TTTGGAGAAATGCTCTCTCATG	2	131-135	0.01	0.03	142	(AAAC) _n
	N (EhC105)		F: SP6-ACCTTCCTTTTGTCTTGACTG R: AACCAGGGCTTTCAATGG	7	181-201	0.71	0.64	142	(AAAC) _n
	Q (EhC120)		F: M13-GATTTCATGCTTCTTCAACAT R: AAATCTGTCCACCCTGCTCTAG	7	198-236	0.42	0.44	142	(AAAC) _n
	X (EhD110)		F: T7-TCAGCCTCCATAATCCTATG R: CCCTCCTTTGGACTTTGTA	9	290-322	0.71	0.79	138	(TAGA) _n
EPK	E (EhA5)		F: SP6-TTGACACAACATGCGAATA R: TTGGCGTTCAGTTATCAG	25	172-203	0.90	0.93	131	(CA) _n
	P (EhC119)		F: M13-CTATCAGAGGGCACATGTATC R: GCACACCACTCACTCAGTTC	3	250-258	0.28	0.35	131	(AAAC) _n
	K (EhB114)		F: T7-CAAAAGGGAAATGGATGTAATC R: GGAAAGTGGAATAAGAAACACC	24	265-323	0.74	0.93	118	(GA) _n
AcAfAi	Ac ^b (EhA115)		F: T7-AATAGAAGCGAGGCTGATG R: TCTCCCCACACTCTCTCCAG	13	213-238	0.74	0.88	141	(CA) _n
	Af (EhD106)		F: M13-TTGGCAGATGACTCTCTTGA R: CAGATTATGGGACTTCACAACA	11	276-316	0.81	0.82	138	(TAGA) _n
	Ai (EhA11)		F: SP6-CCAAGCAGACTCTCTCACC R: CCCTACAGGATACCAAGTGTG	12	302-324	0.83	0.89	136	(CA) _n
AdAh	Ad ^c (EhA107)		F: M13-TCCTGCTCCATTCTACCC R: GCACTCTATCCATCTGTCTCTC	46	232-336	0.53	0.97	132	(CA) _n
	Ah (EhD115)		F: SP6-CTACGGATATTAAGTGGCAGAA R: TGGTTTGGTGATAAGTGTCAG	17	219-252	0.74	0.91	138	(TAGA) _n
AeAg	Ae ^b (EhC4)		F: M13-GCGAGAGACTCATCGGAG R: CCCTGCTGTTCTTTACGC	8	230-259	0.16	0.64	90	(AAAC) _n
	Ag (EhD112)		F: SP6-GAACGGATGAAGATGTGGTAC R: TCTGGCAACCATAGTTTATTC	28	281-344	0.88	0.91	125	(TAGA) _n

^a Fluorochromes added to 5'-end of forward primer: T7 (TAATACGACTCACTATAGGG), SP6 (ATTAGGTGACACTATAGAA), M13 (CAGGACGTTGTAAAACGAC). H_o: Observed heterozygosity, H_e: Expected heterozygosity, n: Sample size. ^b indicates significant deviation from HWE in 1-2 populations, and ^c indicates deviation at 5-6 populations (p<0.05 adjusted with Bonferroni correction).