Phylogeny and historical demography of economically important rodents of the genus *Arvicanthis* (Mammalia: Muridae) from the Nile Valley: of mice and men

E. H. ABDEL RAHMAN AHMED¹, J.-F. DUCROZ², A. MITCHELL³†, J. LAMB⁴, G. CONTRAFATTO⁴, C. DENYS², E. LECOMPTE² and P. J. TAYLOR⁵*

¹Sudan Natural History Museum, Faculty of Science, University of Khartoum, Khartoum, Sudan
²Muséum National d'Histoire Naturelle, Department Systematics & Evolution, UMRCNRS 5202, Paris, France
³Genetics Department, University of KwaZulu-Natal, Pietermaritzburg Campus, Private Bag X01, Scottsville, South Africa
⁴School of Biological and Conservation Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, South Africa
⁵Durban Natural Science Museum, Durban, 4000, South Africa

Received 3 December 2006; accepted for publication 1 June 2007

The present study clarified the taxonomy, phylogeny and historical demography of semicommensal Nile rats (*Arvicanthis*) from the Nile Valley in Sudan. Nile rats are important crop pests and zoonotic disease reservoirs and are closely associated with agricultural settlements in the Nile Valley. Phylogenetic analysis of the complete cytochrome *b* gene (1140 bp) of 23 individuals from six localities in Sudan (from two previously recognized species, *Arvicanthis niloticus* and *Arvicanthis testicularis*), supported the existence of only a single species, *A. niloticus*, from the Nile Valley. Historical demography of the Sudanese Nile Valley population inferred from mismatch coefficients indicated that an exponential population expansion event occurred approximately 144 000–288 000 years ago, corresponding in time with early human expansion and colonization from Africa to the Middle East, Europe, and the world. The inferred high level of gene flow and large size of Sudanese Nile populations of *A. niloticus* is consistent with historically recent (300 years ago) exponential human population growth and intense agricultural activity inferred from archaeological and historical evidence. Two African *Arvicanthis* clades were well supported by the broader phylogenetic analysis: (1) *A. niloticus, Arvicanthis abyssinicus* and *Arvicanthis neumanni* and (2) *Arvicanthis rufinus* and *Arvicanthis ansorgei* from western Africa. Within the first clade, divergence between lineages of *A. niloticus s.s.* from west and north-east Africa (8.9%) suggests specific recognition, but sampling of geographically intermediate localities is required. Based on hypothesized palaeodrainage and palaeoclimatic patterns, we propose a simple model for speciation of *Arvicanthis* in Africa. © 2008 The Linnean Society of London, Biological Journal of the Linnean Society, 2008, 93, 641–655.


INTRODUCTION

Nile rats of the genus *Arvicanthis* (Rodentia: Murinae) are major agricultural pests (Poulet & Poupon, 1978) and zoonotic reservoirs (Weinbren & Mason, 1957; Trape *et al.*, 1991; Gratz, 1997; El Hassan & Zijlstra, 2001), widespread across most of...
northern, western, central, and eastern Africa. Their taxonomy in eastern and northern Africa, particularly along the Nile valley, remains uncertain, especially regarding species composition and distributional limits (Dollman, 1911; Thomas, 1916; Allen, 1939; Ellerman, 1941; Delany, 1972, 1975; Kingdon, 1974; Misonne, 1974; Yalden, Largen & Koch, 1976; Corbet & Hill, 1991; Nowak, 1991; Musser & Carleton, 1993, 2005).

Along the Nile Valley, minor differences in external morphology have led most authors to consider them as belonging to two separate species, Arvicanthis niloticus and Arvicanthis testicularis. Previous taxonomic studies based on traditional morphological (Delany, 1975; Afework et al., 1983; Musser & Carleton, 1993, 2005; Abdel Rahman Ahmed, 2005), protein electrophoretic (Philippi, 1994), geometric morphometric (Fadda & Corti, 1998; Abdel Rahman Ahmed, 2005), and chromosomal (Abdel Rahman Ahmed, 2005) evidence, are contradictory. Although Kingdon (1974) classified these two Nile forms as A. niloticus and Arvicanthis lacernatus (the holotype of the latter was subsequently shown to be a Meriones; Musser & Carleton, 2005), Delany (1975) recognized them as A. niloticus and A. testicularis, in East Africa and along the Nile Valley, differing in tail length, tooth-row length, and the presence or absence of a white belly. His key to differentiate A. niloticus and A. testicularis is provisional, however, because there are overlaps in the ranges of each of these measurements and the designation of female specimens is difficult. Afework et al. (1983) and Fadda & Corti (1998) also recognized A. niloticus and A. testicularis. Protein electrophoretic data from the Nile Valley (Philippi, 1994) and unpublished morphometric, cytogenetic and polymerase chain reaction (PCR)-random amplified polymorphic DNA (RAPD) data from Sudan (Abdel Rahman Ahmed, 2005) revealed only one form, A. niloticus, a view that is followed by Musser & Carleton (2005). Due to these conflicting data, the taxonomy of Arvicanthis from the Nile Valley (Egypt and Sudan) remains unresolved. The present study forms part of a previous study (Abdel Rahman Ahmed, 2005) comprising a multidisciplinary systematic revision of Nile Valley Arvicanthis.

The present study is the first to explore the phylogeny and population genetics of Arvicanthis from Sudan in the context of the taxonomic and geographical range of the genus, by using cytochrome b gene sequence variation. We postulate a general model for explaining the speciation and dispersal of Arvicanthis in Africa based on molecular, palaeontological, palaeodrainage, palaeoclimatic, and biogeographical evidence. Finally, we use population genetic and coalescent-based approaches to model the historical demography of Nile Valley populations of one species, A. niloticus. Because A. niloticus is semicommensal with humans and favoured by human settlement and agriculture, we postulate that inferred patterns of population expansion in this rodent should be linked with known and estimated episodes of human population expansion in the Nile Valley (Deevey, 1960; Rogers & Harpending, 1992; Kates, 1994).

MATERIAL AND METHODS

For the purpose of identification, keys and general body descriptions (Delany, 1975) were used (see Introduction); these are not reliable for identification of female specimens due to substantial character overlap. Thus, all females were labelled as Arvicanthis sp. Our molecular study included 23 specimens from Sudan (nine A. testicularis, five A. niloticus and nine unidentified females) that were collected from six localities along a north–south gradient (Fig. 1, Table 1) which covered different ecological zones in Sudan. Voucher specimens were deposited in the Durban Natural Science Museum. Additionally, 19 sequences were extracted from GenBank from Arvicanthis taxa from the whole distribution range of the genus (Table 1). Mus musculus, Aethomys chrysophilus, and Lemniscomys bellieri sequences from GenBank were used as out-groups (Table 1).

Total genomic DNA was isolated from soft organs (liver, heart, kidney, or muscle tissues preserved in 80% ethanol) by the CTAB method (Winneppeninxkw, Backeljau & De Wachter, 1993). For each individual, the complete mitochondrial cytochrome b gene (1140 bp) was amplified with Perkin-Elmer PCR reagents and conserved oligonucleotide primers, situated in the flanking tRNA genes (L14723 in tRNA-Glu and H15915 in tRNA-Thr) and internal to the cytochrome b gene (H15553 and L15146). The internal specific cytochrome b primers were designed for amplification of the complete cytochrome b gene in two separate, but overlapping, pieces of <750 nucleotide pairs. To obtain pure DNA amplified template for sequencing, each fragment of the Arvicanthis cytochrome b DNA template was amplified separately.

The amplifications were performed by primers L14723 (5′-ACC AAT GAC ATG AAA AAT CAT CGT T-3′); H 15553 (5′-TAG GCA AAT AGG AAA TAT CAT TCT GGT-3′); L15146 (5′-CAT GAG GAC AAA TAT CAT TCT GAG-3′) and H15915 (5′-TCT CCA TTT CTG GTT TAC AAG AC-3′). The PCR profile included 4 min of preheating at 94 °C and 36 cycles of 40 s at 94 °C, 45 s at 45 °C and 40 s at 72 °C. PCR cycles were followed by 10 min of extension at 72 °C. The PCR fragments were purified by using the QiAquick PCR Purification Kit (Qiagen) and spectrophotometri-
cally analysed (Beckman DU 650) to determine the concentration of purified cytochrome b gene fragments. Each cytochrome b gene fragment (100 ng DNA) was sequenced by using a direct PCR sequencing method in an automated sequencer (Beckman) following the manufacturer's protocols. Both the end primers and the internal primers were used for sequences. Sequences were aligned automatically using Sequencer 3.0 (Gene Codes Corporation).

DATA ANALYSIS
Homogeneity of base compositions was defined by Saccone et al. (1987) and bias in base composition by Irwin, Kocher & Wilson (1991). Genetic distances were computed according to Kimura (1980). Aligned sequences were analysed by maximum likelihood (ML) and maximum parsimony (MP) phylogenetic methods. ML models and parameters were determined by Modeltest, version 3.06 (Posada & Crandall, 1998). Unweighted MP analysis was performed by heuristic search (Swofford, 1993) with random addition of taxa. To assess the robustness of relationships, 1000 bootstrap replicates were performed for a Neighbour-joining analysis of ML distances obtained using the parameter estimates derived from Modeltest. The tree was rooted with the composite outgroups consisting of *M. musculus*, *A. chrysophilus*, and *L. bellieri*.

A haplotype network was calculated using TCS 1.20 (Clement, Posada & Crandall, 2000). Population genetic and historical demographic parameters of the
Table 1. List of Arvicanthis taxa and outgroups used in the present study, their geographical origin, and GenBank accession numbers for the cytochrome b gene

<table>
<thead>
<tr>
<th>Taxon</th>
<th>GenBank accession number</th>
<th>Geographic origin</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128073–EF1280877</td>
<td>Khartoum, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128062–EF128064</td>
<td>Dongola, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128065–EF128066</td>
<td>Shandi, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128067–EF128072</td>
<td>El-Sabaloglu, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128078–EF128082</td>
<td>Medani, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>EF128083–EF128084</td>
<td>El-Suki, Sudan</td>
<td>Present study</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>AF004569</td>
<td>Egypt</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>AF004572</td>
<td>Richard-Toll, Senegal</td>
<td>Ducroz et al. (1988)</td>
</tr>
<tr>
<td>Arvicanthis abyssinicus</td>
<td>AF004567</td>
<td>Menagesha, Ethiopia</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis rufinus</td>
<td>AF004583</td>
<td>Tangoumou, Benin</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>AF004571</td>
<td>Kolle, Niger</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis neumannii</td>
<td>AF004574</td>
<td>Berega, Tanzania</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis niloticus</td>
<td>AF004581</td>
<td>Burkina Faso</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis ansorgei</td>
<td>AF004580</td>
<td>Samaya, Mali</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis ansorgei</td>
<td>AF004576</td>
<td>Senegal</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis ansorgei</td>
<td>AF004577</td>
<td>Casamance, Niger</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis ansorgei</td>
<td>AF004578</td>
<td>Casamance, Niger</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis rufinus</td>
<td>AF004584</td>
<td>Koubamba, C.A.R.</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Arvicanthis dembeensi</td>
<td>AF004567</td>
<td>Koka, Ethiopia</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Lemniscomys bellieri</td>
<td>AF004586</td>
<td>Lamto, Ivory-Coast</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Aethomys chrysophilus</td>
<td>AF004587</td>
<td>Berega, Tanzania</td>
<td>Ducroz et al. (1998)</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>V00711</td>
<td>–</td>
<td>Bibb et al. (1981)</td>
</tr>
</tbody>
</table>

Sudan Nile populations were estimated using DnaSP, version 4.10 (Rozas et al., 2003) and Arlequin 3.0 (Excoffer, Laval & Schneider, 2005). Because each Sudanese individual was a unique haplotype and haplotypes showed no geographical structuring, the six Sudanese populations were considered to be a single Nile population. Following Rogers & Harpending (1992) and Russell, Medelin & McCracken (2005), we used the results from haplotype (h) and nucleotide (π) diversity values, neutrality tests (F*: Fu, 1997; Dα and F*: Fu & Li, 1993), and mismatch distribution analysis (distribution of observed pairwise nucleotide differences), to estimate whether the combined population was stationary or had undergone an historical population expansion. High h with low π, a unimodal pairwise difference distribution, significant F*, but nonsignificant Dα and F*, and a high ratio of number of variable sites (S) to average number of pairwise differences (d) (S/d), are indicators of an historical population expansion event (Russell et al., 2005). Based on the distribution of pairwise nucleotide differences, the time since expansion, tau (τ) could be calculated in mutational units. Given a generation time of four months for the species (Novak, 1991), and estimated approximate mutation rate per site per generation (μ) for rodent cytochrome b of between 8.25 × 10^{-9} and 1.65 × 10^{-8} (corresponding to nucleotide divergence rates per million years of 5% and 10%: Yoder & Yang, 2000; Gillooly et al., 2005; Veyrune et al., 2005; Jansa, Barker & Heaney, 2006), this approach allowed approximation of the absolute time of expansion using the formula τ = 2ut, where u is the mutation rate per site per generation (μ) multiplied by sequence length (1140) and t is the absolute time since expansion. Accurate determination of the rate of the cytochrome b clock in rodents is confounded by high heterogeneity (Spradling, Hafner & Demastes, 2001), and a high mutation rate (e.g. compared to primates; Yoder & Yang, 2000), leading to saturation of synonymous sites, which in turn leads to considerable underestimation of actual mutation rates. Most studies use nonsynonymous changes to calibrate molecular dates, which substantially underestimate actual mutation rates. Given estimates of
2–4% divergence rates for human mitochondrial DNA (Rogers & Harpending, 1992), which probably represent the very lower limit of rodent divergence rates, approximate values of 7–9% were obtained for species pairs within the African pygmy mice genus Nannomys (Veyrunes et al., 2005), and an empirically derived divergence estimate (probably an underestimate owing to saturation) of 4.5% from our data (see below: approximate 9% divergence obtained for the west–east split of A. niloticus, which was dated at 2 Mya from the molecular dating analysis), we felt it prudent to use divergence rates of 5% and 10% to bracket values that are likely to be realistic for Arvicanthis.

The programme Mesquite, version 1.06 (Maddison & Maddison, 2005) was used for modelling the haplotype data using a coalescence approach. Having obtained the gene flow statistic (s) (Slatkin & Maddison, 1989) for the degree of mismatch between observed gene trees and the six sampled populations from Sudan, distributions of s were obtained for 1000 simulated trees based on different conservative estimates of effective population size (105–106) and divergence estimates (branch lengths) of 105–106 generations, in order to estimate the divergence value (assuming no gene flow) that best fitted the observed s-value.

**RESULTS**

**Phylogenetic Analysis**

In the present study, guanine was under-represented and adenine over represented (A = 31.9%, C = 26.7%, G = 12.4%, T = 29.0%; G-test: P < 0.001), confirming trends already observed in other mammalian datasets (Irwin et al., 1991) and Arvicanthis from other countries, except Sudan (Ducroz et al., 1998). All taxa shared this nucleotide bias and there was no evidence of heterogeneity of base composition among taxa, thus not interfering with the present results. The number of encoded amino acids within sequences and the absence of unexpected stop codons suggest that our sequences are mitochondrial and not nuclear in origin. GenBank accession numbers for Arvicanthis Sudanese cytochrome b gene sequences are EF128062–EF128084 (Table 1) and the alignment is available from the first author.

For ML analyses, both the chi-squared and log likelihood ratio tests (Kendall & Stuart, 1979; Goldman, 1993) implemented in Modeltest, version 3.06, selected the GTR + I + Γ model for complete cytochrome b gene sequences of all Arvicanthis taxa used in the present study. Hence, all the subsequent analysis used the GTR + I + Γ model with the following parameters: base frequency parameters (ΠA = 0.3167; ΠC = 0.2863; ΠG = 0.1176 and ΠT = 0.2794), rate parameters [relative substitution of AC (rAC) = 04.1545; relative substitution of AG (rAG) = 31.939; relative substitution of AT (rAT) = 05.5922; relative substitution of CG (rCG) = 02.0272; relative substitution of CT (rCT) = 47.3538 and relative substitution of GT (rGT) = 1.0000] and the rate heterogeneity parameters [Prate = 0.5366 and shape parameter (α) = 2.9545]. The unweighted MP analysis resulted in eight most parsimonious trees (CI = 0.5677, RI = 0.6552 and tree length = 1076 steps). The trees generated by maximum Likelihood (ML, model GTR + I + Γ) and MP were identical (Fig. 2).
The monophyly of the genus *Arvicanthis* is strongly supported in both analyses. However, the relationships among the individuals of *Arvicanthis* from Sudan as well as the distinction between *A. niloticus* and *A. cf. testicularis* were unresolved by MP or ML trees (Fig. 2). The haplotype network (Fig. 3) showed that each individual was a distinct haplotype and that haplotypes were distributed randomly with respect to sample populations.

Considering *Arvicanthis* from throughout Africa, there are clearly two distinct well-supported (100% bootstrap support) clades: (1) *A. niloticus*, *A. abyssinicus* and *A. neumanni* and (2) *A. rufinus* and *A. ansorgei* (from west Africa). Clade 1 is further subdivided into four lineages: (1) *A. niloticus* from Sudan, Egypt, and Ethiopia (north-east Africa); (2) west Africa (Niger and Senegal); (3) *A. abyssinicus*; and (4) *A. neumanni*. 

The Sudanese Nile Valley population of *Arvicanthis niloticus* is characterized by very high haplotype diversity ($h = 1.0$) and very low nucleotide diversity ($P = 0.01582$), indicative of recent population expansion (Russell *et al.*, 2005). The ratio of $S/d$ (expansion coefficient) was high (5.46), within the range observed for very large populations of Mexican free-tailed bats in the USA (4.0–8.2; Russell *et al.*, 2005), further supporting the hypothesis of an expanding population. Distribution of pairwise nucleotide differences (mismatch coefficients) followed a unimodal distribution (raggedness statistic: $r = 0.0045$; $P < 0.001$, based on 1000 coalescence simulations of $r$, irrespective of level of recombination assumed; Fig. 4), which would be expected from a model of exponential population growth (a multimodal distribution would be expected from a stationary population). The mismatch distribution analysis also allowed an estimate of time since expansion ($t$) expressed in mutational units. This value was converted into absolute time using the formula $t = 2uN_{e}t$, and the generation time for *A. niloticus* (4 months; Novak, 1991), where $t$ is time in generations, $u$ is the mutation rate per site per generation for the entire cytochrome $b$ gene and $N_{e}$ is the number of generations. From the above formula, and using the two mutation rates in the Introduction ($\mu = 8.25 \times 10^{-9}$ and 1.65 a $10^{-8}$), a sequence length of 1140, and a generation time of 4 months, we obtained a $t$ value of 16.32 and estimates of 144,604–288,120 years BP since expansion.

The significance of the $F_s$ test (Fu, 1997) of neutrality ($-10.44$, $P = 0.003$), especially in the absence of the significance ($P > 0.1$) of $D^*$ ($-1.67$) and $F^*$ ($-1.9$), provides further evidence of a nonstationary population (Fu, 1997).

Coalescence simulations ($N = 1000$ simulations) of the six sampled populations in Sudan allowed estimation of the distribution of the gene flow statistic ($s$) (Slatkin & Maddison, 1989) under a wide range of modelled estimates of effective population size ($N_{e} = 10^6–10^8$) and time since divergence ($10^3–10^6$ generations). Assuming a somewhat arbitrary $N_{e}$ of $10^6$ for the *Arvicanthis* population, the model that best fits the observed $s$ of 14.573 is that with a branch length (divergence time) of approximately 1000 generations (330 years; Fig. 5).

**DISCUSSION**

**BASE COMPOSITION**

The mean base composition and compositional bias of cytochrome $b$ at the different codon positions were found to be consistent with previous reports of trends observed in mammals and other vertebrates in general (Kocher *et al.*, 1989; Irwin *et al.*, 1991; Patton, Da Silva & Malcolm, 1996; Matthee & Robinson, 1997; Halanych & Robinson, 1999) as well as in various rodents (Patton & Smith, 1992; Verheyen, Colyn & Verheyen, 1995; Ducroz, 1998; Maree, 2002; Lecompte Granjon & Denys, 2002a; Lecompte *et al.*, 2002b).

The third position bias (0.310 for *Arvicanthis* from Sudan) was similar to the bias observed in *Arvicanthis* taxa from countries other than Sudan (0.308; Ducroz *et al.*, 1998). Base composition appeared to be relatively low in G+C content (38.6%), overall for the entire gene as well as at the different codon positions (first = 45.5%; second = 37.3% and third = 32.9%). This is consistent with predictions by Brown & Simpson (1982) and Brown (1985) for mitochondrial DNA of other animal species, including *Drosophila* (20–25%) and birds (45–50%).
PHYLOGENETIC CONSIDERATIONS

In the present study, phylogenetic relationships among various lineages were relatively robust (Fig. 2). Phylogenetic analyses support the monophyly of the genus *Arvicanthis*, its association with *Lemniscomys* (100% bootstrap support with MP), and the existence of two major *Arvicanthis* clades (labelled 'C1' and 'C2' in Fig. 2): (1) *A. rufinus* and *A. ansorgei* from west Africa and (2) *A. neumanni* and *A. abyssinicus* from east Africa, together with the widespread *A. niloticus*. It is significant that the west African taxa (*ansorgei* and *rufinus*) were formerly considered part of *A. niloticus s.l.* (ANI-3 and ANI-4 cytotypes; Volobouev et al., 1988, 2002) together with ANI-2 (cytotype from central Africa; Ducroz et al., 1997) and *A. niloticus s.s.* (ANI-1a and ANI-1b cytotypes). The general tree topology is in accordance with the molecular results reported by Ducroz et al. (1998), with the new data providing additional resolution for north African *Arvicanthis* (Fig. 2).

**Figure 5.** Distributions of Slatkin & Maddison’s (1989) gene flow statistic *s* for 1000 simulated gene trees within six population trees (bush tree format). Population trees differ in their branch lengths, expressed as a function of the effective population size (*N_e*; taken to be 10^{-6}). A, branch lengths = *N_e*. B, branch lengths = 0.1*N_e*. C, branch lengths = 0.01*N_e*. D, branch lengths = 0.001*N_e*. Observed *s* = 14, indicating that the observed phylogeny is consistent with 14 gene flow events among the six populations.

**Arvicanthis from the Nile Valley**

*Arvicanthis niloticus* and *A. cf. testicularis* have been considered distinct, chiefly on the basis of traditional morphometrics (Delany, 1975) and, later, by geometric morphometrics (Fadda & Corti, 1998). However, these two putative species (*A. niloticus* and *A. testicularis*) from Sudan are unsubstantiated, with neither forming monophyletic groups (Fig. 2).

The weak interpopulation divergence (0.3–1.9%) and the unstructured haplotype network indicate the absence of any phylogeographical structure in the Sudanese populations (Figs 2, 3). Sequence data therefore compellingly indicate a single species in the Nile Valley, a result that is entirely supported by cytogenetic data (Abdel Rahman Ahmed, 2005) that showed all Sudanese specimens to have 2n = 62 and autosomal fundamental number of arms (NFA) = 62 (i.e. the ANI-1a cytotype of Volobouev et al., 1988, 2002). Allozyme electrophoresis patterns (Philippi, 1994) and PCR-RAPD fingerprints (Abdel Rahman Ahmed, 2005) further support this conclusion.

**PHYLOGENETICS OF A. NILOTICUS POPULATIONS THROUGHOUT AFRICA**

West African populations of *A. niloticus* are highly genetically distinct (8.9%) from north (Egypt), north-east (Sudan), and east (Ethiopia) African populations, whereas populations from the three latter regions show high levels of genetic similarity (Fig. 2). *Arvicanthis* specimens from Sudan are most closely related to those from Egypt, consistent with the electrophoretic study by Philippi (1994).

This divergence across a wide west–east gap is likely to be the result of genetic drift and/or adaptive
divergence operating independently on isolated or semi-isolated demes. Significantly, the two distinct lineages from north-eastern and west Africa correspond in part to chromosome differences (NFA = 62 and 64, ANI-1a and ANI-1b, respectively; Fig. 6), although both ANI-1a and ANI-1b are found in west Africa and only ANI-1a in north Africa. This cytochrome \( b \) divergence falls within the upper limit of intraspecific variation (Bradley & Baker, 2001; Baker & Bradley, 2006) but is higher than that found in many other species of African rodents (Otomys species: Maree, 2002; Mastomys: Lecompte et al., 2002a,b; Dasymys: Mullin, 2003).

Specimens from Egypt, Ethiopia and Niger had \( 2n = 62 \) and NFA = 62 (ANI-1; Volobouev et al., 1988, 2002) as did Sudanese specimens (Abdel Rahman Ahmed, 2005). The slight ANI-1a/ANI-1b distinction (one pericentric inversion of chromosome pair 30 in ANI-1a but not ANI-1b, which retains the presumed ancestral metacentric pair 30) does not merit recognition as species on cytogenetic grounds alone. As argued by King (1993) and documented by Contrafatto (1996) in the vlei rat Otomys irratus, pericentric inversions can be maintained as stable polymorphisms in natural rodent populations without resulting in reproductive isolation and, hence, speciation events.

According to Volobouev et al. (2002), ‘it appears from cytochrome \( b \) sequencing data that despite their chromosomal differentiation, West Africa ANI-1a and ANI-1b are genetically closer to each other than they are to specimens (ANI-1a) from Egypt (Ducroz et al., 1998). The minor karyotypic differentiation between the two cytotypes (ANI-1a and ANI-1b) was proved to be mostly interfertile in captivity based on west African individuals (Petter et al., 1969; Kaminski et al., 1987). Although north-east African populations have the same ANI-1a cytotype, they are morphologically and genetically (8.9%) divergent, suggesting a long period of geographical isolation, and it would be useful to conduct breeding tests between Sudanese ANI-1a and West African ANI-1a and 1b individuals.

Population genetics and historical demography of Sudanese population: of mice and men

Data from various statistical approaches support the hypothesis of a large population (sequential haplotypes separated by multiple nucleotide differences and uncorrelated with sample populations; Russell et al., 2005), which has undergone an historical expansion (high haplotype and low nucleotide diversities, high ratio of number of variable sites to number of pairwise nucleotide differences, and a unimodal distribution of pairwise nucleotide differences). Because direct gene flow estimates are unavailable, it is difficult to establish the relative contributions of incomplete lineage sorting and current gene flow to observed high levels of genetic similarity between populations.

© 2008 The Linnean Society of London, Biological Journal of the Linnean Society, 2008, 93, 641–655
Nevertheless, simulation studies under neutral coalescence suggested that, if it could be assumed that current gene flow is negligible, divergence of sampled populations would have occurred some 300 years ago. As this figure varies by several orders of magnitude from the estimated time since expansion for the Sudanese Nile population of 144 000–288 000 years ago, we must assume that gene flow between populations is not negligible but probably very high, at least partly owing to human factors and the semicommensal habits of Nile rats, as discussed below.

The observed pulse of expansion in Nile Valley *Arvicanthis* at 144 000–288 000 years BP follows closely the date of earliest divergence within the Nile Valley estimated using fossil calibrations from the PAML program (0.315 ± 0.061 Mya), and just before and within the early part of the range of recent estimates (64 000–177 000 years ago; also derived from mismatch distributions) for a postulated global expansion of early humans moving out of Africa (Rogers & Harpending, 1992). Such widescale human movements, presumably northwards along the Nile Valley into the Middle East and Eurasia, even although predating organized agriculture, would probably have created suitable habitat conditions for the geographical spread and population increase of commensal species such as *Arvicanthis*.

Nile rats benefit directly from human settlement and are known to be agricultural pests, reaching densities of more than 100 ha\(^{-1}\) in good rainfall years but crashing to much lower densities in subsequent poor years (Poulet & Poupon, 1978). Although organized agriculture has been established in the Nile Valley for at least 5000 years, it is possible that current high levels of gene flow of *A. niloticus* populations may be linked to human population explosion in the past three centuries, associated with the advent of the industrial revolution (Deevey, 1960; Ehrlich, Ehrlich & Holdren, 1977). Deevey (1960) proposed three surges in human populations over the past million years, associated with the tool making revolution, 1 Mya (leading to an increase in global human population to 5 million), the agricultural revolution 8000 years ago (leading to an increase to 15 million), and the industrial revolution starting in the 18th century (300 years ago) and continuing to this day. Based on human population extrapolations for the Nile Valley, Kates (1994) showed fairly constant human population densities (from 10 to 100 ha\(^{-1}\)) in the region from 6000 years ago until approximately 300 years ago, when populations entered an exponential phase, rapidly reaching densities of 1000 ha\(^{-1}\). The human population of Egypt increased from less than 4 million in 1798 to 4.5 million in the mid-19th century, 7 million in 1882, 11 million in 1907, and 65 million today (Bowman & Rogan, 1999). According to Bowman & Rogan (1999), 'medieval islamic geographers traced an irrigation network back to preislamic times and record that the maintenance of canals and irrigation dams was one of the primary duties of the rulers and fief-holders under the Ayyubids and Mamluks. Those canals became a feature of agricultural organization through to modern times, with periods of extension and neglect. When properly maintained, canals expanded the terrain of cultivation from the immediate banks of the Nile to the wider possible extension'.

Not only did the canals extend the area under cultivation, but also they made possible multiple cropping, thus favouring expansion of Nile rats (*Arvicanthis*). Muhammad Ali became governor of Egypt in 1805 and initiated a programme of state control over land and monopolies over its produces, which allowed him to organize a mass army to conquer other countries. He also initiated agricultural revolutions. Later, British investment in irrigation infrastructure increased agricultural productivity. A network of modern dams and canals from Aswan to the delta was built between 1898 and 1910. Indeed, agricultural land surfaces were multiplied by five at the beginning of 20th century.

These factors suggest that this recent explosive phase of human populations and agricultural activity should have been reflected in a parallel recent expansion in Nile rat populations. This does not appear to be the case based on our data, which instead showed a much older pulse of expansion, as discussed above. However, the cytochrome \(b\) gene probably evolves too slowly, and thus lacks resolution, to detect such a recent event, and future studies using much faster evolving DNA sequences would be predicted to detect such an expansion. Nevertheless, as mentioned above, population genetic parameters obtained in the present study indicate very large populations and very high levels of gene flow that are entirely consistent with hundred of years of intense agriculture and dense human populations along the Nile Valley.

**Model of speciation in *Arvicanthis***

*Arvicanthis* originated in the late Miocene and early Pliocene (Ducroz et al., 1998; Fig. 6), as a constituent of a broader pulse of faunal turnover synchronized by episodes of global climatic change. Molecular, palaeontological, palaeoclimatic, and past and present geomorphological data (e.g. hypothesized past drainage patterns and rifting events) are used here to construct an hypothetical model of speciation in *Arvicanthis niloticus s.l.* and, particularly, *A. niloticus s.s.*
BIOGEOGRAPHY, PALAEOONTOLOGY, PALAEOCLIMATES, AND GEOMORPHOLOGY

All species are habitat-specific (Vrba, 1992; DeHeinezelin et al., 1999). Nile rats (Arvicanthis) require water and grass (Kingdon, 1974). Arvicanthis niloticus s.l. occupies a wide distribution and ecological range, including deserts and savanna grasslands (Fig. 1). Arvicanthis niloticus is the most abundant rodent of the Serengeti (Senzota, 1982; Magige & Senzota, 2004). Its distribution is strongly influenced by cover of the ground layer and it prefers shrub grassland of medium height grass (Misonne & Verschuren, 1966; Senzota, 1982). Arvicanthis have been reported at elevations of up to 3700 m (in Ethiopia; Muller, 1977) and under dense herb mats (in Uganda; Kingdon, 1974).

The highest specific diversity of Arvicanthis occurs in the Ethiopian Highlands (also a centre of high species richness and endemism in other mammals; Yalden & Largen, 1992) and the lowest in north Africa along the Nile valley and in the Sahelian, Sudanian, and Guinean regions (Fig. 1).

The earliest record of Arvicanthis comes from the Lower Pliocene of Uganda (approximately 5–4 Mya, Kakara formation; Mein, 1994). A single molar, dated at 3 Mya (Omo B; Wesselman, 1984) that mostly resembles the extant species A. niloticus originated from the Rift Valley in Ethiopia. Various additional, more recent, fossil sites have yielded Arvicanthis, including the Omo C, F, G, and Asbole sites in Ethiopia (Wesselman, 1984; Geraads et al., 2004); Olduvai Beds I and II and the Natron Peninj site in Tanzania (Jaeger, 1976, 1979; Denys, 1987, 1989, 1990); East Turkana (Black & Krishtalka, 1986); and Isenya (Brugal & Denys, 1989) in Kenya. Even more recent sites, such as Bir Tarfawi in Saharan Egypt (0.2 Mya; Kowalski, 1993) and Laetoli Ngaloba Beds in Tanzania (Day & Magori, 1980), have also yielded fossils attributed to Arvicanthis. Fossil Arvicanthis have also been described from sites well outside the current range of the genus, including Plio-Pleistocene sites of Maghreb (Jaeger, 1976) and from Palestine (Tcherchov, 1968).

The following nodal absolute dates were obtained using the PAML programme (species topology was identical to that shown in Fig. 6): deepest divergence within A. niloticus from Sudan and Egypt (Nile Valley) 0.315 ± 0.061 Mya; divergence of Ethiopia from Nile Valley 0.692 ± 0.112 Mya; divergence of west and north-eastern African lineages = 1.812 ± 0.206 Mya. Inter-species divergence dates were very close to those obtained by Ducroz et al. (1998; Fig. 6): A. abyssinicus from A. niloticus 1.945 ± 0.207 Mya, divergence of A. neumanni from A. niloticus–abyssinicus 2.395 ± 0.240 Mya, divergence of A. rufinus and A. ansorgei 3.407 ± 0.337 Mya.

We hypothesize that the ancestor of Arvicanthis must have been widely distributed in northern and eastern Africa, and probably even more widely in the Middle East during the 'Zeit Wet Phase' of the late Miocene, when heavy monsoon-based rainfall from Asia affected much of northern Africa (Griffin, 2002), providing suitable mesic habitats for Arvicanthis. Four major river systems covered much of the current northern Sahara: the Eonile, the Esahhabi, the Gades, and the Chad. Following the aridification that characterized the early Pliocene, these extensive drainage areas became disrupted. For example, the Gades, which drained an area of 1.1 million km² from a larger Lake Chad to the Tunisian coast at the Gulf of Gabes, became disrupted after approximately 4.6 Myr when the rainfall in that area diminished (Griffin, 2002; Goudie, 2005). The drying up of the Sahara plausibly resulted in shrinking and fragmentation of the range(s) of ancestral Arvicanthis species, resulting in the origin (approximately 5–4 Mya) of primitive Arvicanthis clades from Central and East Africa, probably in mesic savanna grassland-woodland mosaic refugia (e.g. surrounding the hypothesized 'Megachad' Lake in west Africa).

Then followed a period of accelerated speciation of Arvicanthis in East Africa possibly associated with Rift Valley formation. Some 3.1 Mya, the eastern Rift Valley comprised a succession of small basins, 100 km in length and 1000–3000 m deep (Tiercelin, 1981; Chorowicz, 1983), that may have contained a large spectrum of habitats and niches for reproductive isolation of small populations and consequent speciation (Denys, 1999). This burst of speciation gave rise to A. neumanni, A. abyssinicus and A. niloticus (Fig. 6).

Arvicanthis niloticus originated sometime between 3 Mya (date of oldest apparent A. niloticus fossil) and the current molecular date of approximately 2 Mya. According to our molecular dates, the deepest split within A. niloticus (between west and north-east lineages at 1.81 Mya) occurred very soon after A. niloticus split from A. abyssinicus (1.94 Mya). It is feasible that the advent of more open, grassland conditions in the early Pleistocene could have accounted for the westwards spread of A. niloticus at around 1.8 Mya. According to Bobe & Behrensmeyer (2004), the expansion of grasslands in Africa was driven by global climatic change and has been linked to several key events in mammalian evolution. In East Africa, the period of domination by grasslands began 1.8 Mya (Cerling, Bowman & O'Neil, 1988). Correspondingly, the west-north-east divergence of Arvicanthis niloticus occurred 1.8 Mya.
Due to its close association with water and riverine conditions, the subsequent radiation of *A. niloticus* s.s. from east to northern Africa (0.69 Mya between Ethiopia and the Nile Valley) could plausibly be explained by modern hypotheses concerning the evolution of the Nile River (Issawi & McCauley, 1992; Goudie, 2005). During Oligocene to Miocene times, various river systems drained the eastern Sahara in both northerly (Oligocene: Gila River) and southerly (Miocene: Qena River) directions. With the Late Miocene Messinian event (6 Mya) resulting from closure of the Straits of Gibraltar and a subsequent drop of the base level of the Mediterranean by 1000 m, downcutting of the so-called Eonile River established the current north-flowing path of the upper reaches of the current Nile through Egypt. Formation of Lake Victoria 0.4 Mya, and its hypothesized mid-Pleistocene (approximately 0.9 Mya) connection with the main course of the Nile (Goudie, 2005), resulted in the formation for the first time of the Nile as a mighty river with a distant source (in east Africa) reaching the Mediterranean in Egypt. This mid-Pleistocene date for the connection of the Central-East African lakes with the Nile coincides with the molecular date interval for the northward radiation of *A. niloticus* (from Ethiopia to Sudan–Egypt).

The 5–4-Myr evolutionary history of *A. ansorgei* and *A. rufinus* in West Africa occurred during the Pliocene, after a long cooling trend of the global climate, and marks the beginning of strong fluctuations that modified the landscape. These modifications allowed species of animals and plants to exist hundreds of kilometres north of the ranges of their nearest present-day relatives (deMenocal & Bloemendal, 1996). Today, *A. ansorgei* and *A. rufinus* occupy moister savannas (of the Sudano-Guinean domain) than do the other species of the genus ( Ducroz et al., 1998; Sicard et al., 2004), analogous to the forest-savannah mosaic with acacia woodland and mesic savanna, which was postulated by Wesselman (1984) for the 3-Myr-old Omo fossil fauna. This reinforces the earlier conclusion that *Arvicanthis* was widespread in moister Late Miocene and early Pliocene savannas and became restricted to moister savanna refuges in western and eastern Africa following general Pliocene aridification.

ACKNOWLEDGEMENTS

The Service of Molecular Systematics of the Museum National d'Histoire Naturelle, Paris (SSD), and E. Pasquet, G. Lecointre, and A. Tillier, are thanked for providing laboratory facilities for the cytochrome *b* sequences. The National Research Foundation is thanked for financial support to P.J.T.

REFERENCES


Denys C. 1999. Of mice and men. Evolution in East and...


Ellerman JR. 1941. The families and genera of living rodents. 2 Family muridae. London: British Museum (Natural History).


of nucleotide sequences. *Journal of Molecular Evolution* **16**: 111–120.


Spradling TA, Hafner M, Demastes JW. 2001. Differences


