

Molecular Phylogeny of Duiker Antelope (Mammalia: Cephalophini)

Molecular Phylogeny of Duiker Antelope (Mammalia: Cephalophini)

by

Bettine Jansen van Vuuren

**Submitted in partial fulfillment of the
requirements for the degree of**

**Doctor of Philosophy
(Zoology)**

**in the
Faculty of Biological and Agricultural Science
University of Pretoria
Pretoria**

October 1999

Molecular Phylogeny of Duiker Antelope (Mammalia: Cephalophini)

by

Bettine Jansen van Vuuren

Supervisor: Professor T.J. Robinson
Department of Zoology & Entomology
University of Pretoria
Pretoria

ABSTRACT

Molecular sequence data, comparative cytogenetics, and fluorescence *in situ* hybridization (FISH) were employed to study evolutionary relationships within the duiker antelope (tribe Cephalophini). The results of the different data sets are highly concordant. Mitochondrial DNA sequences (the full cytochrome *b* gene and a 767 bp portion of the 12S rRNA) were analyzed from all 19 recognized species. Complete species representation and good cytochrome *b* resolution allowed for the retrieval of four adaptive lineages, the conservative dwarfs (*Cephalophus monticola*, *C. maxwellii*) which were the most basal clade, the savanna specialist (*Sylvicapra grimmia*) which groups apart from all the forest duiker, the giant duiker group (*C. silvicultor*, *C. spadix*, *C. dorsalis*, *C. jentinki*), and the red duiker lineage (*C. leucogaster*, *C. rufilatus*, *C. nigrifrons*, *C. natalensis*, *C. harveyi*, *C. callipygus*, *C. weynsi*, *C. ogilbyi*, *C. rubidus*, *C. niger*). The placement of the endangered *C. zebra* and the enigmatic *C. adersi* remains obscure.

Conventional chromosome banding showed a $2n=60$ complement in *C. spadix* extending previous observations that speciation in duiker antelope does not involve euchromatic rearrangements or variation in diploid number. At a finer level, fluorescence *in situ* hybridization with species specific satellite fragments derived from the chromosomal DNA of *C. maxwellii* and *C. monticola*, resulted in intense fluorescence to the centromeric regions of the autosomes of all species (*S. grimmia*, *C. dorsalis*, *C. maxwellii*, *C. monticola*, *C. natalensis*, *C. silvicultor*, *C. spadix*). However, variation in hybridization to the X and Y chromosomes allowed for some distinction among taxa. These results are consistent with the delimitation of the four adaptive groups suggested by molecular analysis and the published morphological data which, when taken together, question and in some instances support several of the nomenclatural divisions in current use in duiker taxonomy. These include the recognition of *Philantomba* as genus name for *C. monticola* and *C. maxwellii*, an arrangement that would secure *Cephalophus* monophyly, and that *C. harveyi* be relegated to a subspecies of *C. natalensis*.

Keywords Duiker, Cephalophini, systematics, phylogeny, mitochondrial DNA, cytochrome *b*, 12S rRNA, biogeography, comparative cytogenetics, fluorescence *in situ* hybridization.

ACKNOWLEDGMENTS

I am greatly indebted to my supervisor Professor C.J. Robinson for his constant and sustained encouragement during the course of my postgraduate studies. His constant enthusiasm for science in general, and for this project in particular, was an important part of the constructive criticism and thoughtful comments that laid the foundation for much of this work. I am also grateful for the opportunities he provided for me to present my research at the African Medical Congress held in Santiago de Cuba, held in 1997, and for the 4th International Symposium on African Small Mammals, Durban, 1998.

My sincere appreciation is extended to the following people who assisted with the design and production of laboratory dress, sunglasses, and other equipment: P. Christiansen (University of Copenhagen, Denmark), J. Maruy (University of the Witwatersrand, Johannesburg), M. Van der Merwe (Dordeville Zoological Gardens, Pretoria), M. Zuber (University of Zurich, Switzerland), L. Chivers (University of Guelph, Ontario), M. Åberg (Swedish Museum of Natural History, Stockholm), M. Hamman (University of Texas Medical School, Houston, Texas), M. J. S. de la Torre (Hospital General de la Ciudad de Mexico, Mexico), K. Huterea (Albertus Rex, Marburg, Germany), M. S. J. van der Laan (Museum London, United Kingdom), P. Kelly (British Museum, London), M. C. M. van de Kerkhof (Natural History Museum, London), G. van der Kooij (K. Van der Stuyfendael Congo Safari, South Africa), N. Mnyoni (Smithsonian Museum of Natural History, Washington, D.C., USA), T. Swaney (Cape Town, South Africa), M. Viljoen (Fynbos, South Africa), L. Vrolijk (University of Togo, Lomé), A. Williams (Zanzibar Protected Areas Project, Dar es Salaam), V. Williams (Orphan of Wildlife Orphanage, Zimbabwe), and Kevin Konck-Sarik for assistance in the field.

For Mom, Dad, and Miriam

with love

I also thank my dear friend Wilbur Hamman (University of Texas Medical School) for his assistance with fluorescence microscopy techniques. Most of the experiments were conducted while Wilbur was a Acting technician in Terry's laboratory. I also thank Victor Rando for his help and endless patience with the retroviral experiments. The following experiments were carried out in the laboratories of Prof. van den Brink (University of Pretoria) and V. Mizerah (South African Medical Research Council, Pretoria). My dear friend J. Waal is thanked for her time and assistance. Oliver Wittekamp (University of Hamburg)

ACKNOWLEDGMENTS

I am greatly indebted to my supervisor Professor T.J. Robinson for his guidance and sustained encouragement during the course of my postgraduate studies. Terry's unflagging enthusiasm for science in general, and for this project in particular, was truly inspirational. His constructive criticism and thoughtful comments laid the foundation for much of this work. I am also grateful for the opportunities he provided for me to participate in the Euro-American Mammal Congress held in Santiago de Compostela, Spain in 1998 and the 8th International Symposium on African Small Mammals, Paris, France, earlier this year.

My sincere appreciation is extended to the following people and institutions for time and assistance in obtaining duiker samples: Aarn (Australia), P. Arctander (University of Copenhagen, Denmark), P. Bishop (University of the Witwatersrand, South Africa), The Brownsville Zoological Gardens (United States of America), C. Claude (University of Zürich, Switzerland), E. Douzery (University of Montpellier, France), B. Fernholm (Swedish Museum of Natural History, Sweden), P. Grubb (London, United Kingdom), W. Harrison (University of Texas Medical School, United States of America), K. Heckler (Van Rooyen Taxidermists, South Africa), B. Herzig (Vienna Museum for Natural History, Austria), R. Hutterer (Alexander Koenig Museum, Germany), P. Jenkins (Natural History Museum London, United Kingdom), C. Powell (University of Port Harcourt, Nigeria), R. Sabin (Natural History Museum London, United Kingdom), E. Stockenstrom (Congo Safaris, South Africa), S. Strand (Swedish Museum of Natural History, Sweden), C. Stuart (Loxton, South Africa), M. Viljoen (Pretoria, South Africa), L. Vinciguerra (Safariroyal, Tanzania), A. Williams (Zanzibar Protected Areas Project, Tanzania), V. Wilson (Chipangali Wildlife Orphanage, Zimbabwe), and Kevin Kruickshank for assistance in the field.

My thanks go also to Wilbur Harrison (University of Texas Medical School, Houston) for assistance with fluorescence *in situ* hybridization. Most of the hybridization results were generated while Wilbur was a visiting scientist in Terry's laboratory. I also wish to thank Victor Rambau for his help and endless patience with the cytogenetic aspect of the study. Cloning experiments were carried out in the laboratories of Professors H. Huismans (University of Pretoria) and V. Miserahi (South African Medical Research Council). Pamela de Waal is thanked for her time and assistance. Oliver Wittekindt (Institute for Human

Genetics & Anthropology, University of Freiburg) cloned and sequenced the satellite fragments used in this study; I am greatly indebted to him for his time. Colin Groves, Peter Grubb, Ken Halanych, Conrad Matthee, and Marc Robinson are thanked for comments on various aspects of the study.

My sincere appreciation to Hanlie Engelbrecht and her family, Stefanie Freitag, Warren Funston, Conrad Matthee, Sonja Matthee, Sarita Maree, Ben Orban, and Victor Rambau for their continued friendship, interest, and words of encouragement. Babs Potgieter and Simon Maake are thanked for technical assistance, and for keeping me on my toes.

Lastly, my special thanks to Mom, Dad, and Miriam for loving and supporting me throughout.

LIST OF FIGURES

CHAPTER 1

GENERAL INTRODUCTION

AIM AND OBJECTIVES OF THE STUDY

INTRODUCTION

<i>Historical history of <i>Uta stansburiana</i></i>	
<i>Higher order taxonomy</i>	
<i>Evolution of life cycle stages</i>	14
<i>Geography</i>	22
<i>Conservation status</i>	23

CHAPTER 2

MOLECULAR PHYLOGENY OF THE CENTRAL SPECIES

OBJECTIVES

INTRODUCTION

MATERIAL AND METHODS

<i>Sample collection and storage</i>	27
<i>Genomic DNA extraction and restriction sequencing</i>	27

TABLE OF CONTENTS

ABSTRACT	2
ACKNOWLEDGMENTS	5
TABLE OF CONTENTS.....	7
LIST OF TABLES	10
LIST OF FIGURES.....	12
CHAPTER 1	
GENERAL INTRODUCTION.....	16
AIM AND OBJECTIVES OF THE STUDY	16
INTRODUCTION	17
<i>Natural history of duiker antelope</i>	17
<i>Higher order taxonomy</i>	18
<i>Taxonomy of the Cephalophini</i>	19
<i>Speciation</i>	22
<i>Conservation status</i>	23
CHAPTER 2	
MOLECULAR PHYLOGENY OF THE CEPHALOPHINI.....	25
OBJECTIVES	25
INTRODUCTION	25
MATERIAL AND METHODS	27
<i>Sample collection and storage</i>	27
<i>Genomic DNA extraction and nucleotide sequencing</i>	27

<i>Sequence analysis</i>	31
RESULTS AND DISCUSSION	35
<i>Saturation analysis</i>	35
<i>Rate heterogeneity</i>	42
<i>Cytochrome b</i>	42
<i>Cytochrome b phylogeny</i>	44
<i>12S rRNA</i>	51
<i>12S rRNA phylogeny</i>	51
<i>Combined analysis</i>	54
<i>Speciation events</i>	57
<i>Taxonomic implications</i>	59
<i>Recognition of the genus Philantomba</i>	59
<i>Phylogenetic placement of C. zebra and C. adersi</i>	59
<i>Subspecies status of C. harveyi</i>	60

CHAPTER 3

COMPARATIVE AND MOLECULAR CYTOGENETICS	61
OBJECTIVES.....	61
INTRODUCTION.....	61
MATERIAL AND METHODS.....	64
<i>Specimens</i>	64
<i>Comparative cytogenetics</i>	64
<i>Harvesting and slide preparation</i>	64
<i>Banding of chromosomes</i>	64
<i>Molecular cytogenetics</i>	66
<i>Isolation of satellite repeat sequences for FISH</i>	66
<i>Fluorescence in situ hybridization</i>	66
<i>Southern blot hybridizations</i>	67
<i>Cloning and sequencing</i>	67
RESULTS AND DISCUSSION	69
<i>C. spadix karyotype analysis</i>	69
<i>G-banded karyotype and X chromosome morphology</i>	69
<i>C-banding</i>	69
<i>Molecular cytogenetics</i>	72

<i>Satellite fragments</i>	72
<i>FISH</i>	72
<i>Southern blot analysis</i>	76
<i>Characterization of the satellite family</i>	78
<i>Phylogenetic implications</i>	81

CHAPTER 4

COMBINED APPROACH, CONSERVATION GENETICS, AND GENERAL CONCLUSIONS..... 83

TOTAL EVIDENCE.....83

GENETICS AND WILDLIFE CONSERVATION.....84

CONCLUSION.....87

SUMMARY..... 89

OPSOMMING..... 91

REFERENCES..... 93

LIST OF TABLES

- Table 1** Taxonomic classification of the tribe Cephalophini (Bovidae) following Grubb (1993). The synonyms listed in Grubb's revision are presented under earlier classifications. 21
- Table 2** Duiker species used in the present study. Vernacular names, collection localities, and tissues used for DNA extractions are provided. Where multiple specimens were sequenced per species, those included in the phylogenetic analyses are indicated by an * (see text for details). 28
- Table 3** Primers used and their sequences for both amplification and sequencing of duiker DNA. Biotin coated end primers are indicated by *. Customized primer numbers correspond to *Bos taurus* (Anderson *et al.* 1982) and the sequences are reported 5' to 3'. 30
- Table 4** Observed cytochrome *b* pairwise distances (%) between all taxa included in the study. Divergence values above the diagonal are uncorrected p-distances, while the HKY (Hasegawa *et al.* 1985) values are shown below the diagonal. Abbreviations used to indicate sampling localities are: Cam=Cameroon; Gab=Gabon; SA=South Africa; CAR=Central African Republic; Zim=Zimbabwe; Con=Democratic Republic of Congo. 40
- Table 5** Pairwise distances (%) calculated for the small ribosomal subunit. Divergence values above the diagonal line are uncorrected p-distances while HKY (Hasegawa *et al.* 1985) values are shown below the diagonal line. Abbreviations used are: Cam=Cameroon; Gab=Gabon; SA=South Africa; Zim=Zimbabwe; Con=Democratic Republic of Congo. 41

Table 6 Base composition for the complete cytochrome *b* gene, as well as first, second, and third codon positions. Base frequencies were calculated in Paup 4.0b2a (Swofford 1999). Abbreviations used to indicate sampling localities are: Cam=Cameroon; Gab=Gabon; SA=South Africa; CAR=Central African Republic; Zim=Zimbabwe; Con=Democratic Republic of Congo. 43

Table 7 Base composition for the small ribosomal subunit. The composition of the stem and loop domains are also indicated. Frequencies were calculated in Paup 4.0b2a (Swofford 1999). Abbreviations used to indicate sampling localities are: Cam=Cameroon; Gab=Gabon; SA=South Africa; CAR=Central African Republic; Zim=Zimbabwe; Con=Democratic Republic of Congo. 52

Table 8 The seven duiker species used in comparative and molecular cytogenetic analysis. Vernacular species names and their collection localities are provided. 65

Table 9 Five FISH and comparative cytogenetic characters used in the combined analyses. The characters are coded as presence/absence data. The cytogenetic data of Robinson *et al.* (1996) is included. 85

LIST OF FIGURES

- Figure 1** Saturation plots for all duiker taxa included in the present study. The actual number of transitions and transversions are plotted against uncorrected p-distances. Closed symbols indicate ingroup comparisons while open symbols represent ingroup to outgroup comparisons. (A) Complete cytochrome *b* gene (B) Cytochrome *b* first codon position (C) Cytochrome *b* second codon position (D) Cytochrome *b* third codon position (E) Complete 12S rRNA gene (F) 12S rRNA loop domain. 36-39
- Figure 2** The most parsimonious tree obtained under the empirical cytochrome *b* ti/tv weighting of 13:1. This tree is based on 1 140 bp and has a length of 2 739 steps. Nonparametric bootstrap values obtained after 1 000 iterations are shown above branches while the number of unambiguous characters that occur along branches are shown below them. Bootstrap values for nodes that received less than 50% support are not indicated. 45
- Figure 3** The optimal maximum likelihood topology inferred from the cytochrome *b* data. The tree was constructed in Paup 4.0b2a (Swofford 1999) under the HKY (Hasegawa *et al.* 1985) model of sequence evolution with rate heterogeneity (gamma shape parameter=0.168), empirical ti/tv ratio (13:1), and nucleotide frequencies calculated from the data. Bootstrap support, obtained from 100 iterations, is shown. 47
- Figure 4** Neighbor joining tree constructed in Paup 4.0b2a (Swofford 1999) from 1 140 bp of the cytochrome *b* gene. The HKY (Hasegawa *et al.* 1985) correction was applied to the distance matrix, with equal rates of evolution assumed for all lineages. The branch lengths are drawn proportionally to the amount of change occurring along them, and nonparametric bootstrap support is indicated. 49

- Figure 5** The cytochrome *b* parametric bootstrap topology obtained from 1 000 simulated data sets created by the Siminator computer program (Huelsenbeck *et al.* 1996). The most parsimonious tree (Fig. 2) was used as the model tree, with the gamma shape parameter (0.168) and kappa value (27.167) estimated *via* maximum likelihood. The simulated data sets were developed under the HKY (Hasegawa *et al.* 1985) model of evolution. 50
- Figure 6** Strict consensus tree of 14 equally most parsimonious trees obtained from 767 bp of the 12S rRNA gene under a 2:1 ti/tv ratio. Nonparametric bootstrap values for nodes are presented above branches and the number of unambiguous changes that plot along them are indicated below the branches. No changes are shown on branches within polytomies. 53
- Figure 7** Neighbor joining tree constructed from 767 bp of the small ribosomal subunit in Paup 4.0b2a (Swofford 1999). The HKY correction (Hasegawa *et al.* 1985) was applied to the distance matrix. Branches are drawn proportionally to the amount of change occurring on them, and nonparametric bootstrap support for nodes is indicated. 55
- Figure 8** Maximum likelihood topology obtained from the 12S rRNA gene. The tree was constructed in Paup 4.0b2a (Swofford 1999) under the HKY (Hasegawa *et al.* 1985) model of sequence evolution with equal rates, the empirical ti/tv ratio (7:1), and nucleotide frequencies calculated from the data. Bootstrap support was obtained using 100 iterations. 56
- Figure 9** G-banded female karyotype of *C. spadix* (2n=60) from Tanzania. 70
- Figure 10** C-banded metaphase cell of a *C. spadix* female. Arrowheads indicate the X chromosomes. 71

Figure 11 Digestion profiles of 5 duiker species. The unique satellite fragments that were selected as FISH probes are indicated by the arrows. (A) 1.4 kb fragment (lane 2) released in *C. monticola* by *Pst*I digestion, (B) 2.8 kb satellite fragment isolated from *C. maxwellii* (lane 3) after digestion with *Pvu*II, (C) two 2.8 kb fragments (lanes 2 and 3) isolated from *C. monticola* and *C. maxwellii* after digestion with *Eco*RI.

73

Figure 12 FISH results showing hybridization patterns of the four probes isolated from *C. monticola* and *C. maxwellii*. Identical hybridization patterns were found using all four painting probes (see text). As a representative, the *Eco*RI-Max probe results are illustrated here. The X and Y chromosomes, where present and identifiable, are indicated by arrows. (A) *C. maxwellii* and *C. monticola* (shown) had identical FISH patterns; (B) FISH patterns characteristic of *S. grimmia* and *C. natalensis* (shown); (C) FISH patterns characteristic of *C. silvicultor*, *C. spadix* and *C. dorsalis* (shown); (D) absence of signal following FISH to the metaphase chromosomes of cattle (outgroup). Only a partial spread for *B. taurus* is presented.

74-75

Figure 13 Southern blot of *Cfo*I digested DNAs of seven duiker species and an outgroup species after probing with the *C. maxwellii* *Eco*RI fragment. Lane 1=*S. grimmia*; lane 2=*C. maxwellii*; lane 3=*C. monticola*; lane 4=*C. natalensis*; lane 5=*C. dorsalis*; lane 6=*C. silvicultor*; lane 7=*C. spadix*; lane 8=*B. taurus*.

77

Figure 14 Cloned *Eco*RI-Max fragments released after digestion by *Eco*RI. Clones with an insert size spanning approximately 0.6 kb to 0.8 kb were selected for sequencing (lanes A7, B4, B5, B8). The Pharmacia kilobase DNA marker (lanes A1 and B1) was used as the size standard.

79

Figure 15 Partial sequence of four cloned *Eco*RI-Max fragments.

80

Figure 16 Results of the unordered parsimony analyses based on the combined molecular and cytogenetic data sets. (A) The bootstrap consensus tree (1 000 replicates) obtained when all duiker taxa are included. Nodes that received less than 50% support were collapsed. (B) The single most parsimonious tree obtained when only taxa for which cytogenetic data were available are included. The tree was produced with the heuristic search option in Paup 4.0b2a (Swofford 1999) with 100 random replacements. Bootstrap support obtained from 1 000 iterations is shown above the branches.

86

AIMS AND OBJECTIVES OF THE STUDY

In the past, several taxonomic relationships among duiker subgroups (subgenus *Cephalophus* have followed almost exclusively on the morphological and/or morphological and chromosomal characteristics of the group (e.g. Leys 1969, Thebaud 1975, & Quilley 1977, 1981; Kingdon 1982, 1997). This approach was, however, often limited by the number and the resolution of characters and the use of non-molecular characters. It has an attempt to clarify interrelationships of the phylogeny of the subgenus *Cephalophus* was first examined using molecular and cytogenetic data. This work was done on the species and phenotypically divergent group of *M. capensis*.

The objectives were:

- (i) To assess the molecular phylogenetic relationships and interrelationships among the *Cephalophus* based on molecular (mtDNA) and 12S rDNA nucleotide data
- (ii) To determine if the occurrence of the highly conserved (15S) rDNA highly repeated satellite sequences differ between the species of *Cephalophus*
- (iii) To determine the degree of concordance in the phylogenetic placement of taxa with independent data sets (mitochondrial DNA nucleotide sequences, 12S rDNA nucleotide sequences and the fluorescence *in situ* hybridization patterns of chromosomes)
- (iv) To correlate the phylogeny of the taxa with available climatic data to gain insight into the events that have shaped the speciation and evolutionary events within the group