Conservation genetics of the water mouse, *Xeromys myoides* Thomas, 1889

David Benfer[A, E], Andrew M. Baker[A], Tina Ball[B], Ian Gynther[C], Heather Janetzki[D] and Susan Fuller[A]

[A] Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane, Qld, 4001, Australia

[B] Department of National Parks, Recreation, Sport and Racing, PO Box 623, Mackay, Qld, 4740, Australia

[C] Threatened Species Unit, Department of Environment and Heritage Protection, PO Box 64, Bellbowrie, Qld, 4070, Australia

[D] Biodiversity Program, Queensland Museum, PO Box 3300, South Brisbane, Qld, 4101, Australia

[E] Corresponding author. email: d.benfer@qut.edu.au

Abstract

The water mouse, *Xeromys myoides*, is currently recognised as a vulnerable species in Australia, inhabiting a small number of distinct and isolated coastal regions of Queensland and the Northern Territory. An examination of the evolutionary history and contemporary influences shaping the genetic structure of this species is required to make informed conservation management decisions. Here, we report the first analysis undertaken on the phylogeography and population genetics of the water mouse across its mainland Australian distribution. Genetic diversity was assessed at two mitochondrial DNA (Cytochrome b 1000bp, D-loop 400bp) and eight microsatellite DNA loci. Very low genetic diversity was found indicating water mice underwent a recent expansion throughout their Australian range and constitute a single evolutionarily significant unit. Microsatellite analyses revealed the highest genetic diversity was found in the Mackay region of central Queensland; population sub-structure was also identified, suggesting that local populations may be isolated in this region. Conversely, the Coomera region of south-east Queensland revealed very low genetic diversity and the population in this region has experienced a significant genetic bottleneck. These results have significant implications for future management, particularly in terms of augmenting populations through translocations or re-introducing water mice in areas where they have gone extinct.

Introduction

Genetic factors may play a critical role in species extinction (Spielman *et al.* 2004; Frankham 2005; O'Grady *et al.* 2006). The maintenance of genetic diversity enables a species to adapt to changing environments and avoid the negative effects of inbreeding depression through accumulation of deleterious mutations and loss of heterozygote advantage (Spielman *et al.* 2004; Frankham 2005). These issues can be rapidly mitigated through the restoration of gene flow (Lacy 1987; Ingvarsson 2001), but if populations have been separated for long periods of time, they may be highly divergent and adapted to local conditions (Moritz 1994). Restoration of gene flow may therefore be inappropriate and may lead to other problems such as outbreeding depression (Frankham *et al.* 2011) and loss of unique evolutionary lineages (Moritz 2002). As such, an understanding of the natural historical processes and recent anthropogenic influences shaping the genetic structure of threatened species is required to make informed conservation management decisions and efficiently allocate scarce conservation resources.

The water mouse, *Xeromys myoides* Thomas, 1889, is listed in Australia as a vulnerable species under both Queensland and Commonwealth legislation. This predominantly ground dwelling (Gynther and Janetzki 2008) and nocturnal (Van Dyck 1996) rodent occupies mangrove forests, freshwater lagoons, swamps and sedged coastal lakes (Magnusson *et al.* 1976; Van Dyck 1996, Ball 2004; Gynther 2011). It is the only member of the genus *Xeromys* and its closest relative in Australia is the water rat, *Hydromys chrysogaster* (Musser and Carleton 2005). However, morphological and molecular evidence suggests that it is more closely related to the 'moss mice' (*Leptomys* and *Pseudohydromys*) of New Guinea (Rowe *et al.* 2008; Helgen and Helgen 2009). Across the nine genera in the tribe Hydromini, the water rat and water mouse are the only species that occur outside New Guinea. Rather, both species are found in Australia and New Guinea; although to date, only one record exists of the water mouse in New Guinea from a freshwater wetland on the Bensbach River floodplain (Hitchcock 1998). It is therefore thought that both species probably originated in New Guinea (Breed and Aplin 2008; Rowe *et al.* 2008)

and radiated to Australia during the Pleistocene at periods of reduced sea level (Aplin 2006). The vegetation on the Torres land-bridge included extensive low-lying swamps (Torgersen *et al.* 1988), which may have provided suitable habitat for rodents dependent on water. However, the history of rodent exchange between New Guinea and Australia has been complex (see Rowe *et al.* 2008).

The Australian mainland distribution of the water mouse includes coastal areas of central and south-east Queensland from Proserpine south to the Coomera River (Van Dyck and Gynther 2003; Ball 2004), with confirmed records lacking between Cape Palmerston, south of Mackay, and Port Curtis near Gladstone (Ball 2004; Van Dyck and Gynther 2012; QGC 2013). The species is also known offshore from Curtis Island (Brett Taylor, personal communication, 2011), Fraser Island (DERM 2010), Bribie Island (Van Dyck and Gynther 2003; Gynther 2011), North Stradbroke Island (Van Dyck *et al.* 1979; Van Dyck and Durbidge 1992; Van Dyck 1996) and South Stradbroke Island (Van Dyck and Gynther 2003). It has been recorded infrequently at widely separated sites in the Northern Territory including the Glyde, Goyder and Tomkinson Rivers in Arnhem Land, the South Alligator and Daly Rivers and Melville Island (Redhead and McKean 1975; Magnusson *et al.* 1976; Woinarski *et al.* 2000). It is unknown whether the water mouse was once widely distributed along the northern and eastern coastline of Australia. It also remains unclear whether its current disjunct distribution is historical and related to habitat requirements or climatic fluctuations (Russell and Hale 2009), more recent anthropogenic impacts (Gynther and Janetzki 2008; Gynther 2011), or simply a lack of trapping in difficult to access, sometimes crocodile-infested, mangrove habitats (Van Dyck 1996). Nevertheless, the apparent absence of water mice from sites which appear suitable has led to the conclusion that habitat requirements for the species may be highly specific (Russell and Hale 2009).

To date, no published study has examined the distribution of genetic variation amongst the regions occupied by the water mouse and broader phylogenetic analyses have only included a single individual (Rowe *et al.* 2008). The levels of genetic exchange and hence the connectivity

between populations is therefore unknown. Factors influencing the current, potentially fragmented, distribution of the water mouse, such as limited dispersal and barriers to gene flow, genetic drift and inbreeding have also not been considered. Given these considerable knowledge gaps and the current conservation significance of this species, the aim of this study was to examine the distribution of nuclear and mitochondrial DNA variation throughout the geographic range of the water mouse to infer the evolutionary history and contemporary influences on gene flow and provide a better understanding of the long term viability of water mouse populations in Australia.

Materials and Methods

DNA extraction and sequencing

DNA was extracted from 47 *Xeromys myoides* tissue samples (tail or ear clips) from Queensland (Coomera region, Mackay region, Donnybrook and Gladstone region) and two samples (liver) from the Northern Territory (Goyder River and Arafura Swamp) in Australia (see Supplementary Material Table S1 for specimen details). Large tissue samples were extracted using a salt extraction protocol (Miller *et al.* 1988), while small tissue samples were extracted using an Isolate Genomic DNA mini kit (Bioline, Alexandria, Australia) following the manufacturer's protocol.

Approximately 1000bp of the mitochondrial cytochrome b (CytB) gene was amplified for 28 individuals spanning all five localities and 400bp of the D-loop gene was amplified for 24 individuals (see Table 1 for primer details). One nuclear marker was also sequenced: 300bp of the Zona Pellucida 3 glycoprotein (ZP3) for eight individuals from across the sampled range. Polymerase chain reactions (PCR) were performed in a Mastercycler® epgradient S (Eppendorf, Hamburg, Germany). The total volume of the sequencing reaction was 25µl and contained 1µl of both the forward and reverse primers (Geneworks, Adelaide, Australia) at a concentration of 10µM, 5µl of Bioline 5x MyTaq Red reaction buffer, 0.25µl of Bioline MyTaq HS DNA polymerase, between 0.5 and 2µl of the previously extracted DNA and the remaining volume in dH_2O. The following PCR cycling protocol was used: 94°C for 3 min, 36 cycles of 94°C for 15 sec, 50°C for CytB or 55°C for D-loop and ZP3 for 15 sec and 72°C for 30 sec, final extension at 72°C for 5 min and then hold at 15°C. To check for presence and quality of amplification, PCR products were then electrophoresed on a 1.5% TBE agarose gel stained with Gel Red.

PCR products were purified using an Isolate PCR and Gel Kit (Bioline, Alexandria, Australia) following manufacturer's guidelines and eluted into 15µl of elution buffer. Purified PCR product was then amplified in a sequencing reaction with a total volume of 20µl. This

contained 1 to 10µl of PCR product depending on its estimated concentration, 1µl of the forward primer at a concentration of 3.2µM, 3µl of 5x Big Dye sequencing buffer, 0.5µl of Big Dye, with remaining volume composed of dH_2O. The sequencing reaction protocol was the same for all regions amplified and consisted of 94°C for 5 min, 30 cycles of 94°C for 10 sec, 50°C for 5 sec and 60°C for 4 min and a final extension at 15°C for 10 mins. A standard EDTA/ethanol precipitation protocol was used to prepare DNA for sequencing. Samples were sequenced in the Queensland University of Technology's (QUT) Molecular Genetics Research Facility on an ABI 3500 sequencing platform.

Phylogenetic and phylogeographic analyses

For all analyses, individuals were grouped into broader sampling localities. The 'Coomera region' consisted of samples from Coomera and Jacobs Well and the 'Mackay region' consisted of samples from Sarina, Cape Hillsborough, Bucasia and Cape Palmerston National Park. Gladstone and Agnes Water individuals were grouped as 'Gladstone region' for microsatellite analyses (Table S1 in the Supplementary Material); however, only the Agnes Water individual was sequenced for the mitochondrial genes. Chromatograms and sequences for the three genes were aligned by eye using BioEdit Version 7.0.8 (Hall 1999). Further analyses of the ZP3 gene sequences were not conducted as identical sequences were obtained for all eight individuals spanning the five sampling localities. MEGA version 5.1 (Tamura *et al.* 2011) was used to obtain uncorrected pairwise p-distances for the CytB and D-loop sequences. Networks were generated for both CytB and D-loop *Xeromys myoides* sequences using TCS version 1.21 (Clement *et al.* 2000). A statistical parsimony approach as described by Templeton (1992) was used and a 95% connection limit was implemented. The optimal model of molecular evolution for phylogenetic analysis of the CytB data was found using jModelTest2 version 2.1.4 (Guindon and Gascuel 2003; Darriba *et al.* 2012). MrBayes version 3.2 (Ronquist *et al.* 2012) was employed for Bayesian inference of phylogenies using the Metropolis-coupled Markov Chain Monte Carlo (MCMC) method (Geyer 1991). The program was run under the General Time

Reversible model gamma distributed with Invariant Sites (+I+ Γ) for 500000 generations. Model parameters were α = 39.02, proportion of invariant sites = 0.544, substitution rates *R*(a) = 0.046, *R*(b) = 0.092, *R*(c) = 0.0667, *R*(d) = 0.139, *R*(e) = 0.772 and *R*(f) = 0.009. Empirical base frequencies were A = 0.31, C = 0.31, G = 0.12 and T = 0.26. The final CytB tree was then visualised and edited in FigTree version 1.4 (Rambaut 2012). D-Loop sequences were excluded from the broader phylogenetic analyses as sequences for out-group taxa were unavailable. Specimen details and GenBank accession numbers can be found in Supplementary Material Table S1.

Microsatellite amplification

Eight polymorphic microsatellite loci were developed and optimised (Table 2). Next generation sequencing was conducted on one individual (sample IG016 - See Supplementary Material Table S1.) using an Ion Torrent Personal Genome Analyser (Life Technologies, Germany). An Ion Torrent library was produced following the Ion Xpress Plus gDNA Fragment Library Preparation protocol. The library was then purified and prepared for sequencing. The sample was loaded onto the Ion 314 chip and sequenced on the Personal Genome Analyser. The sequence data obtained from the torrent run was checked for quality by excluding low quality sequences and trimming ambiguous bases. These data were then used to locate a number of candidate regions suitable for microsatellite analysis. The program Msat commander version 1.0.8 beta (Faircloth 2008) was used to identify loci consisting of tetra-nucleotide repeats and amplifying at an annealing temperature of 60^{o}C using the default settings. Fluorescently labelled reverse primers were obtained from Applied Biosystems, while forward primers were produced by Geneworks (Adelaide). Eight primer pairs matching these criteria were initially trialled for successful amplification and polymorphism for 25 *Xeromys myoides* individuals from across all five sampling regions. Of the initial eight loci, four were abandoned due to poor

amplification or monomorphism. A further eight primer pairs were selected of which four were again excluded. This resulted in the final eight polymorphic loci employed in the present study. These loci were then amplified for the remaining *Xeromys myoides* individuals.

A Qiagen multiplex kit (Qiagen, Dusseldorf, Germany) was used for microsatellite amplification. This allowed several loci to be amplified in a single reaction. A primer mix was prepared containing the forward and fluorescently labelled reverse primers, diluted in TE buffer, so that each primer was at a final concentration of 2µM. Microsatellite fragments were amplified in a reaction containing 6.25µl of 2x Qiagen multiplex master mix, 2.5µl of Q Solution, 1.25µl of primer mix, 4µl of Rnase free H_2O and 1µl of the previously extracted DNA. The following PCR cycler protocol was used: 95°C for 15 min, 30 cycles of 94°C for 30 sec, 62°C 90 sec and 72°C for 1 min, followed by 60°C for 30 min and ending with a hold at 15°C for 15 min. PCR products were run on a 1.5% TBE agarose gel stained with Gel Red using an electrophoresis rig to check for successful amplification. PCR products were prepared for fragment analysis in a dilution containing 3µl of PCR product in 12µl of HiDi (Applied Biosystems) and 1µl of GeneScan™ 600Liz (Applied Biosystems) sizing standard was added. Samples were then run on an ABI 3500 sequencing platform in the QUT Molecular Genetics Research Facility.

Population genetics analyses

A total of 49 *Xeromys myoides* individuals were included in the population genetics analyses. Microsatellite chromatograms were viewed on GeneMapper® (Applied Biosystems) software and allele sizes were scored manually. Tests for Linkage Disequilibrium (LD) and departure from Hardy-Weinberg Equilibrium (HWE) were implemented in GENEPOP version 4.2 (Raymond and Rousset 1995, Rousset 2008) using the MCMC method (defaults settings of: 1000 dememorizations, 100 batches, 1000 iterations). A comparison across populations revealed that locus pairs Xmyo1a and Xmyo10 ($p<0.021$), and Xmyo1a and Xmyo14 ($p<0.019$)

had a significant signature of linkage disequilibrium (LD). Xmyo1a was therefore excluded from further analyses. Fstat version 2.9.3 (Goudet 1995) was used to calculate F_{is} for Coomera and Mackay regions only as other regions had insufficient sample size. Private alleles (alleles unique to a particular region) were identified by eye from the raw data and Arlequin version 3.11 (Excoffier *et al.* 2005) was used to produce estimates of expected and observed heterozygosity and pairwise F_{st} between Coomera and Mackay regions.

BOTTLENECK version 1.2.02 (Cornuet and Luikart 1996) was used to establish if populations had experienced a recent decline in population size. In this instance, 'recent' was defined as $2N_e$-$4N_e$ generations depending on the mutation rate of the loci being studied and the severity of the bottleneck (Cornuet and Luikart 1996). The program was run for 1000 iterations under the stepwise mutation model and a one tailed Wilcoxin test was performed as recommended for microsatellite data (Piry *et al.* 1999). This test was only run on Mackay and Coomera area samples due to insufficient sample sizes in other regions.

A Bayesian clustering approach implemented in STRUCTURE version 2.3.2 (Pritchard *et al.* 2000) was used to estimate the number of populations (*K*) in a sample and to assign individuals to one or more of these populations (*k*). Ten runs of $K = 1$ to 10 were performed at 100000 MCMC repetitions and 20000 burn-in period using no prior location information, independent allele frequencies and a model of admixture. The posterior probability was then calculated for each value of *K* using the estimated log-likelihood to choose the optimal number of populations. The number of populations (*K*) was estimated from the point where the mean posterior probability Ln *Pr* (X|K) reached a plateau. The Evanno method (Evanno *et al.* 2005) was also employed to confirm this via the website 'STRUCTURE HARVESTER' (Earl 2012). The program CLUMPP (Jakobsson and Rosenberg 2007) was used to produce a best match amalgamation of the ten iterations of the STRUCTURE analysis using the default settings. Final plots from these data were produced using the program Distruct (Rosenberg 2004). A

Factorial Correspondence Analysis (FCA) was also conducted using GENETIX version 4.05 (Belkhir *et al.* 1996).

Results

Phylogenetic and phylogeographic analyses

Eight individuals spanning the range from the Northern Territory to south-east Queensland were sequenced for the ZP3 nuclear gene. However, these sequences were identical and, as they spanned the species' entire known distribution, it was determined that further investigation of this gene would not be informative.

A total of 24 individuals from throughout the range of the water mouse were sequenced for mitochondrial DNA D-loop and seven haplotypes were found (GenBank accession numbers KM582160 to KM582166). Uncorrected p-distances between *Xeromys* haplotypes ranged from 0.3% to 1.4% (Table S2 in the Supplementary Material).

The CytB dataset consisted of 28 *Xeromys myoides* samples from across the geographic range. A total of six CytB haplotypes were found among *Xeromys* (GenBank accession numbers KM603491 to603496). CytB p-distances between haplotypes ranged from 0.1% to 0.6% among *Xeromys* and 0.4% to 1.7% among *Hydromys* (Table S2 in the Supplementary Material). The Bayesian inference phylogeny generated using the CytB dataset revealed that *Xeromys* samples formed a well-supported monophyletic clade (Fig. 1), with *Pseudohydromys ellermani* positioned as sister-taxon. All Australian *Hydromys* also formed a well-supported clade to the exclusion of the New Guinean *Hydromys*. No insertions or deletions were identified in any of the sequence data.

The networks produced from the mitochondrial DNA (CytB, D-loop) sequence data (Fig. 2) both consist of a central haplotype shared between Donnybrook and Mackay area individuals. In the CytB network, the majority of Mackay and Coomera area individuals fall in this central haplotype. However, the D-loop network shows additional structure with unique haplotypes present in the Coomera and Mackay regions. Northern Territory individuals occurred as tip

haplotypes in both networks, with the Goyder River individual being the most divergent. The key difference between the two networks is the position of the Agnes Water sample, which occurs centrally in the CytB network, but as a tip haplotype in the D-loop. Most strikingly, the geographically proximate Goyder River and Arafura Swamp individuals from the Northern Territory are relatively widely separated (3-4 base pairs) in the network.

Population genetic analyses

Allele frequencies for microsatellite data are given in Supplementary Material Table S3. Descriptive statistics for the microsatellite data are given in Table 3. The Mackay area had the highest mean number of alleles. All sites except the Gladstone region had at least one private allele (alleles unique to a specific region), but they were most common in the Mackay area. The Coomera region had only two private alleles across all eight loci. Most alleles were distributed across all regions sampled. Allele sizes often varied by only two or three tetra-repeats and were rarely different by more than one or two repeats within a single region. The mean F_{is} value for the Coomera area indicated a heterozygote excess and for the Mackay region a heterozygote deficiency. Mean observed heterozygosity (H_o) was highest in Mackay. Mean expected heterozygosity (H_e) was highest in Coomera.

Across populations only the Mackay region showed significant departure from Hardy-Weinberg equilibrium (HWE) at loci Xmyo8 ($p<0.0002$), Xmyo10 ($p<0.002$) and Xmyo14 ($p<0.0008$). A pairwise estimate of F_{st} revealed significant differentiation between Coomera and Mackay areas ($F_{st} = 0.39$, $p<0.05$). The BOTTLENECK analysis revealed a significant difference ($p<0.031$) from expected heterozygosity relative to the number of alleles found in the Coomera region, indicating that the population may have been subject to a recent bottleneck, while the Mackay area showed no signature ($p>0.28$).

The STRUCTURE analysis identified two clusters ($K = 2$) across the mainland Australian range of the water mouse (Fig. 3). This was supported by both the Evanno method (Evanno *et al.*

2005) and a plateau in the rate of change in mean posterior probability Ln *Pr* (X|K). A STRUCTURE bar plot showing three clusters was also produced as it provided insight into sub-structuring that was not evident in the two cluster plot (Fig. 3). In the $K = 2$ plot, the Coomera region appeared to form a distinct cluster and the Mackay area and Northern Territory samples formed another. Both Donnybrook and Gladstone region individuals consisted of varying degrees of admixture between the two clusters. The $K = 3$ plot produced a similar result. The Coomera area still formed a distinct cluster; although, admixture was prominent in the other regions.

The FCA plot (Fig. 4) revealed clustering based on sampling region with minimal overlap between groups. Factor one described 48.5%, Factor two 22.9%, Factor three 22.2 % and Factor four (not plotted) explained 19.4% of the variation between all samples.

Discussion

Phylogenetic analyses

We found very low levels of mitochondrial DNA divergence over the known Australian range of the water mouse. The CytB sequences revealed a small amount of divergence between Queensland and the Northern Territory, but shared haplotypes were present across the entire sampled Queensland range. Phylogenetic analyses indicated that *Xeromys* was monophyletic (Fig. 1). *Xeromys* and moss mice (*Leptomys* and *Pseudohydromys*) were positioned as sister-taxa to *Hydromys* and *Parahydromys*. Our results differ to that reported by Rowe *et al.* (2008), who found that *Xeromys, Hydromys* and *Pseudohydromys* formed a clade to the exclusion of *Leptomys*. Their analysis consisted of only a single water mouse and two *Hydromys* individuals, but they included a number of nuclear genes which may have allowed better resolution of deeper level relationships. The single nuclear gene sequenced in the current study (ZP3) showed no evidence of differentiation within *Xeromys*.

Relative levels of divergence and monophyly across the sampled range strongly support the current designation of *Xeromys myoides* as a single species (e.g. maximum of 0.5% CytB divergence in *Xeromys* relative to ~6% between different *Melomys* species [Bryant *et al.* 2011] and ~7-10% between *Pseudomys* species [Rowe *et al.* 2011]). Preliminary morphological examinations also suggest that water mice from Queensland and Northern Territory cannot be distinguished (Steve Van Dyck, unpublished data). However, broader sampling including island representatives may provide further insight.

Phylogeographic relationships and historical biogeography

Given the limited divergence among water mice assayed in the present study and their apparent sister relationship to *Pseudohydromys* of New Guinea (Rowe *et al.* 2008), water mice probably diverged from their moss mice ancestors outside of Australia. Aplin (2006) hypothesised that

the water mouse probably arrived in Australia from New Guinea during the Pleistocene (~2.6myr – 12kyr; Gibbard *et al.* 2010) during periods of reduced sea level. To test this hypothesis, the relationship between Australian and New Guinean individuals would need to be established, but there is only one museum record (Hitchcock 1998) and no tissue currently available.

The limited divergence found in the current study was unexpected considering the apparently disjunct (and broad) distribution of *Xeromys* in Australia. This suggests water mice have recently expanded throughout their range. Populations have either not been isolated long enough for divergence to arise (Avise *et al.* 1987) or are more connected than current trapping records would suggest. *Hydromys chrysogaster* also exhibited low levels of divergence in the current study over large geographic distances (0.5% at CytB from Tully in north Queensland to Torrens in South Australia), providing support for a probable recent range expansion throughout Australia.

The microsatellite data also provide further support for a recent range expansion of the water mouse throughout Australia. Microsatellites are rapidly evolving, non-coding repeat sequences. Under the Stepwise Mutation Model (SMM) microsatellites evolve by either dropping or gaining a repeat number from the sequence. Alleles of a similar size would therefore be expected to be recently diverged (Selkoe and Toonen 2006). The microsatellites examined here exhibited shared alleles across the entire distribution and repeat numbers differed only marginally (generally only one repeat), both within and between regions. The microsatellite mutation process can be more complex than explained by the SMM alone and size homoplasy (the same sized allele arising independently between regions) could explain the shallow divergence found in this study (Selkoe and Toonen 2006), particularly in the case of the shared alleles found between the Northern Territory and Queensland. However, given that the patterns seen in the microsatellites are similar to those found in the mitochondrial DNA analysis, it seems unlikely that size homoplasy has had a large impact in the present study.

Coalescence principles (Kingman 1982) indicate that the centrally located Queensland haplotypes found in both networks (Fig. 2) are ancestral to those on the tips. This suggests water mice most likely initially colonised Australia via Queensland. Although, given their current absence from trapping records in north Queensland, it remains unclear whether this was across the Torres land bridge or a rare rafting or translocation event into central-east Queensland. Both networks suggested that the Northern Territory Arafura Swamp individual has diverged from an ancestral Queensland haplotype (Fig. 2). However, as only two samples from the Northern Territory were available, more detailed sampling is required. This may yield further haplotypes, providing a link to the Northern Territory's shared origins with Queensland and explaining the relationship between Arafura Swamp and Agnes Water haplotypes.

Under a scenario of a simple radiation from a single point of entry, geographically proximal locations are expected to be most similar to each other (Wright 1943; Slatkin 1993). However, the mitochondrial DNA haplotypes of the two Northern Territory individuals were on opposing tips of the networks (Fig. 2), and their origins most likely lie in two independent colonisation events.

These conclusions are dependent on the presence of water mice outside of their current known range. Nevertheless, this appears to be the most parsimonious explanation for the distribution of genetic variation in the current study. Without broader sampling of north Queensland, the Northern Territory and New Guinea it is difficult to draw further conclusions about the colonisation patterns of the water mouse in Australia.

Current population structure and implications for management

While the exact mode of colonisation and the evolutionary relationships among sub-populations can only be hypothesised, our data suggest that populations are genetically similar and recently diverged across the entire Australian range. This has significant implications for the way this species is managed in the future. Water mice consist of a single Evolutionarily Significant Unit

(ESU) as defined by Moritz (1994). They are monophyletic across geographic regions and do not appear to be divergent at nuclear loci (ZP3 sequence identical from Queensland to Northern Territory; minimal divergence in nuclear microsatellites). The presence of a limited amount of structuring in the D-loop and microsatellites suggests that populations may have very recently started to diverge (Avise *et al.* 1987). Each geographic region can be differentiated in the FCA plot, with minimal overlapping individuals between regions (Fig. 4), and unique D-loop haplotypes are present in most regions. This may warrant designation of distinct Management Units (MUs) (Moritz 1994) in Queensland, however, further study is required to determine if populations are demographically distinct or if other distinct MUs exist.

Genetic diversity was much greater in the Mackay area than the Coomera region. While capture records indicate that *Xeromys* are relatively easily trapped in Mackay (Ball 2004), the main Coomera population sampled here has since gone extinct (Van Dyck *et al.* 2006). The population in the Mackay region showed significant departure from Hardy-Weinberg Equilibrium at several loci and may be evidence of a Wahlund effect and the STRUCTURE (K=3) plot (Fig. 3) suggests that the most northerly sampling localities may represent distinct sub-populations potentially isolated from other Mackay sampling localities. If these sub-populations are small and isolated, they could be subject to rapid loss of genetic variability through increased genetic drift and inbreeding (Frankham 2005). The largely positive F_{is} values indicate an excess of homozygotes, which can occur as a result of inbreeding within sub-populations (Wright 1949). Considering the frequency with which mangrove habitats are now altered or removed for development (Valiela *et al.* 2001), there is a real risk of populations becoming more isolated throughout much of the water mouse's current range.

The Coomera population showed the least structure and variability of all regions sampled, but sampling was limited to a smaller geographic area than in the Mackay region (three sample sites near Coomera separated by approximately 10km compared with five sample sites around Mackay over about 80km). Analyses suggest that the low genetic variability may have resulted

from a genetic bottleneck and this is consistent with what is known about the Coomera area population, as these samples were obtained during a recent population decline (Steve Van Dyck, personal communication, 2012). Microsatellite variability in the Coomera population was lower than in the Mackay region from a similar sample size; only two alleles were present at each locus and alleles were fixed at two loci. A large range of threats to this species' persistence have been identified (Gynther and Janetzki 2008; DERM 2010; Gynther 2011; Van Dyck and Gynther 2012) which may have contributed to the decline of the population in the Coomera area. Pressure from loss of habitat and increased predation by feral predators due to proximity to large urban centres (Gynther and Janetzki 2008) may have resulted in an initial decline, accelerating the process of genetic drift and reducing fitness of the population as whole. Only small amounts of migration would be required to mitigate these negative effects (Ingvarsson 2001). However, isolation by urbanisation associated with the cities of Brisbane and the Gold Coast, together with extensive modification of mangrove habitat in this region, probably completely halted any migration into this population from surrounding areas. Coomera is considered to be the southernmost limit within the species' mainland distribution (Van Dyck and Gynther 2003, 2012) and this could have also impacted on migration rates over an historical timeframe, particularly if the area was initially colonised by only a few individuals. The presence of a unique D-loop haplotype provides some support for this conclusion and indicates populations may have started to diverge prior to anthropogenic impacts.

Despite a sample of only three individuals, the Donnybrook population appears to be more variable than the one in the Coomera region. All three individuals were from the same small, relatively isolated patch of mangroves yet still possessed greater allelic diversity than all of the Coomera area samples combined. Gynther (2011) proposes that Pumicestone Passage is a stronghold for the species in south-east Queensland and the limited data available in the present study suggest that the Donnybrook population is not suffering from the same lack of genetic variability found in the Coomera region. If remnant populations are large and sufficient habitat

exists in Donnybrook, this area may provide a valuable source of individuals for re-introduction. Considering that divergence between water mice populations is so low, the use of translocations could be implemented safely without risk of loss of unique lineages or outbreeding depression (Frankham *et al.* 2011). Particularly in localised areas, translocation could certainly be considered as an option to augment existing populations, and to reduce the chance of inbreeding depression in small isolated sub-populations. However, given scarce conservation resources, the question is whether this is the most important issue facing the species' immediate persistence. In the short term, greater benefit may come from concentrating on more immediate threats to water mouse persistence such as clearing of mangrove habitat and feral predators (Gynther 2011).

Future directions

Our results indicate that Australian water mice constitute only a single species and are recent arrivals on an evolutionary timescale. Inclusion of New Guinean individuals (if they can be located) would prove invaluable in further resolving the phylogeography of the species. Furthermore, broad sampling is required to define the boundaries of separate management units in Queensland and to establish if gaps are truly present across the species' distribution. Future genetic studies may benefit from the inclusion of MHC genes to provide insight into levels of functional genetic diversity important for immune response and critical for the persistence of endangered species (Ellegren *et al.* 1993, Gutierrez-Espeleta *et al.* 2001).

Acknowledgements

We thank all those that spent long difficult hours in mangrove habitat collecting tissue samples, particularly Steve Van Dyck (Queensland Museum). We would also like to thank Jessica Worthington-Wilmer (Queensland Museum) and the South Australian Museum for providing tissue samples. The Wildlife Preservation Society of Queensland also generously funded part of this research through their student grants program.

References

Aplin, K. (2006). Ten million years of rodent evolution in Australasia: phylogenetic evidence and a speculative historical biogeography. In 'Evolution and Biogeography of Australasian Vertebrates'. (Eds J. R. Merrick, M. Archer, G. M. Hickey and M. S. Y. Lee) pp. 707-744. (Auscipub: Sydney).

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., and Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489-522.

Ball, D. (2004). Distribution and habitat of the false water rat, *Xeromys myoides* Thomas, 1889 (Rodentia: Muridae) in intertidal areas of central eastern Queensland. *Memoirs of the Queensland Museum* **49**, 487-494.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (1996). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire Génome, Populations, Interactions, CNRS UMR* 5000, Université Montpellier II, Montpellier, France.

Breed, W. G., and Aplin, K. (2008). Family Muridae. In 'The Mammals of Australia'. (Eds S. Van Dyck and R. Strahan) pp. 575-576. (New Holland Publishers: Sydney).

Bryant, L. M., Donnellan, S. C., Hurwood, D. A., and Fuller, S. J. (2011). Phylogenetic relationships and divergence date estimates among Australo-Papuan mosaic-tailed rats from the Uromys division (Rodentia: Muridae). *Zoologica Scripta* **40**, 433-447.

Clement, M., Posada, D., and Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**(8), 772.

Department of Environment and Resource Management [DERM] (2010). National Recovery Plan for the Water Mouse (False Water Rat) *Xeromys myoides*. Report to Department of Sustainability, Environment, Water, Population and Communities, Canberra. DERM, Brisbane.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Ellegren, H., Hartman, G., Johansson, M., and Andersson, L. (1993). Major histocompatibility complex monomorphism and low levels of DNA fingerprinting variability in a reintroduced and rapidly expanding population of beavers. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8150-8153.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Excoffier, L., Laval, G., and Schneider S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47.

Faircloth, B. C. (2008). Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**, 92-94.

Frankham, R. (2005). Genetics and extinction. *Biological Conservation* **126**, 131-140.

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., and Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology* **25**, 465-475.

Geyer, C. J. (1991). Markov chain Monte Carlo maximum likelihood. In 'Computer Science and Statistics: Proceedings of the 23rd symposium on the interface'. (Eds E. M. Kerimidas) pp. 156-163. (Fairfax Station: Interface Foundation).

Gibbard, P. L., Head, M. J., and Walker, M. J. (2010). Formal ratification of the Quaternary system/period and the Pleistocene series/epoch with a base at 2.58 Ma. *Journal of Quaternary Science* **25**, 96-102.

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-6.

Guindon, S., and Gascuel, O. (2003). A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* **52**, 696-704.

Gutierrez-Espeleta, G. A., Hedrick, P. W., Kalinowski, S. T., Garrigan, D., and Boyce, W. M. (2001). Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? *Heredity* **86**, 439-50.

Gynther, I. (2011). Distribution and ecology of the water mouse *Xeromys myoides* on Bribie Island, South-eastern Queensland. *Proceedings of the Royal Society of Queensland* **117**, 275-296.

Gynther, I., and Janetzki, H. (2008). The water mouse (*Xeromys myoides*). In 'The Mammals of Australia'. (Eds S. Van Dyck and R. Strahan.) pp. 664-666. (New Holland Publishers: Sydney.)

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucleic Acids Symposium Series* pp. 95-98.

Helgen, K. M., and Helgen, L. E. (2009). Biodiversity and Biogeography of the Moss-mice of New Guinea: A Taxonomic Revision of *Pseudohydromys* (Muridae: Murinae). *Bulletin of the American Museum of Natural History* **331**, 230-313.

Hitchcock, G. (1998). First record of the false water-rat *Xeromys myoides* (Rodentia: Muridae), in New Guinea. *Science in New Guinea* **23**, 141-144.

Ingvarsson, P. K. (2001). Restoration of genetic variation lost – the genetic rescue hypothesis. *Trends in Ecology and Evolution* **16**, 62-63.

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

Kingman, J. F. C. (1982). On the genealogy of large populations. *Journal of Applied Probability* **19**, 27-43.

Lacy, R. C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conservation Biology* **1**, 143-158.

Magnusson, W., Webb, G., and Taylor, J. A. (1976). Two new locality records, a new habitat and a nest description for *Xeromys myoides* Thomas (Rodentia: Muridae). *Wildlife Research* **3**, 153-157.

Michaux, J., Chevret, P., and Renaud, S. (2007). Morphological diversity of Old World rats and mice (Rodentia, Muridae) mandible in relation with phylogeny and adaptation. *Journal of Zoological Systematics and Evolutionary Research* **45**, 263-279.

Miller, S., Dykes, D., and Polesky, H. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215.

Moritz, C. (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution* **9**, 373-375.

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* **51**, 238-254.

Musser, G., and Carleton, M. (2005). Superfamily Muroidea. In 'Mammal Species of the World: a Taxonomic and Geographic Reference'. (Eds D. E. Wilson and D. M. Reeder.) pp. 894-1531. (Johns Hopkins University Press: Baltimore.)

O'Grady, J. J., Brook, B. W., Reed, D. H., Ballou, J. D., Tonkyn, D. W., and Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation* **133**, 42-51.

Piry, S., Luikart, G., and Cornuet, J. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 502-503.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2), 945-959.

QGC (2013). Environmental Management Plan – Water Mouse (*Xeromys myoides*). Revision 6

Rambaut, A. (2012). FigTree v1.4. *Published by the author*.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 4.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Redhead, T., and McKean, J. L. (1975). A new record of the false water-rat, *Xeromys myoides*, from the Northern Territory of Australia. *Journal of the Australian Mammal Society* **1**, 347-354.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539-542.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4**(1), 137-138

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Rowe, K. C., Reno, M. L., Richmond, D. M., Adkins, R. M., and Steppan, S. J. (2008). Pliocene colonization and adaptive radiations in Australia and New Guinea (Sahul): multilocus systematics of the old endemic rodents (Muroidea: Murinae). *Molecular Phylogenetics and Evolution* **47**, 84-101.

Rowe, K. M. C., Rowe, K. C., Elphinstone, M. S., and Baverstock, P. R. (2011). Population structure, timing of divergence and contact between lineages in the endangered Hastings River mouse (*Pseudomys oralis*). *Australian Journal of Zoology* **59**, 186-200.

Russell, T. L., and Hale, P. T. (2009). Conservation of the false water rat (*Xeromys myoides*) depends on landscape complementation. *Australian Mammalogy* **31**, 81-87.

Selkoe, K. A., and Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**(1), 264-279.

Spielman D., Brook, B. W., Frankham, R., and Schaal, B. A. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15261-15264.

Swann, C. A., Cooper, S. J., and Breed, W. G. (2007). Molecular evolution of the carboxy terminal region of the zona pellucida 3 glycoprotein in murine rodents. *Reproduction* **133**, 697-708.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731-2739.

Templeton, A. R., Crandall, K. A., and Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619-633.

Torgersen, T., Luly, J., De Deckker, P., Jones, M., Searle, D., Chivas, A., and Ullman, W. (1988). Late Quaternary environments of the Carpentaria Basin, Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology* **67**, 245-261.

Valiela, I., Bowen, J. L., and York, J. K., (2001). Mangrove forests: one of the world's threatened major tropical environments. *Bioscience* **51**, 807-815.

Van Dyck, S. (1996). *Xeromys myoides* Thomas, 1889 (Rodentia: Muridae) in mangrove communities of North Stradbroke Island, southeast Queensland. *Memoirs of the Queensland Museum* **42**, 337-366.

Van Dyck, S., Baker, W., and Gillette, D. (1979). The false water rat, *Xeromys myoides* on Stradbroke Island, a new locality in southeastern Queensland. *Proceedings of the Royal Society of Queensland* **90**, 84.

Van Dyck, S., and Durbidge, E. (1992). A nesting community of false water rats (*Xeromys myoides*) on the Myora sedgelands, North Stradbroke Island. *Memoirs of the Queensland Museum* **32**, 374.

Van Dyck, S., and Gynther, I. (2003). Nesting strategies of the water mouse *Xeromys myoides* in southeast Queensland. *Memoirs of the Queensland Museum* **49**, 453-480.

Van Dyck, S., Janetzki, H., Amey, A., Sherman-Hayes. L., Baker, A., Cox, T., Frost, A., and Avern, P. (2006). Mangroves, mansions and mice - the demise of a population of Water Mice (*Xeromys myoides*) adjacent to a Gold Coast canal housing development. Australian Mammal Society Newsletter, October 2006 No. 65.

Van Dyck, S., and Gynther, I. (2012). Water Mouse *Xeromys myoides.* In: 'Queensland's threatened animals' (Eds L. K. Curtis, A. J. Dennis, K. R. McDonald, P. M. Kyne and S. J. S. Debus.) pp. 408-409. (CSIRO: Canberra).

Woinarski, J., Brennan, K., Dee, A., Njudumul, J., Guthayguthay, P., and Horner, P. (2000). Further records of the false water-rat *Xeromys myoides* from coastal Northern Territory. *Australian Mammalogy* **21**, 245-247.

Wright, S. (1943). Isolation by distance. *Genetics* **28**, 114.

Wright, S. (1949). The genetical str ucture of populations. *Annals of Human Genetics* **15**, 323-354.

Table 1: Primer sequences and references for the three genes sequenced.

620

Table 2: Characteristics of the eight microsatellite loci developed for *Xeromys myoides*.

Table 3: Descriptive statistics for the microsatellite data. H_e, H_o and F_{is} were only calculated for the Coomera and Mackay regions as sample size limited the utility of these
625 **measures for other regions.**

Figure 1: Bayesian inference phylogeny generated from the CytB dataset using a General Time Reversible model gamma distributed with Invariant Sites (+I+ Γ). Nodes are marked with their Bayesian posterior probabilities. Unique haplotypes are as indicated in the names of each individual. The scale bar indicates the number of substitutions per site.

Figure 2: Statistical parsimony network analysis showing the relationship between haplotypes for *Xeromys myoides*. Each line indicates a one base pair difference. Dots on branches represent putative nucleotide changes between haplotypes. Each region is assigned an unique identifying pattern and the frequency of each haplotype is indicated by its size. Each haplotype is coded by the percentage contribution from each region. The network for CytB is shown on the bottom left and for D-Loop on the right.

Figure 3: Results of the structural analysis showing estimated membership coefficients for all individuals (*n*=49) for both two and three clustered plots. Each individual is represented by a single vertical column divided into colours representing its proportional membership of each cluster. Samples are grouped by pre-determined localities and divided by a thin black line. Samples are also ordered left to right from south to north.

Figure 4: Results of the Factorial Correspondence Analysis showing values for all individuals and a population average. Coomera area (*n*=20), Donnybrook (*n*=3), Agnes Water/Gladstone (*n*=2), Mackay area (*n*=22) and the Northern Territory (*n*=2). Each region formed a distinct cluster and is outlined and labelled. Each axis represents one factor (numbered to match the corresponding axis) that explains a component of the variability between samples.