Mitochondrial DNA diversity and population structure of a forest-dependent rodent, Praomys taitae (Rodentia: Muridae) Heller 1911, in the fragmented forest patches of Taita Hills, Kenya

Nyakaana, S.; Tumusiime, C.; Oguge, N.; Siegismund, Hans Redlef; Arctander, Peter; Muwanika, V.

Published in:
South African Journal of Science

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Mitochondrial DNA diversity and population structure of a forest-dependent rodent, *Praomys taitae* (Rodentia: Muridae) Heller 1911, in the fragmented forest patches of Taita Hills, Kenya

Silvester Nyakaana\textsuperscript{a,}* , Christopher Tumusiime\textsuperscript{a}, Nicholas Oguge\textsuperscript{b}, Hans R. Siegismund\textsuperscript{d}, Peter Arctander\textsuperscript{a} and Vincent Muwanika\textsuperscript{c}

The population genetic structure of the forest-dependent rodent, *Praomys taitae*, sampled from nine indigenous forest fragments distributed over three ranges of the Taita Hills in Kenya, was determined using mitochondrial DNA (mtDNA) control region sequence variation. Thirty-three unique haplotypes were observed in a total sample of 132 individuals, whereas the number of different haplotypes per population ranged from two to 10. An overall low nucleotide diversity of 0.9% was observed in the total sample but varied widely between populations (0.2–1.3%). Significant genetic differentiation was observed in 30 of the 36 possible pair-wise comparisons between populations, while a hierarchical AMOVA revealed significant genetic subdivision between groups of populations on the three hill ranges of Dabida, Mbololo and Kyulu (*F_{ST} = 0.404, P < 0.01*), among populations on each of the hill ranges (*F_{SC} = 0.112, P < 0.01*) and among populations in the total sample (*F_{ST} = 0.471, P < 0.001*). Demographic history analyses based on pair-wise nucleotide sequence mismatch distributions revealed that all the populations were in mutation-drift disequilibrium except the populations of the Kyulu and Ronge forest fragments.

Introduction

The Taita Hills in southeast Kenya (3°25’S, 38°20’E) represent the northernmost end of the eastern arch mountain ranges and lie at an altitude of 1500–2230 m above sea level. They are composed of three distinct mountain isolates of Sagalla, Dabida and Mbololo. They have existed since the Oligocene (c. 100 million years ago) and have been forested for most of this period.\textsuperscript{1} Extensive forest loss over the last four decades, however, fueled by increased demand for agricultural land and hardwood products, has reduced the original indigenous forest cover to a patchwork of eleven forest fragments scattered across the hilltops and ridges totalling less than 400 hectares.\textsuperscript{2,3} Because of their age, isolated location, and comparatively stable climatic conditions, the Taita Hills are a global biodiversity hotspot characterized by high levels of endemism.\textsuperscript{15} Some of the endemic fauna include the Taita reed frog (*Hyperolius viridisflavus glandicola*), the Taita thrush (*Turdus helleri*), the Taita white-eye (*Zosterops polygaster silvanus*) and the Taita apalis (*Apalis thoracica fuscigularis*).\textsuperscript{4}

One of the attendant consequences of such fragmentation of terrestrial landscapes is the subdivision of a species into small and often localized populations, where mating is more likely to occur among closely related individuals, a phenomenon that culminates in reduced intra-population genetic diversity, increased genetic differentiation between populations, reduced fitness and accelerated local extinction of the populations.\textsuperscript{6–8} Whereas in highly vagile species the attendant consequences of habitat fragmentation can be mitigated by frequent migration between populations across the heterogeneous landscape,\textsuperscript{9} sedentary species with limited dispersal ability are particularly vulnerable. In this study we examined the impact of this habitat fragmentation on the distribution of genetic diversity within and between the little-studied *Praomys taitae* populations in nine forest patches in the Taita Hills at the northernmost end of the eastern arc mountain ranges. *P. taitae* is a sedentary species with an average home range of approximately 0.22 hectares.\textsuperscript{20} Although many ecological and molecular genetic analyses of fauna of the Taita Hills have been accomplished,\textsuperscript{11–16} no assessment of the genetic status of this forest-dependent murid has been carried out. It is therefore imperative that long-term effective management strategies for species in fragmented habitats be designed with the aim of mitigating the deleterious effects of inbreeding and loss of genetic diversity; and this can be achieved only if an inventory of the genetic parameters of the species in question is carried out.

We used mtDNA control region sequence variation to examine the partitioning of genetic diversity within and between *P. taitae* populations in the fragmented landscape of the Taita Hills.

Methods

Sample collection and DNA extraction

Tissue samples were obtained from toe clips of 132 samples of *Praomys taitae* trapped using small (7.5 × 9 × 23 cm) Sherman’s live traps from eight forest patches distributed on two large ridges of the Taita Hills (Dabida and Mbololo; 3°25’S, 38°20’E) and also from the Kyulu forest fragment (24°12’S, 37°52’14’E) located on a volcanic mountain ridge 40 km northwest of the Taita Hills. The study area has been described in detail.\textsuperscript{11} The spatial distribution of the forest fragments, characteristics of the study areas and number of samples obtained from each locality are summarized in Fig. 1 and Table 1, respectively. Samples were stored in cryo-vial tubes containing 25% dimethylsulphoxide (DMSO) in saturated sodium chloride\textsuperscript{8a} at ambient temperature in the field and at ~80°C in the laboratory. DNA extraction was done using the DNeasy tissue kit (Qiagen) following the manufacturer’s instructions.

DNA amplification and sequencing

We used primer pairs Ham1 (5′-CCACCATCAGCACCAACCCCAA-3′) and Ham2 (5′-CCAGAGTCTTATTAAAGGGGGA-3′)\textsuperscript{10} to amplify a 356-bp fragment of the 5′ hyper-variable segment of the mitochondrial control region. Amplifications were done in 50 µl reaction volumes containing 2–5 ng of genomic DNA, 50 mM dNTPs, 10 pmol of each primer, 1 unit of *Taq* DNA polymerase enzyme and 20 mM of 10X PCR reaction buffer containing 1.5 mM MgCl\textsubscript{2}. An Eppendorf mastercycler gradient (Eppendorf) was used with the following cycling parameters: an initial denaturation at 94°C for 5 min, 35 cycles at 94°C for 1 min, at 54°C for 2 min, at 72°C for 2 min and a final extension of 5 min at 72°C. Negative control PCR reactions containing all the reactants minus template DNA were included in all amplifications to check for possible contamination. PCR products were cleaned using the QIAquick PCR purification kit (Qiagen) and both strands of template DNA sequenced using an automated ABI 377 DNA sequencer (Applied Biosystems) using the Big Dye Terminator cycle sequencing kit version 2.0 (Applied Biosystems), following the manufacturer’s standard protocol. The light and heavy
Mitochondrial DNA phylogenetic analyses

We used the computer program package MODELTEST to select the appropriate nucleotide substitution model that best fitted our data set. Consequently, the HKY + I + G model was chosen and incorporated in subsequent analyses. The evolutionary relationships between unique mtDNA haplotypes were estimated using distance methods implemented in TREECON version 1.3b,22 taking into consideration the model of DNA evolution identified by MODELTEST. The statistical support of each node in the haplotype tree was estimated using 100 bootstrap replicates of the data set.

Genetic diversity and population structure analyses

Nucleotide diversity within populations (π),23 haplotype diversity (H),23 inter-haplotype divergences and net percentage sequence divergence between populations were estimated using the program ARLEQUIN version 3.01.24 The net inter-population distances were then used to estimate a distance-based population tree using the Neighbor program of the package PHYLIPI version 3.5c.27 A hierarchical AMOVA as implemented in ARLEQUIN version 3.0124 was used to analyse the partitioning of genetic diversity within and between populations and also between groups of populations on the three hill ranges of Dabida, Mbololo and Kyulu (see Table 1). The extent of genetic differentiation between populations was quantified using the F-statistics based on haplotype frequency and nucleotide sequence variations. The statistical significance of FST estimates was determined using 1000 permutations. We tested for correlation between within-population genetic diversity and forest fragment size using the program XLSTAT Pro version 7.5, a Microsoft Excel add-in.

Population demographic history analysis

We computed pairwise nucleotide mismatch distributions among haplotypes and tested the data against the population sudden-expansion model26 using the programs ARLEQUIN version 3.0124 and DnaSP version 4.0.27 A scenario of a recent demographic expansion or bottleneck is expected to generate a unimodal distribution of pairwise nucleotide differences between haplotypes, whereas in populations at demographic equilibrium, the pairwise mismatch distribution curve is expected to be multimodal.28,29 We validated the conformity of our data to the chosen population model using Tajima’s D and Fu’s Fs summary statistics plus the distribution of the SSD test statistic (the sum of squared differences) between the observed and the estimated mismatch distribution. We estimated the significance of the D and Fs statistics using a coalescent simulation algorithm under the hypothesis of selective neutrality and population equilibrium.22 We interpreted significantly negative D and Fs values as being indicative of recent population expansion, significantly positive values as signifying significant effects of genetic drift (which are normally accentuated under conditions of a bottleneck) and significant SSD values as evidence for significant departure from the demographic model of sudden population expansion.

Results

mtDNA control region sequence variation and evolutionary relationships between haplotypes

Thirty-three unique haplotypes defined by 22 polymorphic sites were identified among the 132 individuals sampled from the nine localities. The individual sequences have been submitted to GenBank (FJ606713–FJ606745). The frequencies and distribution of the haplotypes observed in the different localities are shown in Fig. 2. Within-population haplotype diversities were relatively high (0.3–0.9), while the number of haplotypes per population ranged from two (MW) to ten (CW) (Table 2).

Twenty-four of the haplotypes were locality specific, of which 19 were singletons. The most frequent haplotype was CW693, occurring in 22% of all the individuals sampled and distributed across five localities in the same hill range of Dabida. Only two haplotypes occurred on more than a single hill range, haplotype YA692 (shared between Dabida and Mbololo) and CW704 (occurred in all three hill ranges of Kyulu, Dabida, and Mbololo).

Table 1. Summary of sampled forest fragments, showing number of samples from each locality, approximate area of patches, elevation, and degree of disturbance.
Six out of the seven individuals from Kyulu (KY) shared a locality-specific A/G transition and T/G transversion at positions 220 and 221, respectively, while three out of the seven individuals from the same population had a unique insertion of thymine at position 242 and four individuals had a unique A/G transition at position 243 (Fig. 2). These population-specific substitutions are indicative of limited or complete absence of contemporary gene flow between this forest patch and the rest of the fragments. Likewise, a population-specific A/T transversion substitution at position 7 was found in 53% of the individuals from the Mbololo population. An overall transition: transversion ratio of approximately 4:1 was observed. The overall nucleotide diversity in the total sample was 0.9% but varied widely across populations, ranging from 0.2% in Mwachora (MW) to 1.3% in Kyulu (KY) (Table 2). A significantly positive correlation was observed between the amount of genetic diversity in the different populations and the forest patch size (as indicated in Table 1) from which they were sampled ($r = 0.84; P < 0.01$).

While the neighbor-joining tree showed no locality-specific structuring of the different haplotypes (Fig. 3), the population tree grouped the populations in three distinct clades concordant with the three hill ranges from which they were sampled (Fig. 4).

**Population structure**

Inter-population genetic divergence quantified using $F_{ST}$ statistics showed significant genetic differentiation ($P < 0.05$) in 83% of the 36 possible pair-wise population comparisons, with all comparisons of populations from different hill ranges being significant, thus suggesting limited gene flow between the contemporary populations in the different forest patches on each hill range and complete isolation of the populations on the different hill ranges (Table 3). A hierarchical analysis of molecular variance (AMOVA) also confirmed the highly significant genetic subdivision between the forest fragments on the three hill ranges of Dabida, Mbololo and Kyulu ($F_{CT} = 0.404, P < 0.01$), among populations on each of the hill ranges ($F_{SC} = 0.112, P < 0.01$) and among populations in the total sample ($F_{ST} = 0.471, P < 0.001$). The AMOVA attributed 53% of the total variation to variations between populations sampled from the same hill range.

**Demographic history of the populations**

The mismatch frequency spectra for the nine populations are shown in Fig. 5. The observed unimodal mismatch frequency distributions of all the populations sampled from the Dabida hill range (CW, ND, FU, YA, MA and MW) showed excesses of mutations that appeared in only a few individuals and a deficiency of mutations shared by many individuals—a scenario in conformity with the recent expansion model. On the other hand, the mismatch distributions of populations sampled

---

**Table 2. Summary statistics for the d-loop sequence variation in the Praomys taitae populations.**

<table>
<thead>
<tr>
<th>Population</th>
<th>CW</th>
<th>ND</th>
<th>FU</th>
<th>YA</th>
<th>MA</th>
<th>MW</th>
<th>MB</th>
<th>RO</th>
<th>KY</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>19</td>
<td>15</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>12</td>
<td>16</td>
<td>15</td>
<td>7</td>
<td>132</td>
</tr>
<tr>
<td>$H$</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>$h$</td>
<td>0.9</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.9</td>
<td>0.9</td>
<td>0.3</td>
<td>0.8</td>
<td>0.6</td>
<td>0.9</td>
</tr>
<tr>
<td>$P_i$</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>$\pi$</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
<td>0.4</td>
<td>0.6</td>
<td>0.2</td>
<td>1.0</td>
<td>0.6</td>
<td>1.3</td>
<td>0.9</td>
</tr>
</tbody>
</table>

$N$ = sample size, $H$ = number of haplotypes in each population, $h$ = haplotype diversity, $P_i$ = fraction of segregating sites, $\pi$ = percentage nucleotide diversity.

**Table 3. Population pair-wise $F_{ST}$ values based on pair-wise haplotype nucleotide differences between populations.**

<table>
<thead>
<tr>
<th></th>
<th>CW</th>
<th>ND</th>
<th>FU</th>
<th>YA</th>
<th>MA</th>
<th>MW</th>
<th>MB</th>
<th>RO</th>
<th>KY</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>F_{ST}</td>
<td>0.037**</td>
<td>0.257**</td>
<td>0.176**</td>
<td>0.121*</td>
<td>0.068**</td>
<td>0.174**</td>
<td>0.281**</td>
<td>0.284**</td>
<td>0.422**</td>
<td>0.422**</td>
</tr>
</tbody>
</table>

n.s., not significant at 5% level; * $P < 0.05$; ** $P < 0.01$
from Kyulu Hills (KY) and Mbololo Hill range (MB and RO) were bimodal and fitted poorly with their corresponding distributions expected under the recent expansion model.

Fu’s $F_s$ neutrality test, however, revealed significant signatures of recent population expansion in only two populations; Chawia ($F_s = -4.81; P < 0.01$) and Ndiwenyi ($F_s = -4.32; P < 0.01$). Tajima’s $D$ and SSD statistics were in contrast not significant (Table 4).

Conversely, the populations of Kyulu and Ronge had mismatch frequency spectra that significantly deviated from what would be expected under the sudden expansion model as indicated by the significantly positive values of the SSD test statistics (Table 4).

**Discussion**

**Population genetic structure**

An overall nucleotide diversity of 0.9% observed in this study is less than what has previously been found in other rodents such as the subterranean mole rats (*Spalax ehrenbergi*, $\pi = 3.86%$; *Heterocephalus glaber*, $\pi = 2.9%$; *Cryptomys hottentotus*, $\pi = 1.8%$), Amazonian spiny rats (*Proechimys steerei*, 7.1%; *Proechimys simonsi*, 11.8%), the Siberian lemming (*Lemmus sibiricus*, 1.8–2.8%), and the dusky-footed woodrat (*Neotoma fuscipes*, $\pi = 1.0–3.9%$). However, this value is higher than what was recently observed in the plains woodrat, *Neotoma micropus* ($\pi = 0.8%$). The low nucleotide diversity indices observed in most of the studied populations (Table 2) indicate that the populations possess a large number of closely related haplotypes, a scenario that would be expected after a recent demographic event such as a bottleneck or a population expansion. Although the spatial distribution of the haplotypes indicated that they were locality-specific (Fig. 2), they exhibited a shallow and poorly resolved phylogenetic structure (Fig. 3), confirming their very recent common ancestry.

Our data show limited female dispersal between populations as evidenced by the observed significant genetic differentiation between most of the populations. The observed limited gene flow, high haplotype diversity, and low nucleotide diversity and a significant correlation between nucleotide diversity and fragment size all suggest that genetic drift is the key factor controlling the distribution of genetic diversity in these fragmented populations. These observations contradict the previously reported panmixia and absence of differentiation in *P. taitae* populations that was based on morphometric and cytogenetic studies. Genetic studies on other co-distributed avian species present differing scenarios: strong differentiation between fragments and mutation-drift disequilibrium have been reported for the endangered Taita thrush (*Turdus helleri*), whereas habitat fragmentation was found to have no apparent effect on the dispersal of the white-starred robin (*Pogonocichla stellata*) between different

---

**Table 4. Demographic history summary statistics of the nine *Praomys taitae* populations.**

<table>
<thead>
<tr>
<th>Population</th>
<th>SSD (P-value)</th>
<th>$D$ (P-value)</th>
<th>$F_s$ (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chawia (CW)</td>
<td>0.011 (0.09)</td>
<td>0.01 (0.58)</td>
<td>$-4.81(&lt; 0.01)^*$</td>
</tr>
<tr>
<td>Ndiwenyi (ND)</td>
<td>N/A</td>
<td>-1.24 (0.11)</td>
<td>$-4.32(&lt; 0.01)^*$</td>
</tr>
<tr>
<td>Fururu (FU)</td>
<td>0.034 (0.11)</td>
<td>0.87 (0.83)</td>
<td>$-1.23 (0.21)$</td>
</tr>
<tr>
<td>Yale (YA)</td>
<td>$&lt;0.001 (0.95)$</td>
<td>$-1.14 (0.13)$</td>
<td>$-0.54 (0.36)$</td>
</tr>
<tr>
<td>Macha (MA)</td>
<td>0.007 (0.35)</td>
<td>0.185 (0.61)</td>
<td>$-1.97 (0.07)$</td>
</tr>
<tr>
<td>Mwachora (MW)</td>
<td>0.102 (0.07)</td>
<td>$-0.248 (0.34)$</td>
<td>1.38 (0.70)</td>
</tr>
<tr>
<td>Mbololo (MB)</td>
<td>0.029 (0.33)</td>
<td>1.044 (0.87)</td>
<td>$-1.02 (0.29)$</td>
</tr>
<tr>
<td>Ronge (RO)</td>
<td>0.325 ($&lt;0.01)^*$</td>
<td>0.763 (0.79)</td>
<td>3.08 (0.92)</td>
</tr>
<tr>
<td>Kyulu (KY)</td>
<td>0.166 (0.04)^*</td>
<td>0.10 (0.56)</td>
<td>1.79 (0.80)</td>
</tr>
<tr>
<td>Total sample</td>
<td>–</td>
<td>$-0.758 (0.10)$</td>
<td>$-18.60 (&gt;0.10)$</td>
</tr>
</tbody>
</table>

$^*P < 0.05$; N/A = least-squares procedure to fit model of mismatch distribution and observed distribution did not converge after 1800 steps.

SSD = sum of squared differences, $D = $ Tajima’s statistic, $F_s = $ Fu’s statistic.
The occurrence of significant genetic differentiation between fragmented populations of *P. taitae* coupled with reduced intra-population diversity has serious conservation implications for such a non-dispersing, resource-limited and forest-dependent species. It is a well-established ecological fact that altered edge effects due to forest habitat fragmentation favour invasions of fragmented patches by highly competitive grassland generalist species with no specialized requirement for habitat use. Therefore, any intensive inter-specific competitive interaction between the invasive generalists and the forest-dependent *P. taitae* is bound to impact negatively on the survival of the latter, and may even act as a major driving force for local extinctions of the species in some of the forest patches.

In the event that some of the populations become locally extinct, rapid recolonization from nearby populations is unlikely given the limited dispersal ability of the species plus the human-induced loss of connectivity between the different forest patches.

### Population demography

Tajima’s and Fu’s tests for neutrality of mutations used in this study have been reported to be unable to disentangle the effects of true selective departures from neutrality (such as those caused by purifying selection and genetic hitchhiking) from those due to demographic processes such as population expansion on mismatch distribution profiles. For example, both statistics can be significantly negative due to purifying selection, genetic hitchhiking or population expansion while both of them can be significantly positive under scenarios of population bottlenecks and balancing selection at linked loci. However, because the data used in this study are derived from the selectively neutral non-coding region of the mitochondrial genome (control region), it is highly unlikely that the observed significant deviations from neutrality are due to selection. The high haplotype diversity, low nucleotide diversity, unimodal mismatch distributions, and significantly negative values of Fu’s $F_s$ statistic (see Table 2, Fig. 5 and Table 4) all support a recent population expansion from a smaller founder population as the most plausible explanation for the observed significant deviations from neutrality in the Chawia and Ndiwenyi populations. The observation that population expansion in the Chawia and Ndiwenyi populations was detected by Fu’s $F_s$ but not Tajima’s $D$ statistic attests to the fact that Fu’s $F_s$ statistic is a more powerful tool for detecting deviations from neutrality when testing for population expansion and genetic hitchhiking in populations.

This research was funded by the Danish International Development Agency (DANIDA) under the Wildlife Genetics Project, a collaborative project between the University of Copenhagen (Denmark) and Makerere University (Uganda). DNA samples were obtained through research collaboration with the Zoology Department, Kenyatta University, Kenya.

Received 21 July 2007. Accepted 14 October 2008.

---

skewed sex ratios in the critically endangered Taita thrush as revealed by CHD genes. 


