Using Ancient DNA to Investigate Extinction, Extirpation and Past Biodiversity of Australian Macropods

Dalal Haouchar

Bsc (Honours) in Molecular Biology

(Murdoch University)

This thesis is presented for the degree of doctor of philosophy in biological sciences

Murdoch University

School of Veterinary and Life Sciences

2016

Declaration

I declare that this thesis is my own account of my research and contains, as its main content, work that has not previously been submitted for a degree at any tertiary education institution.

Dalal Haouchar

Abstract

The field of ancient DNA (aDNA) involves the isolation and retrieval of trace amounts of degraded DNA from a variety of substrates including fossils, sediments and historical material. The fragmentary nature of aDNA necessitates the use of methods with the ability to capture and amplify short segments of DNA. Collectively aDNA studies have made significant and unique contributions to a wide field of research including conservation, population genetics, taxonomy and phylogeny. The primary aim of this thesis research is to explore the utility of aDNA techniques to study extirpation, extinction and past biodiversity of Australian macropods. Using a combination of historical, Holocene and Pleistocene aged fossils, this research will attempt to investigate what ancient mitochondrial DNA (mtDNA) can add to our knowledge of Australia's macropods.

Traditional aDNA techniques have largely been used to isolate mtDNA from single fossil samples - an example of this approach is shown in Chapter Two where a well-preserved wallaby fossil bone from Depuch Island (Western Australia) was studied. The ancient mtDNA (cytochrome *b* and control region) data produced strong phylogenetic signal and shows that the Depuch Island rock-wallaby specimen is most similar to the mainland *Petrogale lateralis lateralis*. This finding has conservation implications for ongoing rehabilitation and translocation efforts in the Pilbara region.

Chapter Three of this thesis also uses mitochondrial aDNA techniques, to explore questions regarding interrelationships and former distribution of a macropod species complex; *Bettongia* spp. Cytochrome *b* and control region data retrieved from 88 historical samples, along with ~214 already sequenced samples, place the most recent common ancestor of the brush-tailed bettongs at *c*. 2.5 Myr. Ancient mtDNA is suggestive of connectivity between what are now highly fragmented populations, a result that has implications for how critically endangered brush-tailed bettongs should be managed.

Ancient DNA analyses and DNA sequencing technology have evolved over recent years and during the course of this study. Therefore in keeping up with the latest high-throughput sequencing (HTS) technology, aDNA analyses in ~70 bones and 20 sediment samples excavated from a Late Pleistocene–Holocene cave deposit on

Kangaroo Island, South Australia was undertaken. Samples were selected from 15 stratigraphic layers, ranging in age from >20 ka to ~6.8 ka. The successful retrieval of *bona fide* aDNA sequences, back to at least 20 ka, demonstrates excellent long-term DNA preservation at the site. All unidentified bones that were screened revealed a number of taxa from the assemblage including, *Macropus, Onychogalea, Potorous, Bettongia, Dasyurus, Rattus* and *Notechis*. The results from this study add significant value to the late Pleistocene-mid-Holocene paleontological record, detailing the past diversity of flora and fauna on Kangaroo Island.

Lastly, Chapter 5 introduces the latest molecular techniques in capturing and enriching highly fragmented aDNA bone from four sites across Australia. Ancient DNA extractions techniques, targeting ultra-short DNA fragments, were employed in an attempt to obtain Pleistocene-aged material. The warm conditions, a factor common in Australian caves, are not conductive to long-term DNA preservation at many sites. Shotgun sequencing was only successful on six bone samples (including one incisor) from a total of 25 samples that were screened. Three samples were successfully captured and enriched for endogenous DNA; one bettong sample generated 89.6% of a mtDNA genome with 5.4X coverage. Overall, the decay rate of DNA and preservation across all four sites was high, and extremely degraded, with an average fragment length between 47 bp and 57 bp. These data demonstrate that recovery of Pleistocene-aged aDNA from warm climate sites across Australia will remain a challenge and that better ways to screen and predict DNA survival are needed.

This thesis presents a combination of work from multiple sites across Australia using a range of aDNA techniques and sequencing technologies that have evolved over the tenure of this thesis. Collectively, this body of work has demonstrated the value of integrating aDNA data into modern-day conservation decision-making and has contributed to a wider understanding of Australian macropods both past and present.

Acknowledgments

It has been one massive journey, at times I thought I would never get there but with all good things, they come to an end. With just a little over four years in the making (yes it has been four years, minus all the suspensions), I would like to express my sincere gratitude to my primary supervisor, Prof Mike Bunce. Your patience and understanding throughout the many years we have worked together, even prior to taking on this challenge is greatly appreciated and will always form a milestone in my life. I have learnt so much along the way, thank you for encouraging me to stay motivated.

I would also like to thank Dr James Haile for the support over the years. Dr Charlotte Oskam, thank you for taking me on board and supporting me in the last year as my Murdoch supervisor. A special thank you to Dr Carlo Pacioni (can we please have a very big break from bettongs for a very long time). We have both had our fair share of frustrations over the computer crashes and always hitting wall time on the cluster computer, but we got there in the end. Thank you for your continuous support, your time and super thorough plans and ideas along the way. Your encouragement and advice in every aspect is greatly appreciated.

To the TREND group - you are the nicest bunch of people to work with and have been so supportive over the years. Our short coffee breaks are something I am going to dearly miss. In particular I would like to thank Alicia Grealy, Dáithí Murray and Dr Megan Coghlan, for the last minute edits on the chapters, support and advice. Also Tina Berry and Dr Nicole White, thank you for your support throughout the years.

During the last year of my candidature was one of the most exhausting, exciting but by far, the saddest time of my life. After being so overwhelmed at the arrival of my beautiful twins, I was saddened by the passing of my brother. What I thought was a challenge before this incident, became even more of a challenge and a struggle to emotionally get back on track. But with the support of my family especially my husband Ayman and my beautiful, supportive, mother Samar and father Talal, I was eager to continue and not give up. The encouragement from my sisters especially Donia will always be appreciated. I will always use this lesson in my life to teach my

children and anyone else that, although we cannot expect things in life to go as we plan we must always have a positive attitude, take a moment to pull ourselves together and continue in our life journey.

A final thank you to some of the collaborators I have got to know over the years. Thank you for all your feedback and input into the manuscripts and support, Dr Matthew Phillips and Dr Alexander Baynes. Also I cannot forget the field trip to Kangaroo Island, thank you Dr Gavin Prideaux and Dr Matthew McDowell for an unforgettable experience! I felt like "Dr Indiana Jones" in my caving suit and matching hardhat, with the little flashy light on top.

Photo of myself taken by Dr James Haile at one of the entrance points of Kelly Hill Cave, Kangaroo Island.

Table of Contents

List of Figures

Figure 6.2.1 **187**

Predictive map and experimental design strategy for screening Late Pleistocene and Holocene Australian fossil sites. Regions on the map show some key megafaunal assemblage sites found to contain a number of extinct macropods. Site numbers: 1, Devil's Lair; 2, Kudjal Yolgah Cave; 3, Mammoth Cave; 4, Tight Entrance Cave; 5, Nailtail Cave; 6, Kelly Hill Cave; 7, Naracoorte Caves; 8, Cuddie Springs; 9, Mooki River; 10, Ned's Gully; 11, Riversleigh World Heritage Area. Green circles on the map show sites where fossil material have not been previously genetically assessed, but have megafaunal deposits formerly dated to Late Pleistocene, hence red boxes with 'L' and should be genetically assessed using the experimental design outlined above. The 'H' blue filled boxes, denote to areas, which contain Holocene material, that are successfully yielding DNA. Orange-filled shapes on the map indicate those locations where DNA preservation has previously been assessed, although these sites should be revisited using freshly excavated material (i.e. from the locations highlighted in step 1) targeting the megafaunal layers. An experimental design (steps 2 – 3) should then be followed, which includes [i] an initial test for the preservation of DNA at the site [ii] using a web tool predicting the rate of DNA decay and the possibility of DNA surviving into the Late Pleistocene [iii] DNA hybridisation and capture to enrich for endogenous DNA and [iv] high-throughput sequencing (HTS) followed by mapping/analysis.

List of Tables

CHAPTER ONE

INTRODUCTION

Figure 1.1 Mosquito trapped inside amber fossil. (Public domain image from https://aztalksci.wordpress.com/category/science).

1.1 GENERAL INTRODUCTION AND THESIS LAYOUT

The content presented in this thesis is formatted to include two scientific manuscripts that are published (Chapters Two and Four), one that is accepted (Chapter Three), and one manuscript 'in preparation' (Chapter Five). Given the manuscript-style format of this thesis, the chapters consist of an abstract, introduction, materials and methods, results and discussion, conclusions and references sections. For the purposes of thesis cohesion and clarity between sections, each chapter begins with a preface and concludes with a chapter summary. In each preface section, a brief outline of the purpose of the chapters and any additional information not included in the manuscript is provided. Closing each of the chapters is a chapter summary section, which has been included to recapitulate the main findings and ideas and introduce the subsequent chapter. To maintain a consistent layout throughout this thesis, slight modifications were made to the published papers to enable referencing between chapters.

This multidisciplinary thesis covers an array of topics within the field of molecular biology, and some outside of this field (i.e. phylogenetics and phylogeography). Collaborative work was therefore required from researchers specialising in such areas to fulfil the requirements of each study. The roles of co-authors who assisted in this research are listed after the titles within the thesis chapters, others are mentioned in the acknowledgments section of each manuscript.

This introduction chapter is divided into four main categories. Each section introduces topics in order to summarise the scope of research covered throughout this entire thesis. Each main category is also supplemented with sub-categories that explore specific case studies and examples. The first section of Chapter One begins with a general introduction to provide an overview of the field of ancient DNA (aDNA - section 1.2). This section is further divided into six sub-categories exploring general topics on aDNA, including a brief history on aDNA through time. The next main category (Section 1.3) is further organised by seven key aspects, focusing on the evolution of DNA sequencing technologies and methods. In section 1.4 the applications of aDNA are explored with a particular focus on conservation (1.4.1), then phylogenetics and evolutionary history (1.4.2). These are some of the focal points of this thesis research. In section 1.5, Australia (a biodiversity hotspot) is explored, taking a closer look into Australian macropods – a major theme in this thesis research (section 1.5.2). Lastly, the general aims and scope will be outlined at the conclusion of this chapter (section 1.6). To overcome the repetitiveness common in a manuscript-style thesis layout, cross-referencing between chapters will be used to assist maintaining the overall flow between consecutive chapters.

1.2 ANCIENT DNA

1.2.1 ANCIENT DNA FAST TRACKED

The field of ancient DNA (aDNA), a division of molecular biology, has become a highly valuable and constructive science, pushing the boundaries of molecular biology, producing novel data, and yielding publications in high impact journals (e.g. Science; Poinar *et al.*, 2006; and Nature; Orlando *et al.*, 2013). Ancient DNA studies involve the isolation and retrieval of trace amounts of degraded DNA from fossils and historical material. In 1984, for the very first time, researchers were able to use a cloning-based approach, to recover traces of aDNA from an extinct museum specimen, the quagga, (*Equus quagga*) (Higuchi *et al.*, 1984). Prior to this study, researchers focused extensively on modern DNA sampling, placing restrictions on aged samples.

Today, we can appreciate an array of sources being utilised for aDNA sampling (see section 1.2.2). Although the field of aDNA has quickly developed over the years, in the past, this science was not-so-popular when a ground-breaking discovery revealed an 80 million year old dinosaur DNA sequence (published in Science) was, in actual fact, modern human contamination (Woodward *et al.*, 1994). The novel, 'Jurassic Park' by Michael Crichton (later created into the Jurassic Park film) was inspired by the idea that dinosaur DNA contained inside insects (such as mosquitoes) fossilised and preserved in amber (Figure 1.1), could then be extracted and cloned in an attempt to bring dinosaurs back to life. At the time, aDNA studies that attempted to prove that this technique was a practical approach were mostly published in high impact journals, which damaged the reputation of the field once researchers showed that this was actually possible. In a more recent study, scientists from Manchester University used a more advanced sequencing approach called high through-put sequencing technology (HTS – see section 1.3.2 for more on this technique) to preferentially produce sequence information of all DNA molecules in an extract (Penney *et al.*, 2013). The results revealed that this technique was unlikely to succeed, given that they were not successful in detecting aDNA, in a much younger (10,612 $\pm$ 62 cal yr BP) sub-fossilised insect preserved in Copal (the sub-fossilised resin precursor of amber; Penney *et al.*, 2013). It took many years for the field of aDNA to overcome this and progress to what it is today (Shapiro *et al.*, 2012).

Despite the near downfall, researchers are now free to explore the potential that aDNA has to offer by broadening the types of substrates used for aDNA analysis. This will form the focus of the following section.

1.2.2 ANCIENT DNA SOURCES

Earlier aDNA studies used mainly soft tissue, based on the assumption that materials such as muscle contain a lot of DNA in living organisms, and therefore should retain more DNA *post-mortem* than other less DNA-rich tissues (Shapiro *et al.*, 2012). It was only after several reports were made that the suggestion arose that ancient bone contains on average more DNA than ancient soft tissue. This is because the hydroxyl-apatite contained within the bone, where the DNA presumably adheres to (Campos *et al.*, 2012), creates a barrier that protects the DNA from degradation. From there on, researchers concentrated mostly on non-mineralised fossil bone (or sub-fossils) as a source of aDNA as it clearly preserved DNA (with some exceptions) and is fairly abundant in excavations. However, the discovery of biological sources from which DNA can be extracted did not stop there. In some instances, *post-mortem* DNA is transmitted into the environment upon death and can survive in the environment for some time depending on the environmental conditions; this is known as sedimentary aDNA, or *sed*aDNA (Haile *et al.*, 2007). An example of *sed*aDNA can be seen in permafrost from Siberian cores yielding aDNA from ages ranging between 400 ky (thousand years) and 10 ky (Willerslev *et al.*, 2003). Even more extreme, the presence of aDNA from ice cores sampled 2km down from Greenland's surface has been found (Willerslev *et al.*, 2007).

A study exploring the potential of isolating DNA from coprolites (sub-fossil faeces) was also successful in amplifying ground sloth aDNA from coprolites that were dated to approximately 11 ky to ~ 28 ky (Hofreiter *et al.*, 2000). Other substrates from which aDNA has been obtained include eggshells (Oskam *et al.*, 2010), urine (Valiere *et al.*, 2000), feathers (White *et al.*, 2011), and hair (Almeida *et al.*, 2011). Although it may seem as if DNA can be readily utilised from the majority of sources, this is not the case because the preservation of DNA is highly variable from one sample to another. The degradation of DNA, influenced by the rate at which DNA is broken down, is a major factor contributing to the variance between samples and sources, and is explored further below.

1.2.3 ANCIENT DNA AND DEGRADATION

As previously mentioned, DNA can be extracted from a variety of biological materials. These materials can originate from museum skins/skeletal material, or archaeological/paleontological remains, all of which can be degraded at different rates depending on their preservation conditions (Leonard, 2008). Within a living organism, the cells are continually being maintained and repaired by enzymatic repair processes. Following the death of the organism, cellular repair processes no longer counteract cellular damage. Damage therefore accumulates progressively, to the point of irreversible loss of nucleotide sequence information (Pääbo *et al.*, 2004; Roberts *et al.*, 2008). DNA degradation can also be carried out by enzymes such as lysosomal nucleases, and further attacked by bacteria, fungi, and insects that feed on and digest the macromolecules (Pääbo *et al.*, 2004). Processes such as oxidation, as well as direct and indirect effects of background radiation, also modify the nitrous bases and the sugar-phosphate backbone of the DNA (Figure 1.2.1).

Furthermore, other degrading processes such as deamination, depurination and other hydrolytic processes will also lead to destabilisation and breaks in DNA (Lindahl, 1993; Hofreiter *et al.*, 2001). Under certain circumstances (e.g. cool climates), however, DNA has a chance at survival. Although degradation results in very low concentrations of short damaged fragments, the sequence information can still be recovered by amplification with a routine molecular biology technique – polymerase chain reaction (PCR). The invention of PCR made it possible to clonally amplify and study single surviving molecules, allowing aDNA research to increase rapidly (Willerslev and Cooper. 2005). However, this procedure later uncovered sequencing artefacts and preferential amplification of undamaged contaminant DNA (Lindahl, 1993). Also, the common bias towards CG→TA transitions is regularly seen in aDNA-amplified products. Other factors, such as nucleotide misincorporation during the amplification step, are also commonly observed. In addition, miscoding lesions (Figure 1.2.1), they manifest as base modifications in a PCR amplified sequence changing the appearance of the DNA template and potentially misleading haplotype analysis (Ho *et al.*, 2007).

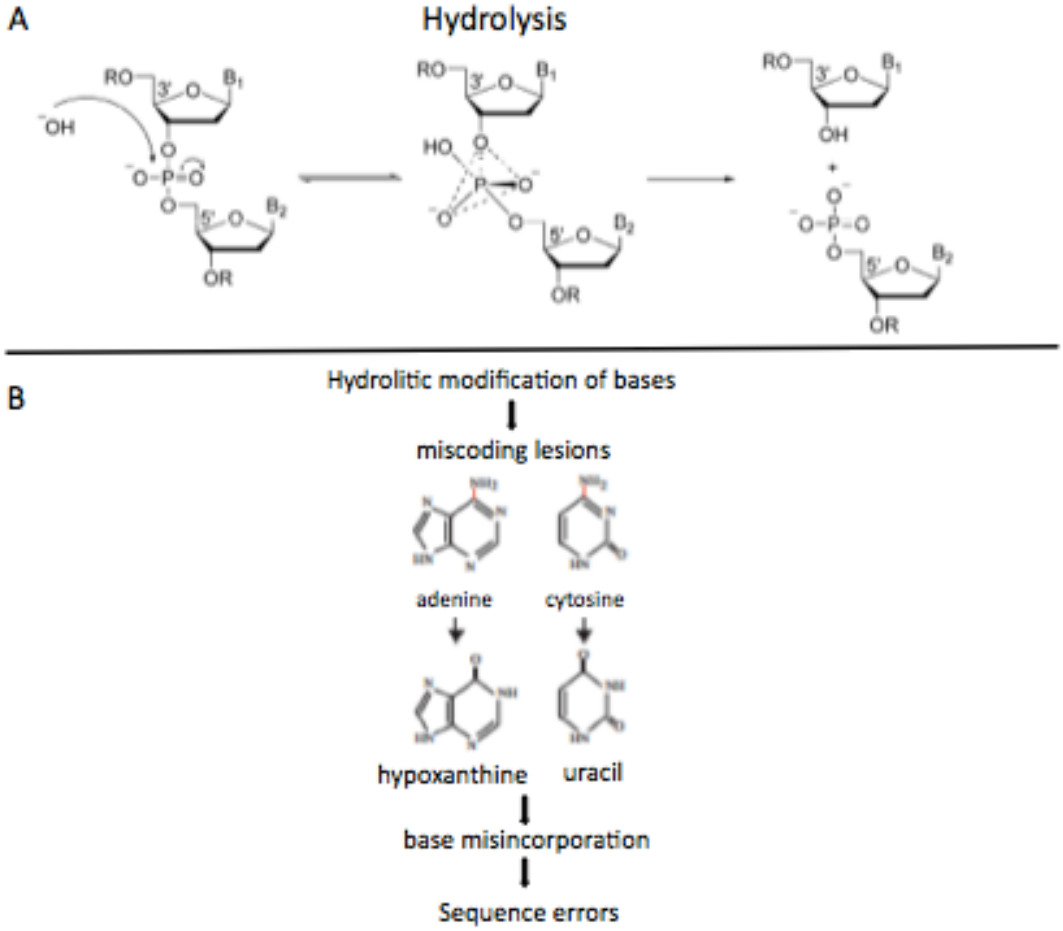


Figure 1.2.1 (A) DNA hydrolysis reaction, a nucleophilic attack at the DNA phosphate backbone. (B) Hydrolytic modification of bases resulting in miscoding lesions. Modified from (Sangeetha *et al.,* 2014).

1.2.4 CONTAMINATION

Even though bone, dentine and tissue are the most obvious sources of preserved nucleic acids, they are still greatly susceptible to environmental contamination. Older samples in particular are more vulnerable to exogenous environmental contamination, introducing erroneous miscoding lesions producing DNA artefacts and preferential amplification of contaminant sequences (Pääbo *et al.*, 2004; Gilbert *et al.*, 2005; Axelsson *et al.*, 2008). Contamination by modern exogenous DNA is the most serious concern in aDNA studies and can be a problem from the moment the sample

is collected to contamination arising from within the lab and throughout the PCR process (Hofreiter *et al.*, 2001; Pääbo *et al.*, 2004; Willerslev *et al.*, 2005; Deguilloux *et al.*, 2011). Due to the severe sensitivity of aDNA samples, a routine process, commonly known as "criteria of authenticity", is employed as a means of determining the authenticity of the data by following a 'check list'. An example of these criteria can be found at the following (Hofreiter *et al.*, 2001; Pääbo *et al.*, 2004; Gilbert *et al.*, 2005; Willerslev *et al.*, 2005).

1.2.5 KEEPING IT COOL

As discussed above, it is clear that aDNA can persist in a number of biological materials (section 1.2.2). However, post-mortem degradation (section 1.2.3) can change the outcome of DNA survival and environmental contamination can pose as a serious risk (section 1.2.4). In addition to all these factors affecting DNA survival, there still remains another key influence impacting DNA preservation – temperature. Constant low temperatures and dry conditions are generally desirable for the prolonged existence of aDNA molecules (Lindahl, 1993; Willerslev *et al.*, 2005; Leonard, 2008). Collectively, environmental conditions (e.g. temperature, moisture and pH) in combination with time since death (post-mortem interval-PMI) are primary factors influencing DNA degradation, but the relative effects of environment and time appear to be strongly situation dependent (Higgins *et al.*, 2015). Also over long PMIs (hundreds to thousands of years), mtDNA appears to degrade at a slightly slower rate than nuclear DNA (see later for more discussions on nuDNA). Recent work by Allentoft *et al.* (2012) suggest that under a range of conditions, DNA degradation follows a random fragmentation model and, at least in bone, the rate of mtDNA degradation can be predicted based solely on PMI and ambient temperature. Determining the rate at which DNA degrades in the past has been difficult because it is rare to find large sets of DNA-containing fossils with which to make meaningful comparisons. Adding to this is the variable environmental conditions such as temperature, degree of microbial attack and oxygenation, all altering the speed of the decay process. However, understanding the rate at which DNA decays can provide useful information about paleontological cave sites, giving palaeogeneticists an indication of whether preservation is present. Using the DNA fragmentation theory as adopted by Allentoft *et al.* (2012) (a mechanism that was firstly proposed by

Lindahl, 1993), is one way that the average DNA half-life within geographically constrained fossil assemblages can be estimated. Another approach is using the Thermal Age Web Tool (Collins *et al.*, 2009), a computer model and web interface for quantifying chemical reactions in archaeological specimens, which can produce results rapidly and aims to predict the amount of decay in ancient DNA (as found in bone and hair, which often survive for a long period of time). Information about the site of interest such as the source of the sample, estimated age, the surrounding conditions the material was found in (e.g. in a cave? what layer? what type of soil? estimated temperatures?), are required to do the predictions using mathematical models. The Thermal Age tool is purely predictive, whereas the method from Allentoft *et al.* (2012) study is based on empirical data. Such models, are an integral way to scope-out archaeological sites that have the potential of preserving DNA before materials, funds and resources are injected into such projects, which would otherwise be costly and time consuming (Smith *et al.*, 2003). The Thermal Age tool is also another way in which valuable, rare and unique museum specimens can be protected from unnecessary destructive sampling.

Since temperature is a major driver of long-term DNA preservation, aDNA studies utilising samples from permafost environments have been commonplace, yielding the best preserved and oldest aDNA sequences to date. This includes the greater than 65 kyr old bison mtDNA from North America (Shapiro *et al.*, 2004) to the 450 - 800 kyr old insect DNA from Greenland ice-cores (Willerslev *et al.*, 2007). Although the cooler climates result in better preservation of aDNA than the warmer regions around the world, questions continue to arise about the authenticity of the data. Nevertheless, aDNA data on its own is a powerful tool and can provide insights into a species past, which would otherwise be lost with the use of modern DNA data alone.

1.2.6 ANCIENT BIOMOLECULES – A HOP BACK IN TIME

Information about a species' past biodiversity, prehistoric dispersal and colonisation events, consequences of environmental changes, and extinction processes (Hofreiter *et al.*, 2001; Willerslev *et al.*, 2005), are all matters that can be explored through the utility of aDNA. However, of all the aDNA studies in the literature, it is by far the charismatic extinct megafauna of the Pleistocene that has caught the attention of

molecular biologists around the world. Some of the species that have been publicised globally include the woolly rhino (*Coelodonta antiquitatis*) (e.g. Orlando *et al.*, 2003; Willerslev *et al.*, 2009), woolly mammoth (*Mammuthus primigenius*) (e.g. Greenwood *et al.*, 1999; Gilbert *et al.*, 2008; Nogues-Bravo *et al.*, 2008), steppe bison (*Bison priscus*) (e.g. Shapiro *et al.*, 2004), and cave bear (*Ursus spelaeus*) (e.g. Hofreiter *et al.*, 2004; Stiller *et al.*, 2010), among others.

Over the last 100 ka, there were at least 340 species of land mammals in Australia, 67 of them are now extinct, which has been a result of three waves of extinctions (Johnson, 2006). The first occurred during the late Pleistocene between 130 ka and 10 ka; within the last glacial cycle, over 50 species of mainly large marsupials disappeared. The extinction of these large marsupials is referred to as the late Pleistocene megafauna extinction, which dramatically reduced the number of vertebrates in Australia. Today, palaeontologists have discovered many megafaunal assemblages from Australian caves, retrieving a large amount of megafaunal bones. Of these, ancient kangaroo remains are abundant (Prideaux, 2004; Sears, 2005; Prideaux and Warburton 2010). The extinct giant sthenurine kangaroo (Diprotodontia; often referred to as the "short-faced" kangaroo; Figure 1.2.2), is an example of one of the Australian megafauna species, amongst many others, that went extinct from the continent during this time (Johnson, 2006).

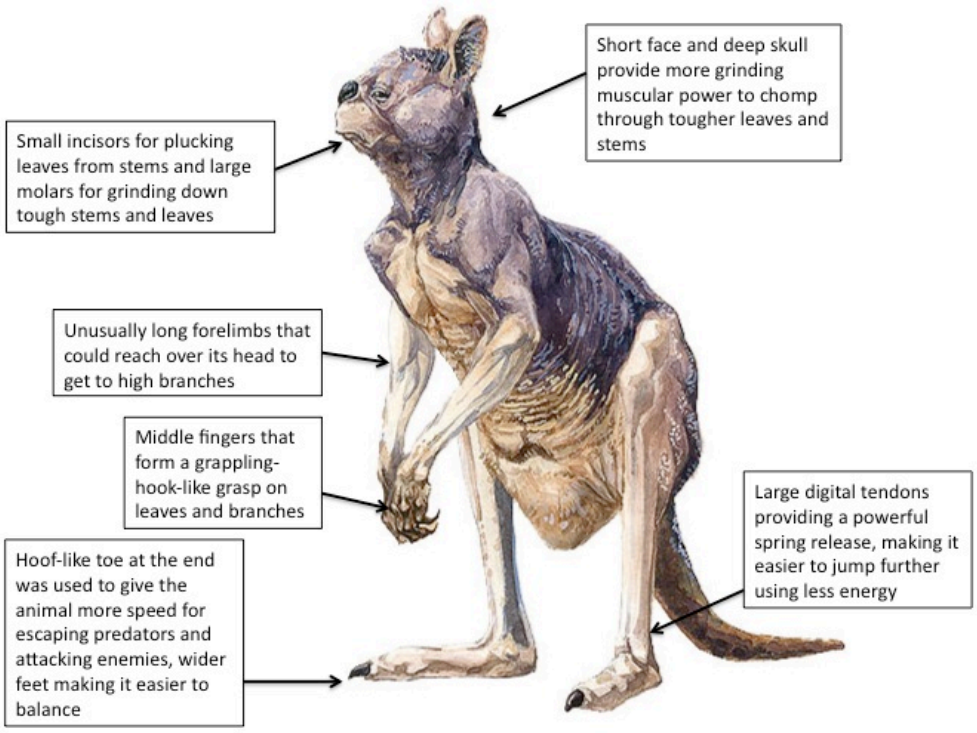


Figure 1.2.2 Adaptations of the Giant short-faced kangaroo – *Procoptodon goliah.* The largest of the Pleistocene sthenurine kangaroos weighing over 200 kg, about double the size of a modern day kangaroo. JoyZine. (2009). Giant Short-faced Kangaroo. Retrieved Feb 2016 from http://www.artistwd.com/joyzine/australia/articles/megafauna/procoptodon_goliah.php.

From a morphological perspective, a large selection of extinct kangaroo bones from a number of cave sites across Australia have been studied extensively (Prideaux, 2004; Sears, 2005; Prideaux *et al.*, 2010); however, from a molecular perspective, no studies to date have produced DNA information from the extinct megafauna kangaroo from the Australian mainland. One recent study, however, managed to isolate mtDNA sequences from two extinct kangaroo fossils found in the cool climate caves within the mountainous Tasmanian region. The short lengths of mtDNA sequences provided limited information at first, until the aDNA nucleotides were combined in a meta-analysis with modern molecular DNA and morphological data (Llamas *et al.*, 2015). In order to fully resolve the taxonomy within the extinct kangaroo lineage, molecular aDNA sequences are clearly needed to supplement current morphological studies (Llamas *et al.*, 2015). Using a multidisciplinary approach to explore taxonomy related questions and past biodiversity is becoming common practise, especially when biologists need to investigate both extinct and extant species.

Comparing ancient and modern DNA data using molecular techniques is a framework that has previously been used to interpret the population origin of species or lineages (e.g. Steeves *et al.*, 2010). Furthermore, adding a temporal context to combined analyses (i.e. aDNA, modern DNA and morphology), the time of divergence of a species compared to other taxa can be determined; this is done by integrating a molecular clock analysis, which enables researchers to determine an accurate time frame of when a species lineage diverged from a common ancestor (Mourier *et al.*, 2012). Using aDNA techniques to explore questions (about species divergence) is difficult when faced with challenges of DNA preservation, especially in localities that are not commonly conducive to DNA survival, such as mainland Australia. The rapid advancement in sequencing technology and DNA extraction techniques, is therefore pivotal in progressing the field of aDNA forward, and can help overcome many of the obstacles encountered in the aforementioned discussions.

1.3 THE EVOLUTION OF DNA SEQUENCING TECHNOLOGIES

Sampling and DNA extraction techniques have remained largely unchanged since the earliest publications in aDNA research, over two decades ago. The ability of HTS platforms to sequence ultra-short fragments (i.e. not reliant on PCR amplifiable fragments) has recently changed these approaches. More recent methods are cognisant of DNA fragment sizes and the (often) contaminated nature of paleo-samples. The first steps in DNA analysis of historical remains are critical and can have major impact on the amount and integrity of recovered endogenous DNA, contamination, and co-extraction PCR inhibitors, thereby affecting the success of downstream analysis. The following section provides an overview of some molecular biology techniques and their progression through time, leading into the most current and up-to-date techniques used for optimal DNA extraction and sequencing of aDNA substrates.

1.3.1 SANGER SEQUENCING AND CLONING DNA

The field of genomics, has been constantly evolving since Watson and Crick discovered the structure of DNA in the 1950s (Watson *et al.*, 1953). In the 1970s, the F. Sanger's group further elucidated the DNA structure through the development of

the di-deoxy chain termination method; as a result, the first genomic landmark of deciphering the phiX174 bacteriophage genome was achieved (Sanger *et al.*, 1977; Men *et al.*, 2008). From there on, the ability to replicate DNA *in vitro* resulted in the development of the PCR which irrevocably changed the field of molecular biology (Willerslev *et al.*, 2005). Current sequencing methods introduced in 2000s have come a long way, enabling DNA to be sequenced at unprecedented levels and speed, producing a multitude of data from a single sample (Shendure *et al.*, 2008). This new technology, known as high-throughout sequencing (HTS), was introduced in order to explore a wider scope of applications than could previously be explored with Sanger sequencing (see section 1.3.2 and 1.3.3 for more detail). Some of the limitations of past technology are exemplified in the traditional methodology of cloning; that is, cloning foreign DNA in vectors that are compatible with replication machinery of the *E. coli* cells (Figure 1.3.1). Each clone needs to be amplified separately and Sanger sequenced one-by-one, a very low-throughput approach to verifying authentic DNA, as apposed to simultaneously (which is the approach HTS takes). Other limitations of Sanger sequencing and by far the most significant, is the cost. At $1 per kilobase, it would cost $10 million to sequence a 1 Gb genome to 10X coverage (Men *et al.*, 2008). Nonetheless, combining Sanger outputs with HTS information is useful, and will be discussed in the following sections.

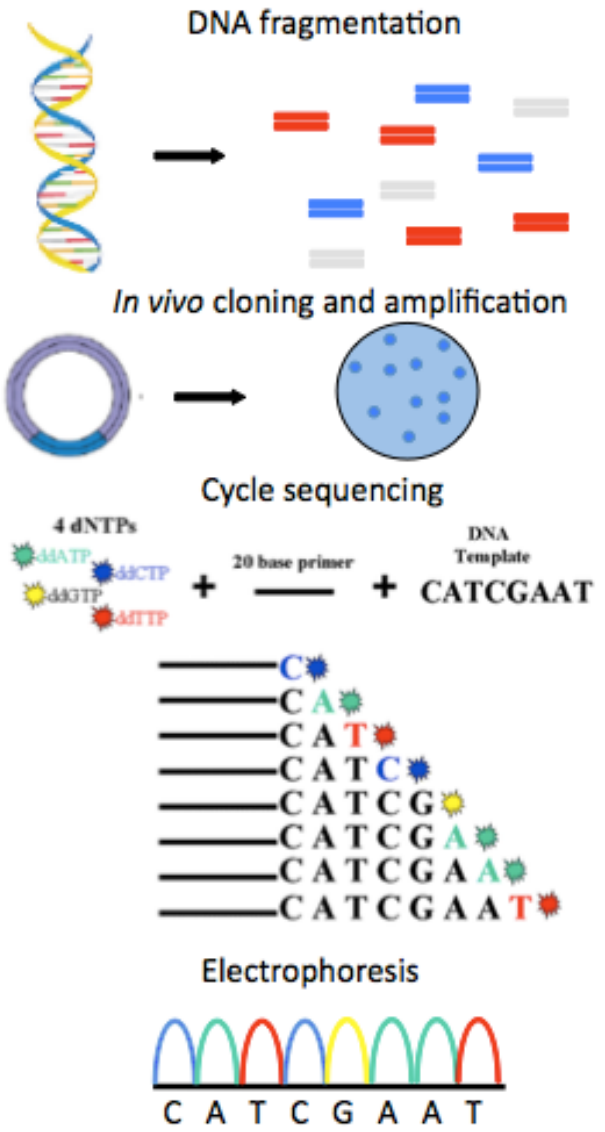


Figure 1.3.1 A brief over view of the cloning technique and Sanger sequencing. DNA is fragmented and cloned into a vector such as a plasmid and clonally amplified. Cyclic sequencing is carried out followed by separation by electrophoresis. Sequence reads and chromatograph readings are produced (1 read/capillary).

1.3.2 ANCIENT DNA AND HIGH THROUGHPUT SEQUENCING

Ancient DNA studies have been massively hindered due to the fragmentary nature of surviving DNA; highly degraded aDNA samples are mostly restricted to short-sized fragments and the majority of DNA being exogenous (non-target DNA). A common problem of DNA sequencing with Sanger methods is poor quality in the first 15-40 bases of the sequences due to its limitations in resolving the short dye-terminated products at either end of the sequence. Therefore, Sanger sequencing products that are less than 100 bp (a common fragment size in palaeo-samples) are not capable of

producing enough informative data for aDNA analysis. Although the DNA fragments may be short in size, recent developments in sequencing technologies have the ability to reproduce and help authenticate the data, largely changing the scope of aDNA research. It is with these HTS platforms (see below) that have helped achieve several orders of magnitude more sequence data compared to traditional Sanger sequencing methods.

1.3.3 SEQUENCING PLATFORMS

The establishment of Roche's (454) GS FLX Titanium instrument in 2005 (Margulies *et al.*, 2005), followed by other technology like Ion Torrent (Life Technologies) (Shendure *et al.*, 2005) and MiSeq (Illumina) platforms (Bentley *et al.*, 2008), has made it possible to generate gigabases of sequence data within a few days (Kircher, 2012). This would be impossible using the more conventional method of Sanger sequencing, mainly because of its inability to sequence many fragments in parallel. The rapid development of sophisticated bench-top sequencers such as these have made HTS technologies applicable to a wide range of fields (Tillmar *et al.*, 2013) and consequently revolutionised the field of genomics. An example of this was seen from within only a few months of the introduction of HTS, where Poinar *et al.* (2006) published 13 million bp from the nuclear genome of an extinct woolly mammoth, representing the largest nuclear data set available from an extinct species during the pre-HTS era. This study paved the way for more ambitious projects; for example, in 2008, the first low-coverage (0.8-fold) draft genome of an extinct mammoth species was published (Miller *et al.*, 2008) and in 2010, the first high coverage (20-fold) ancient human genome was obtained from a 4,000 year-old permafrost-preserved hair sample from a paleo-eskimo (Rasmussen *et al.*, 2010).

Given that now there are a number of HTS platforms available, the choice of instrument comes down to the desired application and the quality of the input DNA. Currently there are six 2nd and 3rd generation sequencing platforms available, and a seventh is in development (Table 1.3.1). Costs for these systems range from $1,000/run (e.g. 454 GS jr. Titanium – not shown in Table 1.3.1), to over $20K/run (e.g. Illumina HiSeq 2000), with instrument run times ranging from several hours to weeks, and data producing from approximately <100,000 reads/run (e.g Ion Torrent – ′314′ chip) to >1 million reads/run (e.g. SOLiD 5500xl) (Glenn, 2011). Although

the instrument of choice is project dependent, each comes with advantages and disadvantages that are briefly summarised below (Table 1.3.1).

Table 1.3.1 Brief comparison of the primary high-throughput sequencing (2nd and 3rd generation) platforms with advantages/disadvantages and primary applications listed, as of 2013 (Glenn, 2011; Liu *et al.*, 2012).

Platform/Company	Sequencing method	Advantages	Disadvantages	Applications
454/Roche	Synthesis (pyrosequencing)	Long read length	Prone to homopolymer error	1*, 2, 3*, 4, 7
Illumina/Illumina (HiSeq, Miseq)	Synthesis	High-throughput (short-read sequencer)	Short read length	1*, 2, 3*, 4, 5, 6, 7
SOLiD/Life Technologies	Ligation	High accuracy-low error rates (short-reads)	Short read length	3*, 5, 6
HeliScope/Helicos	Synthesis	Single-molecule sequencer	Short read length	5, 8
Ion Torrent/Life Technologies	Ion semiconductor	Fast sequencing	Prone to homopolymer error	1, 2, 3, 4
PacBio/Pacific Biosciences	Synthesis	Longest read length	Lower throughput	1, 2, 3, 7
Nanopore /Oxford Nanopore technologies	IONpore	Long read length	High error rate	1, 2, 3, 4

1 = *de novo* BACs, plastids, microbial genomes
2 = transcriptome characterization
3 = targeted re-sequencing
4 = *de novo* plant and animal genomes
5 = re-sequencing and transcript counting
6 = mutation detection
7 = metagenomics
* = pooling multiple samples with sequence tags (i.e. MIDs or indexes) is required for efficient use of this application

After briefly exploring and comparing some HTS platforms, the following sections will now take a closer look at some sequencing strategies, and highlight their applications and importance in aDNA research. Some of the applications in which HTS technology can be used for include amplicon sequencing (1.3.4), whole genome shotgun sequencing (section 1.3.5), and metagenomics.

1.3.4 AMPLICON SEQUENCING OF aDNA USING HTS METHODS

Amplicon sequencing can be briefly summarised by a process that involves the amplification of targeted regions of DNA, which are subsequently pooled together and sequenced in parallel on the most suitable platform (Murray *et al.*, 2015). Amplicon sequencing is a cost-effective approach, because it is both targeted and conducive to multiplexing. This method requires the use of a unique six to ten base pair multiplex identifier (MID), which is attached to a gene specific primer. It is the use of these MID tags that allow for samples to be uniquely identified from one another and allows for a large number of samples to be analysed concurrently. This method also has the potential to provide enormous amounts of data, albeit with the possibility of high error rates, although this is variable between platforms (Murray *et al.*, 2015). The use of error correction algorithms are therefore required depending on the choice of the HTS platform used. For example, Ion Torrent, 454 and PacBio sequencing platforms typically generate insertions and deletions during homopolymer tracts, while Illumina platforms tend to generate substitution errors most frequently (Quail *et al.*, 2012). Although amplicon sequencing may present a way forward for retrieving genetic sequence information, the technique becomes a limitation when dealing with highly fragmented and degraded DNA, as the sequence length on average is shortened to around 50 bp. Using amplicon sequencing techniques on short fragment lengths becomes a challenge due to the requirement of the PCR primer binding on the fragment, leaving not much room for primer design. Shotgun sequencing of DNA has, in many respects, helped circumvent this problem.

1.3.5 SHOTGUN SEQUENCING OF aDNA

Shotgun sequencing represents the most basic approach that utilises HTS technology. It has been used for sequencing complete genomes of extinct species from well preserved permafrost mammoths (Miller *et al.*, 2008), to more poorly preserved Neanderthal remains (Green *et al.*, 2006). Shotgun sequencing is utilised when isolated DNA is sequenced without any *a priori* selection (Rizzi *et al.*, 2012). The complete mammalian genome is about 2.8–4 billion bp, while the mitochondrial genome is only about 17,000 bp long. Thus, the mitochondrial genome makes up about 0.0004% of the entire genome. Keeping in mind that aDNA already exists in very low copy number, sequencing DNA from poorly preserved Neanderthal remains, for example, can result in less than 0.03% mitochondrial reads (Green *et al.*, 2008).

Exogenous DNA, like bacteria, usually take up a large part of the overall reads produced, making this approach somewhat problematic when dealing with ancient samples (Knapp *et al.*, 2010). Due to the high level of bacterial contamination in the shotgun approach, a strategy to decrease the proportion of microbial DNA sequences captured should be a priority when dealing with heavily degraded and contaminated samples, common in historical material. This strategy is the newest advancement in aDNA methods known as target enrichment strategies (see section 1.3.6 for more detail).

A shortcoming of the shotgun sequencing approach is the presence of repeated sequences being clonally amplified during the PCR steps. Shotgun sequencing is based on the assumption that a pair of reads would overlap a region, enough so that they form contiguous sequences in order to provide a homogenous coverage of the genome (Figure 1.3.2, I). A high number of unique reads can produce appropriate coverage and confirmation of sequence data (Figure 1.3.2, II), which allows you to check for error rates and homopolymer runs that are increasingly common in HTS output (although are notably inconsistent between platforms). However, many repeated sequences could indicate a sample has been sequenced to saturation point creating clonally amplified DNA fragments that are not informative (Figure 1.3.2).

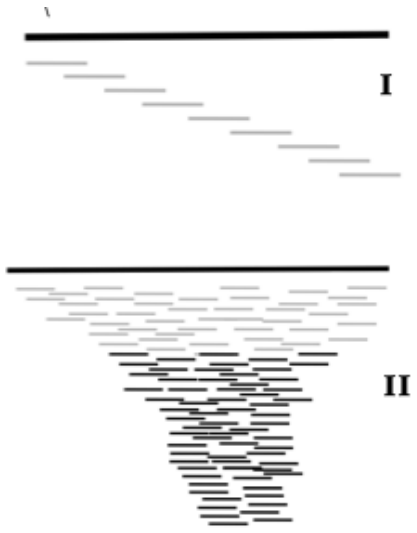


Figure 1.3.2 Repeated sequence assembly, black bars indicate sequence reads in which the same region has been amplified multiple times, grey bars indicate reads which give unique sequence information. Solid black line resembles the reference sequence. I. Contiguous sequence reads that appear to overlap across the entire reference genome, this pattern is ideal. II. The resultant assembly of clonally amplified sequences reaching saturation point.

1.3.6 DNA ENRICHMENT STRATEGIES

Efficient targeted capture and enrichment of specific DNA sequences from complex genomes as an alternative to PCR-based methods, are becoming more widely applied in molecular biology, especially in aDNA research (Albert *et al.*, 2007). The problem previously within the field of aDNA lay in the inability to capture on-target material to the exclusion of exogenous DNA in a sample. The lack of 'front-end' technology in the past had previously hindered the potential for aDNA research to take control of such capture methods; however, today this technique can be fully leveraged through the power of HTS technologies. Ancient DNA that can be successfully captured is usually derived from the mtDNA genome, and outweigh the proportion of nuclear DNA (nuDNDA) sequences found in a cell (Knapp *et al.*, 2010). In order to balance out the system, numerous hybridisation capture methods have been introduced and include on-array capture as well as in-solution capture (Gnirke *et al.*, 2009; Knapp *et al.*, 2010; Carpenter *et al.*, 2013; Dabney *et al.*, 2013; Tillmar *et al.*, 2013).

Due to the highly fragmented nature of aDNA, an ideal enrichment technique would target as much of the endogenous genome as possible so as not to discard informative sequences (Carpenter *et al.*, 2013). There are two main categories of target enrichment techniques commonly employed in aDNA research: PCR amplification of target regions (amplicon sequencing – see previous section 1.3.4) and DNA hybridisation capture (Knapp *et al.*, 2010). Targeting specific genomic loci in ancient populations is more beneficial than trying to sequence the complete genome via shotgun sequencing as described above (section 1.3.5) because it allows the sequencing of contaminating exogenous DNA to be minimised. Endogenous DNA exists in historical material at extremely low copy number; therefore, enriching for DNA prior to sequencing is highly useful in aDNA studies (Horn, 2012). An 'in-solution' capture enrichment approach is an unbiased mean to increase the proportion of endogenous DNA in aDNA sequencing libraries, and is investigated further in the case study below.

1.3.7 DNA ENRICHMENT AND CAPTURE – AN EXAMPLE FROM ANCIENT CAVE BEAR

It is not surprising that permafrost environments have yielded the oldest credible records of DNA survival. This includes short stretches of plant and invertebrate

DNA with an estimated age of up to 800 ka that were amplified by PCR from Artic ice cores (Willerslev *et al.*, 2003; Willerslev *et al.*, 2007). To date, only a single study has convincingly raised the possibility of DNA survival extending far into the Middle Pleistocene (c. 780 – 126 ka) outside permafrost (Valdiosera *et al.*, 2006). However, in a recent study, a cave site from Sima de los Huesos (Atapuerca, Spain) has produced a number of interesting bones, some of which were estimated to be >300 ka (Dabney *et al.*, 2013). The technique used to capture the DNA from bear samples was based on retrieving ultra-short PCR products approximately 50 bp at a time – to piece together a draft of the bear's genome. Achieving this firstly involved the development of a new extraction technique to preferentially target ultra-short fragments of DNA. This was done using silica spin columns that contained custom adapted reservoirs to hold large amounts of buffer during the DNA extraction step. Once the DNA was extracted from the sample, the solution was used to build customised libraries using single-stranded library preparation methods as shown in (Gansauge *et al.*, 2013 - see also Figure 1.3.3 for an experimental diagram of library build and enrichment - capture protocol). Shotgun sequencing was then carried out on the library using 10% of a MiSeq lane (Illumina). A significant number of sequences did not primarily map to the modern bear genome; therefore, an enrichment technique was used to capture and specifically amplify the endogenous DNA (Maricic *et al.*, 2010). After mtDNA captures were performed, the final consensus of 16,305 bp of the cave bear mitochondrial genome was produced (Dabney *et al.*, 2013).

Although the development of HTS technology has greatly impacted most fields of molecular biology, few have benefited as much as the field of aDNA. Techniques such as DNA enrichment and capture (Figure 1.3.3 for an example) have now allowed for very old DNA to be extracted from highly degraded and contaminated samples. These new methods, coupled with the latest HTS technologies, allow information to be utilised from archaeological and paleontological remains giving researchers the ability to 'go back in time' and study the genetic relationships between extinct organisms and their contemporary relatives (Rizzi *et al.*, 2012). Ancient DNA studies have also made significant and unique contributions to a wide field of research including conservation, population genetics, taxonomy and phylogeny. Some of these fields of research and applications will form the discussions of the following sections.

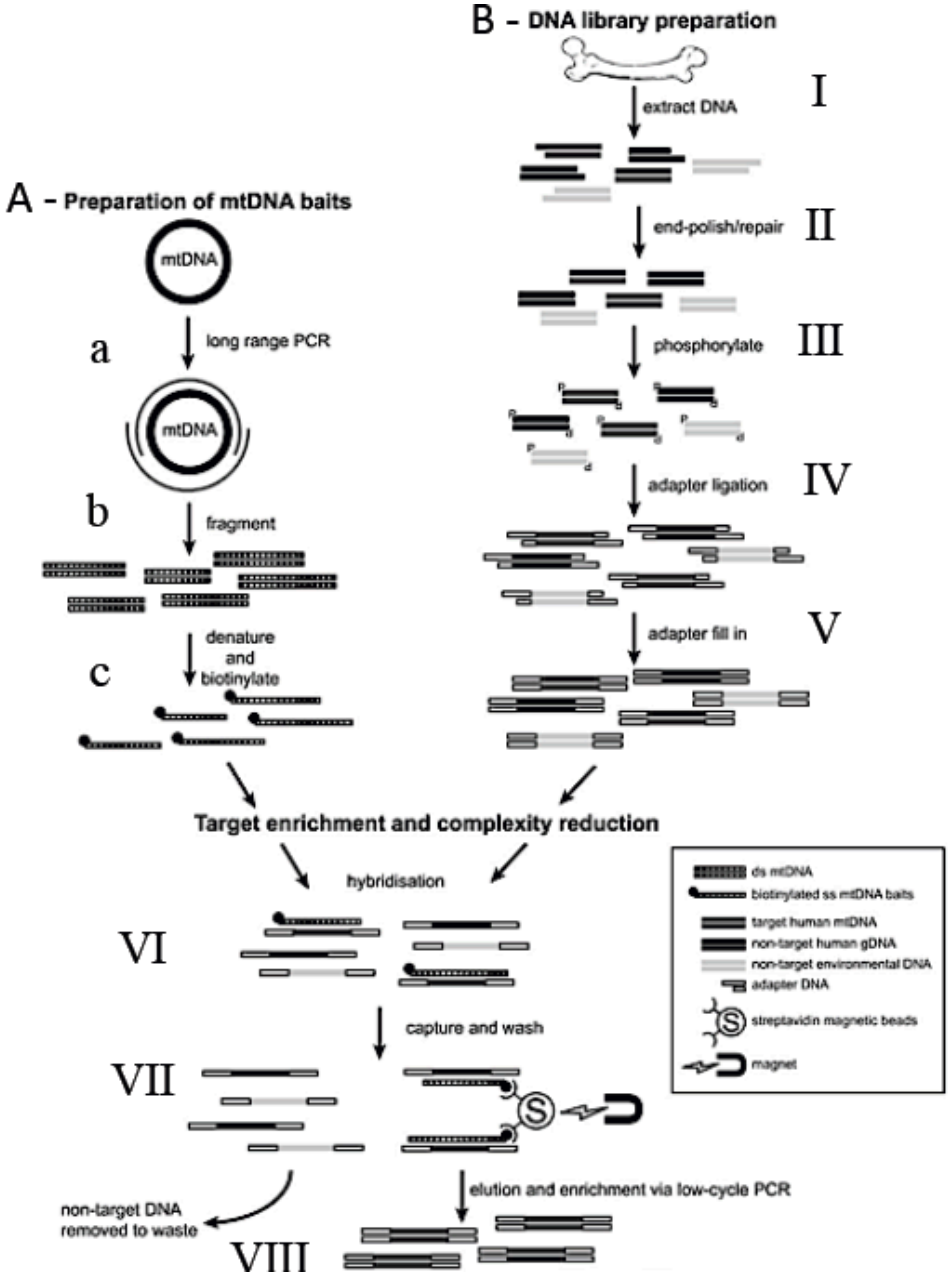


Figure 1.3.3 Example of a mtDNA hybridisation enrichment protocol. **A**. MtDNA baits are prepared using complete mitochondrial genomes from extant taxa. The products from the mtDNA long-range PCR (a) are fragmented by physical shearing (b) to create short fragments of the desirable size (e.g. size range from 100 bp to 500 bp). The ends of the sheared DNA are then biotinylated (c). **B**. To prepare a DNA library, DNA is first extracted (I). DNA damage leaves 5' and 3' overhangs. T4 DNA polymerase is used to polish the DNA (II) by creating blunt ends and T4 PNK phosphorylated 5' ends (III), which is required for adaptor ligation. T4 ligase attached universal hybridization adaptors (uni-hyb A and uni-hyb B) to the phosphorylated ends (IV). Adaptor fill (V) in is then required to create double-stranded adaptors (through the use of deoxyribonucleotide triphosphates – dNTPs). Adaptor complementary primers and Taq polymerase is required to amplify the entire library and immortalise the sample. Single-stranded probe DNA is mixed with single-stranded DNA and left to hybridise over night (VI - in the presence of blocking oligos). Biotinylated probe and bound library DNA are fixed to streptavidin beads on a magnetic rack, and non-specific or weakly bound library DNA is washed away through a series of washes (by increasing temperature and decreasing salt concentration) from the library-probe-streptavidin interaction (VII). The single-stranded library is then converted to double stranded DNA and eluted from the probe-streptavidin interaction (VIII). Eluted library DNA is enriched through low cycle PCR. Library DNA can now be prepared for high-throughput sequencing (taken from open access journal; Templeton *et al.*, 2013 with modifications).

1.4 APPLICATIONS OF aDNA

Ancient DNA analyses of historical remains can be utilised by a wide array of fields including conservation, phylogenetics, evolutionary history, archaeology, palaeontology and forensics (Herrmann *et al.*, 1994; Alonso *et al.*, 2003). In relevance to the scope of this thesis research, the utility of aDNA in conservation and evolutionary histories is explored and will form the focus of the subsequent sections.

1.4.1 UTILITY OF aDNA IN CONSERVATION

The recent advances in sequencing methods and technology as described in section 1.3 have opened up even greater prospects in the field of aDNA, particularly within the conservation field. At first, the practical applications of aDNA in conservation genetics were not widely recognised. Today, with a greater understanding of what time-stamped data can offer, means aDNA is becoming more widely applied. The possibility of observing changes in genetic diversity through time (Shapiro *et al.*, 2014) is one aspect that could not have been easily achieved using 'modern' data alone. The incorporation of time-stamped aDNA data (i.e. aDNA obtained from dated fossils) provides direct insights into a population's past diversity (Ramakrishnan *et al.*, 2009). The demographic history of bison (Shapiro *et al.*, 2004), musk ox, (Campos *et al.*, 2010) and other megafauna (Lorenzen *et al.*, 2011) demonstrate the value of using aDNA to explore demography and evolutionary history. However, the use of aDNA data to make informed conservation decisions is rare, despite the fact that it can provide vital information regarding past population demography (e.g. bottlenecks), former connectivity, and the extent of biodiversity loss (see Leonard, 2008 for a review). Genomic analysis of ancient samples can establish baseline levels of genetic parameters in ancestral populations before demographic declines. Therefore, information on the timing of population fragmentation and how this is related to past changes in the environment (e.g., anthropogenic impact, climate change) can provide valuable insight into current processes influencing population viability (Paplinska *et al.*, 2011; Shafer *et al.*, 2015).

Sample quality and quantity are often direct concerns in the field of conservation biology, since sample collection and quality are become increasingly difficult when the species of interest is rare or elusive in the wild. Conservation biologists often

need to resort to museum samples (Casas-Marce *et al.*, 2009) or non-invasive sampling (Fernandes *et al.*, 2007). Such samples are being utilised more frequently within conservation-related studies as they harness the power of HTS technologies (Murray *et al.*, 2012; Russello *et al.*, 2015). A common factor that is now being observed in many aDNA studies is that they have largely been based on mtDNA sequencing. In some conservation scenarios it may be important to integrate aDNA information to provide a direct temporal context. For example, a study by Pacioni *et al*. (2015) recently looked at the critically endangered woylie (brush-tailed bettong) and found that the genetic connectivity in woylie was detected between fragmented populations, but the level of past connectivity was more accurately evaluated when including direct measures from the past, through the use of aDNA data. Furthermore, expanding aDNA studies beyond the use of mtDNA in an attempt to explore evolution at a nuclear DNA (nuDNA) level is becoming increasingly important. This can be achieved by investigating microsatellite data, single nucleotide polymorphisms (SNPs), and by targeting specific nuclear loci for phylogenetic resolution (Meredith *et al.*, 2008; Meredith *et al.*, 2010; Westerman *et al.*, 2012). Such approaches can be aided by HTS platforms and are becoming more useful in certain studies; for example, where male dispersal may skew interpretations of diversity, connectivity and phylogeography based on matrilineal DNA (Pacioni *et al.*, 2015).

The applications of aDNA in the conservation field are continually growing, although now with the rapid development of genomic methods, the prospects of utilising aDNA techniques are growing even faster and have already seen researchers move aDNA from genetic to genomic research. Although the question may linger as to whether conservation studies are keeping up with the rapid development of molecular techniques, it is clear that applying genomics to conservation studies is a powerful approach (McMahon *et al.*, 2014).

As an exemplar of how aDNA can be integrated into conservation science, Collins *et al.* (2014) explored extinction and recolonisation patterns in New Zealand's coastal megafauna. This study shows the importance of aDNA data integrated into modern DNA analysis. Immediately after the Polynesian arrival approximately AD 1280, a vast number of vertebrate extinctions took place, of which two endemic mainland species (*Phocarctops* spp. Sea lion and *Megadyptes* spp. Penguins) were completely

wiped out. These species were replaced by a genetically divergent clade from the remote subantarctic region, a process that was carried out within the space of a few centuries. The inclusion of fossil data implies that *Phocarctos* persisted on mainland New Zealand throughout the Late Glacial (ca. 14-11.6 kya) and Holocene (11.6 kya-present), and the extinction-recolonisation event had potentially replaced an endemic taxa with an invasive species (Collins *et al.*, 2014). In such cases, extinction-replacement scenarios are thought to reflect the retraction and/or expansion of lineages, although the genetic data reveals that pre-historic mainland and subantarctic specimens are highly distinct, showing no evidence of genetic turnover. From a conservation perspective, extinction and replacement events like these can dramatically reshape biological communities. Ancient DNA is clearly a suitable means for exploring and providing information about a species past biodiversity. However, the absence of a temporal component, which is often a missing factor when making conservation decisions, can make it difficult to establish meaningful conservation targets and strategies (Pacioni *et al.*, 2011).

1.4.2 UTILITY OF ADNA IN PHYLOGENETICS AND EVOLUTIONARY HISTORY

The use of aDNA to resolve controversy about evolutionary relationships, and to estimate divergence times between species, is becoming more readily used in molecular studies; however, as briefly mentioned above, one of the limitations of aDNA research is that it remains heavily reliant on mitochondrial loci. Using aDNA as a tool in phylogenetic studies may not be utilised to its full potential without having access to information from the nuclear genome (see below for more on nuDNA). Nonetheless, ancient mtDNA can be incredibly informative, especially if areas of the fast evolving regions (such as the control region) are targeted (Lambert *et al.*, 2002; Navascués *et al.*, 2010). This alone can provide a wealth of knowledge about a species' evolutionary history. As a result, one of the earliest uses of aDNA sequences was for phylogenetic studies, where sequences of extinct species were included in molecular phylogenies to test taxonomic relationships that had previously been hypothesised based upon morphological characteristics alone (Foote *et al.*, 2011); this was done for the quagga (*Equus quagga quagga*; Higuchi *et al.,* 1984), marsupial wolf (*Thylacinus cynocephalus;* Thomas *et al.,* 1989), and mammoth (*Mammuthus primigenius;* Hoss *et al.,* 1994).

The use of mtDNA to explore questions about evolutionary history is usually targeted at species-level questions (Meyer *et al.*, 2003). However, with the advancements in sequencing technologies since Sanger sequencing, the use of mtDNA as a phylogenetic marker has become more applicable and helped achieve greater levels of phylogenetic resolution within a taxon. The ability to assemble trees that show the relationships between all of the species found in a given assemblage allows us to answer questions about a species' history. However, when we make use of other molecular markers such as nuclear DNA, we begin to acquire information at a population-level. Molecular phylogenies built using both modern DNA and aDNA sequences provide the most optimal approach in building phylogenies, which can offer insights into a species past biodiversity, patterns of gene flow, inter- and intra-species variations, signals of genetic loss, and other evolutionary history uncertainties proposed about a species. Although aDNA is extremely useful in accompanying any data set, there are important considerations when applying aDNA to phylogenetic or population genetic analyses; for instance, distinguishing between true mutations and artefacts arising from post mortem damage is difficult (Hofreiter *et al.,* 2001). Using appropriate bioinformatic methods for heterochronous data, for example, can overcome this (Navascués *et al.*, 2010; Murray *et al.*, 2015).

Although the utilisation of aDNA in conservation and phylogenetic studies has been positive, there still remains a challenging aspect to it. Extracting DNA from historical specimens within Australia, for example, with climate conditions that are mainly dry and warm, decreases the chances of DNA survival (see previous section 1.2.5 for more information). The southwest region of Western Australia in particular is a biodiversity hotspot, and contains a high concentration of endemic species. This region, however, has also experienced severe biodiversity loss. Using aDNA to explore the loss of species diversity is fundamental in understanding this unique and diverse region. The following section explores some aspects of utilising aDNA studies in Australia, and provides some insight into Australian macropod studies.

1.5 AUSTRALIA AND ITS BIODIVERSITY

Antarctica, the last continent to be connected to Australia, began to separate from Australia as early as 70 million years ago (Johnson, 2006). This isolation has led to

the continent (Australia) supporting large numbers of species found nowhere else in the world (Hanson *et al.*, 2008). Western Australia (WA) in particular, is home to some of the most unique and diverse biota on earth, which can be attributed to the state's size that spans a range of biological diversity, variable soil conditions, and climatic zones (Stefen *et al.*, 2009). Biodiversity is classified at three levels: genetic diversity, species diversity, and ecosystems diversity (Frankham *et al.*, 2002). Because WA has a high concentration of endemic species at all of these levels, it is formally recognised as a global biodiversity hotspot (Mayr, 2000). Although there still appears to be a wealth of biodiversity in Australia today, over the past 200 years, Australia has experienced more mammal extinctions than any other country in the world (Johnson, 2006). This has resulted in a severe loss of species biodiversity, as well as a decline in genetic diversity.

1.5.1 FOSSILS – SERVING AS GENETIC TIME CAPSULES

The ability to capture genetic information from species that no longer exist is achievable. This can be done by utilising historical samples, through aDNA analysis, which provide the ability to explore "lost" genetic signatures in a population. An important insight into how ecosystems functioned prior to the arrival of European settlement is a common example of what historical information from fossils can provide. Also, genetic data of past biodiversity may also assist in developing scientifically sound conservation management responses and may even provide missing signals within a population to be able to piece together or change our knowledge of macropod phylogeography, for example. Attaining this type of information from historical species does not come without a challenge, and is especially problematic in harsh (hot and dry) climates, a common characteristic of most Australian habitats. In Australia, the superfamily Macropodoidea has some well-documented information based on living taxa (Cardillo *et al.*, 2004; Phillips *et al.*, 2013), although the genetic characterisation of extinct macropod taxa remains unexplored.

1.5.2 AUSTRALIA'S CLIMATE CHALLENGES AND aDNA APPLICATIONS IN MACROPOD STUDIES

Morphology-based studies have dominated the majority of mammal research undertaken in Australia. Prideaux and Warburton (2010), for example, used a number of kangaroos and wallabies in an osteology-based approach to explore species evolution and provide a phylogeny of the Macropodidae. Another study also applied a morphology-based approach by looking at the morphometric variation of bones among populations of the mountain brushtail possum (Lindenmayer *et al.*, 1995). To complement current morphology based studies, provide temporal analyses, evolutionary and ecological information, DNA work is required (Lindenmayer *et al.*, 1995).

Previous modern DNA studies have seen the completion of mitochondrial genomes of some of Australia's unique marsupials, for example; the fat-tailed dunnart (*Sminthopsis crassicaudata*)(Janke *et al.*, 1994), Tasmanian devil (*Sarchophilus harrisii*) (Miller *et al.*, 2011) and the tammar wallaby (*Macropus eugenii*) (Renfree *et al.*, 2011); these are all extant species, and to date, there have been no systematic attempts to add mitochondrial genomes of extinct Australian taxa to the list.

Historical samples used for molecular analysis have largely been collected from curated and preserved museum specimens (Paplinska *et al.*, 2011), or from modern environmental samples (Tringe S *et al.*, 2005), or both (Macqueen *et al.*, 2011). The following table (Table 1.5.1) briefly summarises aDNA studies that have successfully produced molecular information from archaeological sites within Australia. The utilisation of old museum specimens for macropod research in Australia is important, as the number of mammals declining over time is increasing. This has been a direct result of habitat destruction and fragmentation, amongst other leading factors, which are discussed later on in this thesis research. Having access to genetic information of endangered species persisting in these fragmented habitats, for example, develops an understanding of the historical changes and becomes a crucial part of future conservation decisions.

Table 1.5.1 Chronological overview of ancient DNA research in Australia showing main findings from studies. Those marked with an * represent additions to the literature after the start date of this thesis research

REFERENCE	TITLE	TARGET DNA	MAIN FINDINGS
Thomas *et al.*, 1989	DNA phylogeny of the extinct marsupial wolf	mtDNA (12S rRNA)	One of the earliest ancient DNA studies in an attempt to ascertain the thylacine's phylogenetic position based on 94 bp of 12S rRNA. Results were inconclusive.
Krajewski *et al.*, 1992	Phylogenetic relationships of the thylacine (Marsupialia: Thylacinidae) among dasyuroid marsupials: evidence from cytochrome *b* DNA sequences	mtDNA (Cytochrome *b*)	Partial mitochondrial cytochrome *b* sequences showed that dasyurids are monophyletic with respect to *Thylacinus*.
Krajewski *et al.*, 1997	DNA phylogeny of the marsupial wolf resolved	mtDNA (Cytochrome *b* and 12S rRNA) + nuDNA (protamine gene)	A mtDNA (1546 bp) and nuDNA (841 bp) reanalysis shows that thylacines are members of the Dasyuromorphia, and suggest a late Oligocene or very early Miocene divergence.
Oskam *et al.*, 2010	Fossil avian eggshell preserved ancient DNA	mtDNA (12S rRNA) + nuDNA	aDNA from eggshell was sourced from Australia among other countries and revealed to contain DNA up to 19 ka old.
Heupink *et al.*, 2011*	Ancient DNA suggest Dwarf and 'Giant' Emu are conspecific	mtDNA (d-loop and COXI) + nuDNA (melanocortin 1 receptor gene)	Evolutionary relationships between the King Island Emu and modern Emu share a recent common ancestor and possibly underwent dwarfism as a result of phenotypic plasticity.
Haouchar *et al.*, 2012*	The identity of the Depuch Island Rock-wallaby revealed through ancient DNA	mtDNA (control region and cytochrome *b*)	Using ancient mtDNA *Petrogale lateralis lateralis* was identified as the most likely species on Depuch Island, among the other mainland candidates. This provides essential information for conservation aspects.
Murray *et al.*, 2012*	High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens	mtDNA (12S and 16S rRNA) + chloroplast DNA	The application of high-throughput sequencing approaches to profile the biotic remains preserved in midden material in Australia uncover ancient plant and animal DNA.
Haouchar *et al.*, 2013*	Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late pleistocene-holocene cave deposit on Kangaroo Island, South Australia	mtDNA (12S, 16S rRNA, Cyt *b*) + Chloroplast DNA	Kelly Hill Cave (Kangaroo Island, South Australia) has excellent DNA preservation revealing an array of plant and animal species, some not previously known to entirely exist on the Island such as the red kangaroo.

Murray *et al.*, 2013*	A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages	mtDNA (12S and 16S rRNA)	A novel bulk bone method was used to characterise numerous amounts of morphologically indistinct fossil bones, quickly and cost effectively. Some interesting species like tree kangaroo were revealed from the site in southwest Australia.
Grealy *et al.*, 2015*	A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages	mtDNA (12S and 16S rRNA	The novel bulk bone metabarcoding method is critically analysed, using fossil deposits from two sites, Kangaroo Island, Australia and Canterbury, New Zealand. Results from the molecular and morphological methods used showed variances between certain steps. The bulk bone method requires less sampling than traditional methods.
Llamas *et al.*, 2015*	Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies	mtDNA shotgun sequencing	Using a cross-species DNA capture method on fossils from high latitude, high altitude caves in Tasmania. The retrieval of aDNA from two megafaunal kangaroo species dated > 40 ka successfully yielded sequence information. Phylogenetic evidence confirm three primary lineages: Sthenurines, macropodines and banded hare-wallaby
Pacioni *et al.*, 2015*	Genetic diversity loss in a biodiversity hotspot: ancient NDA quantifies genetic decline and former connectivity in a critically endangered marsupial	mtDNA (12s rRNA control region) + nuDNA	The extent of genetic diversity loss and former connectivity between fragmented populations of the Western Australian woylie was explored using both mtDNA and nuDNA from modern and historical samples. Results show a high degree of gene flow in the woylies historical range and a both mtDNA and nuDNA found evidence of genetic diversity loss.
Grealy *et al.*, 2016*	An assessment of ancient DNA preservation in Holocene-Pleistocene fossil bone excavated from the world heritage Naracoorte Caves, South Australia	mtDNA (12S r RNA) + shotgun sequencing	The first aDNA study into the fossil site deposits of the Naracoorte Caves (South Australia). Ancient DNA from diverse taxa was retrieved using a bulk bone method. Bone as old as 18 600 ka, was found to be successful in the amplification of DNA. Overall, the DNA is highly degraded and contains low proportions of endogenous DNA.

1.6 USING ANCIENT DNA TO INVESTIGATE EXTINCTION, EXTIRPATION, AND PAST BIODIVERSITY OF AUSTRALIAN MACROPODS: AIMS AND SCOPE OF THESIS

The ultimate aim of this thesis research is to use a combination of historical, Holocene and Pleistocene fossils in an attempt to explore what ancient mtDNA can add to our knowledge of Australia's macropods and their ecosystems. Due to the rapid advancements of molecular biology techniques in recent years, acquiring information about a species past has become more accessible with developments made in post-DNA-extraction techniques, the design and optimisation of short-amplicon DNA capture and enrichment, and lastly the advances in HTS platforms and bioinformatic analysis programs. Utilising these latest tools, the extinction, extirpation, and past biodiversity of Australian macropods was investigated, despite the challenging DNA preservation conditions present on the Australian mainland.

Due to the nature of this research, each chapter has specific questions and ideas that are addressed independently and will be highlighted in the front of the introduction to each chapter. Below is a brief summary of the chapters' focus and aims used to accomplish them.

Chapter Two aims to identify the species of rock wallaby that was previously found on Depuch Island before they became completely extirpated. A Sanger sequencing and DNA cloning approach was used to produce nucleotide information from two mtDNA genes. A mtDNA phylogeny was used to resolve the relationships between the island and mainland species.

Chapter Three investigates a geographically wide-ranging macropod species complex, to explore questions regarding interrelationships and former distribution of the brush-tailed bettongs: *Bettongia gaimardi, B. tropica* and *B. penicillata.* The molecular techniques used in this chapter include DNA extraction and mtDNA sequencing of two gene regions and placed within a mtDNA phylogeny with previously published, modern bettong sequence data.

Chapter Four explores the Kelly Hill Cave archeological site on Kangaroo Island and investigates the aDNA preservation from ~70 late Pleistocene-Holocene fossil bones

and 20 sediment core samples. One of the focuses of this study was to use a molecular approach utilising advanced HTS technology and amplicon sequencing methods to compliment the morphology-based studies.

Chapter Five examines four cave sites around Australia, in an attempt to assess preservation at each site. The latest molecular techniques were utilised in capturing and enriching ultra-short, highly fragmented aDNA in paleontological samples. Using a combination of amplicon and shotgun sequencing, sites were assessed on their ability to preserve DNA and the potential for DNA survival into the Late Pleistocene.

In summary, this thesis presents a combination of work from multiple sites across Australia using a range of aDNA techniques and sequencing technologies that have evolved over the course of this thesis. Collectively, this body of work has demonstrated the value of integrating aDNA data into modern-day conservation decision-making and aided in a wider understanding of Australian macropods both past and present.

1.7 REFERENCES

Albert, T. J., R. A. Gibbs, C. M. Middle, M. N. Molla, D. M. Muzny, L. Nazareth, C. J. Packard, T. A. Richmond, M. J. Rodesch, X. Song, G. M. Weinstock & D. Wheeler (2007). Direct selection of human genomic loci by microarray hybridization. *Nature Methods* **4**, 903-905.

Allentoft, M. E., M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, T. M. P. Gilbert, E. Willerslev, G. Zhang, P. Scofield, R. N. Hodaway & M. Bunce (2012). The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of Royal Society B-Biological Sciences* **279**, 4724-4733.

Almeida, M., E. Betancor, R. Fregel, N. Suarez & J. Pestano (2011). Efficient DNA extraction from hair shafts. *Forensic Science International: Genetics Supplement Series* **3**, 319-325.

Alonso, A., P. Martín, C. Albarrán, P. García, D. Primorac, O. García, L. F. de Simón, J. García-Hirschfeld, M. Sancho & J. Fernández-Piqueras (2003). Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. *Croatian Medical Journal* **44**, 273-280.

Axelsson, E., E. Willerslev, T. M. P. Gilbert & R. Nielsen (2008). The effect of Ancient DNA damage on inferences of demographic histories. *Molecular Biology and Evolution* **25**, 2181-2187.

Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes & H. R. Bignell (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53-59.

Campos, P. F., O. E. Craig, G. Turner-Walker, W. Peacock, E. Willerslev & M. T. P. Gilbert (2012). DNA in ancient bone - Where is it and how should we extract it? *Annals of Anatomy* **194**, 7-16.

Campos, P. F., E. Willerslev, A. Sher, L. Orlando, E. Axelsson, A. Tikhonov, K. Aaris-Sørensen, A. D. Greenwood, R.-D. Kahlke, P. Kosintsev, T. Krakhmalnaya, T. Kuznetsova, P. Lemey, R. MacPhee, C. A. Norris, K. Shepherd, M. A. Suchard, G. D. Zazula, B. Shapiro & M. T. P. Gilbert (2010). Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (Ovibos moschatus) population dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 5675-5680.

Cardillo, M., O. R. Bininda-Edmonds, E. Boakes & A. Purvis (2004). A species-level phylogenetic supertree of marsupials. *The Zoological Society of London* **264**, 11-31.

Carpenter, M. L., J. D. Buenrostro, C. Valdiosera, H. Schroeder, M. E. Allentoft, M. Sikora, M. Rasmussen, S. Gravel, S. Guillén & G. Nekhrizov (2013). Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *The American Journal of Human Genetics* **93**, 852-864.

Casas-Marce, M., E. Revilla & J. A. Godoy (2009). Searching for DNA in museum specimens: a comparison of sources in a mammal species. *Molecular Ecology Resources* **10**, 502-507

Collins, C. J., N. J. Rawlence, S. Prost, C. N. K. Anderson, M. Knapp, R. P. Scofield, B. C. Robertson, I. Smith, E. A. Matisoo-Smith, B. L. Chilvers & J. M. Waters (2014). Extinction and recolonization of coastal megafauna following human arrival in New Zealand. *Proceedings of the Royal Society of London B: Biological Sciences* **281**, 20140097.

Collins, M. & D. Harker (2009). thermal-age.eu. Available from http://thermal-age.eu/).

Dabney, J., M. Knapp, I. Glocke, M. T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. Garcia, S. Paabo, J. Arsuaga & M. Meyer (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15758-15763.

Deguilloux, M. F., S. Ricaud, R. Leahy & M. H. Pemonge (2011). Analysis of ancient human DNA and primer contamination: One step backward one step forward. *Forensic Science International* **210**, 102-109.

Fernandes, C. A., C. Ginja, I. Pereira, R. Tenreiro, M. W. Bruford & M. Santos-Reis (2007). Species-specific mitochondrial DNA markers for identification of non-invasive samples from sympatric carnivores in the Iberian Peninsula. *Conservation Genetics* **9**, 681-690.

Foote, A. D., M. Hofreiter & P. A. Morin (2011). Ancient DNA from marine mammals: studying long-lived species over ecological and evolutionary timescales. *Annals of Anatomy* **194**, 112-120.

Frankham, R., J. Ballou & D. A. Briscoe (2002). *Introduction to conservation genetics,* England, Cambridge University Press.

Gansauge, M.-T. & M. Meyer (2013). Single-stranded DNA library preparation for the sequencing of ancinet or damaged DNA. *Nature Protocols* **8**, 737-748.

Gilbert, M. T. P., H. J. Bandelt, M. Hofreiter & I. Barnes (2005). Assessing ancient DNA studies. *Trends in Ecology & Evolution* **20**, 541-544.

Gilbert, M. T. P., D. I. Drautz, A. M. Lesk, S. Y. W. Ho, J. Qi, A. Ratan, C. H. Hsu, A. Sher, L. Dalén & A. Götherström (2008). Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8327-8332.

Glenn, T. C. (2011). Field guide to next-generation DNA sequencers. *Molecular Ecology Resources* **5**, 759-769.

Gnirke, A., A. Melnikov, J. Maguire, P. Rogov, E. M. LeProust, W. Brockman, T. Fennell, G. Giannoukos, S. Fisher & C. Russ (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology* **27**, 182-189.

Grealy, A., A. Macken, M. E. Allentoft, N. J. Rawlence, E. Reed & M. Bunce (2016). An assessment of ancient DNA preservation in Holocene–Pleistocene fossil bone excavated from the world heritage Naracoorte Caves, South Australia. *Journal of Quaternary Science* **31**, 33-45.

Grealy, A. C., M. C. McDowell, P. Scofield, D. i. C. Murray, D. A. Fusco, J. Haile, G. J. Prideaux & M. Bunce (2015). A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages. *Quaternary Science Reviews* **128**, 37-47.

Green, R. E., J. Krause, S. E. Ptak, A. W. Briggs, M. T. Ronan, J. F. Simons, L. Du, M. Egholm, J. M. Rothberg & M. Paunovic (2006). Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330-336.

Green, R. E., A. S. Malaspinas, J. Krause, A. W. Briggs, P. L. F. Johnson, C. Uhler, M. Meyer, J. M. Good, T. Maricic & U. Stenzel (2008). A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**, 416-426.

Greenwood, A. D., C. Capelli, G. Possnert & S. Paabo (1999). Nuclear DNA sequences from late Pleistocene megafauna. *Molecular Biology and Evolution* **16**, 1466-1473.

Haile, J., R. Holdaway, K. Oliver, M. Bunce, M. T. P. Gilbert, R. Nielsen, K. Munch, S. Y. W. Ho, B. Shapiro & E. Willerslev (2007). Ancient DNA chronology within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor? *Molecular Biology and Evolution* **24**, 982-989.

Hanson, T., T. M. Brooks, G. A. B. D. Fonseca, M. Hoffmann, J. F. Lamoreux, G. Machlis, C. G. Mittermeier, R. A. Mittermeier & J. D. Pilgrim (2008). warfare in biodiversity hotspots. *Conservation Biology* **23**, 578-587.

Haouchar, D., J. Haile, M. C. McDowell, D. C. Murray, N. E. White, R. J. N. Allcock, M. J. Phillips, G. J. Prideaux & M. Bunce (2013). Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late Pleistocene-Holocene cave deposit on Kangaroo Island, South Australia. *Quaternary Science Reviews* **84**, 56-64.

Haouchar, D., J. Haile, P. B. S. Spencer & M. Bunce (2012). The identity of the Depuch Island rock-wallaby revealed through ancient DNA. *Australian Mammalogy* **35**, 101-106.

Herrmann, B. & S. Hummel (1994). *Ancient DNA: recovery and analysis of genetic material from paleontological, archaeological, museum, medical, and forensic specimens,* New York, Springer-Verlag.

Heupink, T. H., L. Huynen & D. M. Lambert (2011). Ancient DNA suggests Dwarf and 'Giant' Emu are conspecific *Plos ONE* **6**, e18728.

Higgins, D., A. B. Rohrlach, J. Kaidonis, G. Townsend & J. J. Austin (2015). Differential nuclear and mitochondrial DNA preservation in post-mortem teeth with implications for forensic and ancient DNA studies. *Plos ONE* **10**, e0126935.

Higuchi, R., B. Bowman, M. Freiberger, O. A. Ryder & A. C. Wilson (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature* **312**, 282-284.

Ho, S. Y., T. H. Heupink, A. Rambaut & B. Shapiro (2007). Bayesian estimation of sequence damage in ancient DNA. *Molecular Biology and Evolution* **24**, 1416-1422.

Hofreiter, M., H. N. Poinar, W. G. Spaulding, K. Bauer, P. S. Martin, G. Possnert & S. Paabo (2000). A molecular analysis of ground sloth diet through the last glaciation. *Molecular Ecology* **9**, 1975-1984.

Hofreiter, M., D. Serre, H. N. Poinar, M. Kuch & S. Pääbo (2001). Ancient DNA. *Nature Reviews Genetics* **2**, 353-359.

Hofreiter, M., D. Serre, N. Rohland, G. Rabeder, D. Nagel, N. Conard, S. Munzel & S. Paabo (2004). Lack of phylogeography in European mammals before the last glaciation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12963-12968.

Horn, S. (2012). Target Enrichment via Hybridization Capture. *Ancient DNA: Methods and Protocols*. (ed. by B. Shapiro and M. Hofreiter), pp. 177 -188. Springer Science.

Janke, A., G. Feldmaier-Fuchs, W. K. Thomas, A. Von-Haeseler & S. Paabo (1994). The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* **137**, 243-256.

Johnson, C. (2006). *Australia's Mammal Extinctions: a 50000 year history*, America; Cambridge University Press.

Kircher, M. (2012). Analysis of high-throughput ancient DNA sequencing data. *Ancient DNA: Methods and Protocols*. (ed. by B. Shapiro and M. Hofreiter), pp. 197 - 228. Springer Science

Knapp, M. & M. Hofreiter (2010). Next generation sequencing of ancient DNA: requirements, strategies and perspectives. *Genes* **1**, 227-243.

Krajewski, C., L. Buckley & M. Westerman (1997). DNA phylogeny of the marsupial wolf resolved. *Proceedings of Royal Society of London B* **264**, 911-917.

Krajewski, C., A. C. Driskell, P. R. Baverstock & M. J. Braun (1992). Phylogenetic relationships of the thylacine (Marsupialia: Thylacinidae) among dasyuroid marsupials: evidence from cytochrome *b* DNA sequences. *Proceedings of Royal Society of London B* **250**, 19-27.

Lambert, D. M., P. A. Ritchie & C. D. Millar (2002). Rates of evolution in ancient DNA from Adelie penguins. *Science* **295**, 2270-2273.

Leonard, J. A. (2008). Ancient DNA applications for wildlife conservation. *Molecular Ecology* **17**, 4186-4196.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715.

Lindenmayer, D., K. Viggers, R. Cunningham & C. Donnelly (1995). Morphological Variation Among Populations of the Mountain Brushtail Possum, Trichosurus-Caninus Ogilby (Phalangeridae, Marsupialia). *Australian Journal of Zoology* **43**, 449-458.

Liu, L., Y. Li, S. Li, Y. Hu, R. He, D. Pong, L. L. Lin & M. Law (2012). Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology* **2012**, 251364.

Llamas, B., P. Brotherton, K. J. Mitchell, J. E. Templeton, V. A. Thomson, J. L. Metcalf, K. N. Armstrong, M. Kasper, S. M. Richards, A. B. Camens, L. S. Y. Michael & A. Cooper (2015). Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies. *Molecular Biology and Evolution* **32**, 574-584.

Lorenzen, E. D., D. Nogues-Bravo, L. Orlando, J. Weinstock, J. Binladen, K. A. Marske, A. Ugan, M. K. Borregaard, M. T. P. Gilbert, R. Nielsen, S. Y. W. Ho, T. Goebel, K. E. Graf, D. Byers, J. T. Stenderup, M. Rasmussen, P. F. Campos, J. A. Leonard, K.-P. Koepfli, D. Froese, G. Zazula, T. W. Stafford, K. Aaris-Sorensen, P. Batra, A. M. Haywood, J. S. Singarayer, P. J. Valdes, G. Boeskorov, J. A. Burns, S. P. Davydov, J. Haile, D. L. Jenkins, P. Kosintsev, T. Kuznetsova, X. Lai, L. D. Martin, H. G. McDonald, D. Mol, M.

Meldgaard, K. Munch, E. Stephan, M. Sablin, R. S. Sommer, T. Sipko, E. Scott, M. A. Suchard, A. Tikhonov, R. Willerslev, R. K. Wayne, A. Cooper, M. Hofreiter, A. Sher, B. Shapiro, C. Rahbek & E. Willerslev (2011). Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359-364.

Macqueen, P., J. M. Seddon & A. W. Goldizen (2011). Effects of historical forest contraction on the phylogeographic structure of Australo-Papuan populations of the red-legged pademelon (Macropodidae: Thylogale stigmatica). *Austral Ecology* **37**, 479-490.

Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen & Z. Chen (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380.

Maricic, T., M. Whitten & S. Paabo (2010). Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *Plos ONE* **5**, e14004.

Mayr, E. (2000). The biological species concept. *Species concepts and phylogenetic theory: a debate. Columbia University Press, New York*, 17-29.

McMahon, B. J., E. C. Teeling & J. Hoglund (2014). How and why should we implement genomics into conservation? *Evolutionary Applications* **7**, 999-1007.

Men, A. E., P. Wilson, K. Siemering & S. Forrest (2008). *Sanger DNA Sequencing,* Weinheim, Wiley.

Meredith, R. W., M. A. Mendoza, K. K. Roberts, M. Westerman & M. S. Springer (2010). A phylogeny and timescale for the evolution of Pseudocheiridae (Marsupialia: Diprotodontia) in Australia and New Guinea. *Journal of Mammalian Evolution* **17**, 75-99.

Meredith, R. W., M. Westerman & M. S. Springer (2008). A timescale and phylogeny for "Bandicoots" (Peramelemorphia: Marsupialia) based on sequences for five nuclear genes. *Molecular Phylogenetics and Evolution* **47**, 1-20.

Meyer, A. & F. Zardoya (2003). Recent advances in the Molecular phylogeny of vertebrates. *Annual Reviews of Ecology and Evolutionary Systematics* **34**, 311-338.

Miller, W., D. I. Drautz, A. Ratan, B. Pusey, J. Qi, A. M. Lesk, L. P. Tomsho, M. D. Packard, F. Zhao & A. Sher (2008). Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* **456**, 387-390.

Miller, W., V. M. Hayes, A. Ratan, D. C. Petersen, N. E. Wittekindt, J. Miller, B. Walenz, J. Knight, J. Qi, F. Zhao, Q. Wang, O. C. Bedoya-Reina, N. Katiyar, L. P. Tomsho, L. M. Kasson, R.-A. Hardie, P. Woodbridge, E. A. Tindall, M. F. Bertelsen, D. Dixon, S. Pyecroft, K. M. Helgen, A. M. Lesk, T. H. Pringle, N. Patterson, Y. Zhang, A. Kreiss, G. M. Woods, M. E. Jones & S. C. Schuster (2011). Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian devil). *Proceedings of the*

National Academy of Sciences of the United States of America **108**, 12348-12353.

Mourier, T., S. Y. W. Ho, M. T. P. Gilbert, E. Willerslev & L. Orlando (2012). Statistical guidelines for detecting past population shifts using ancient DNA. *Molecular Biology and Evolution* **94**, 8-21.

Murray, D. C., M. L. Coghlan & M. Bunce (2015). From benchtop to desktop: Important considerations when designing amplicon sequencing workflows. *PloS ONE* **4**, e10.

Murray, D. C., J. Haile, J. Dortch, N. White, D. Haouchar, M. I. Bellgard, R. Allcock, J, G. J. Prideaux & M. Bunce (2013). Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports.* **3**, 3371.

Murray, D. C., S. G. Pearson, F. Richard, B. M. Chase, J. Houston, J. Atchison, N. E. White, M. I. Bellgard, E. Clarke, M. Macphail, T. M. P. Gilbert, J. Haile & M. Bunce (2012). High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens. *Quaternary Science Reviews* **58**, 135-145.

Navascués, M., F. Depaulis & B. C. Emerson (2010). Combining contemporary and ancient DNA in population genetic and phylogeographical studies. *Molecular Ecology Resources* **10**, 760-772.

Nogues-Bravo, D., J. Rodriguez, J. Hortal, P. Batra & M. B. Arauo (2008). Climate change, humans, and the extinction of the woolly mammoth. *PLoS Biology* **6**, 79.

Orlando, L., A. l. Ginolhac, G. Zhang, D. Froese, A. Albrechtsen, M. Stiller, M. Schubert, E. Cappellini, B. Petersen & I. Moltke (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74-78.

Orlando, L., J. A. Leonard, A. Thenot, V. Laudet, C. Guerin & C. Hanni (2003). Ancient DNA analysis reveals woolly rhino evolutionary relationships. *Molecular Phylogenetics and Evolution* **28**, 485-499.

Oskam, C. L., J. Haile, E. McLay, P. Rigby, M. E. Allentoft, M. E. Olsen, C. Bengtsson, G. H. Miller, J. L. Schwenninger & C. Jacomb (2010). Fossil avian eggshell preserves ancient DNA. *Proceedings of the Royal Society B: Biological Sciences* **277**, 1991-2000.

Pääbo, S., H. Poinar, D. Serre, V. Jaenicke-Després, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant & M. Hofreiter (2004). Genetic analyses from ancient DNA. *Annual Review of Genetics* **38**, 645-679.

Pacioni, C., H. Hunt, M. E. Allentoft, T. G. Vaughan, A. F. Wayne, A. Baynes, D. Haouchar, J. Dortch & M. Bunce (2015). Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. *Molecular Ecology* **24**, 5813-5828.

Pacioni, C., A. F. Wayne & P. B. S. Spencer (2011). Effects of habitat fragmentation on population structure and long-distance gene flow in an endangered marsupial: the woylie. *Journal of Zoology* **283**, 98-107.

Paplinska, J. Z., D. A. Taggart, T. Corrigan, M. D. B. Eldridge & J. J. Austin (2011). Using DNA from museum specimens to preserve the integrity of evolutionarily significant unit boundaries in threatened species. *Biological Conservation* **144**, 290-297.

Penney, D., C. Wadsworth, G. Fox, S. L. Kennedy, R. F. Preziosi & T. A. Brown (2013). Absence of ancient DNA in sub-fossil insect inclusions preserved in 'Anthropocene' Colombian Copal. *PloS ONE* **8**, 73150-73156.

Phillips, M. J., D. Haouchar, R. C. Pratt, G. C. Gibb & M. Bunce (2013). Inferring kangaroo phylogeny from incongruent nuclear and mitochondrial genes. *PloS ONE* **8**, e57745.

Poinar, H. N., C. Schwarz, J. Qi, B. Shapiro, R. D. E. MacPhee, B. Buigues, A. Tikhonov, D. H. Huson, L. P. Tomsho & A. Auch (2006). Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* **311**, 392-394.

Prideaux, G. J. (2004). *Systematics and evolution of the sthenurine kangaroo,* Berkeley and Los Angeles, California, University of California Press.

Prideaux, G. J. & N. M. Warburton (2010). An osteology-based appraisal of the phylogeny and evolution of kangaroos and wallabies (Macropodidae: Marsupialia). *Zoological Journal of the Linnean Society* **159**, 954-987.

Quail, M. A., M. Smith, P. Coupland, T. D. Otto, S. R. Harris, T. R. Connor, A. Bertoni, H. P. Swerdlow & Y. Gu (2012). A tale of three next genreation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13**, 341.

Ramakrishnan, U. M. A. & E. A. Hadly (2009). Using phylochronology to reveal cryptic population histories: review and synthesis of 29 ancient DNA studies. *Molecular Ecology* **18**, 1310-1330.

Rasmussen, M., Y. Li, S. Lindgreen, J. S. Pedersen, A. Albrechtsen, I. Moltke, M. Metspalu, E. Metspalu, T. Kivisild & R. Gupta (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757-762.

Renfree, M., A. Papenfuss, J. Deakin, J. Lindsay, T. Heider, K. Belov, W. Rens, P. Waters, E. Pharo, G. Shaw, E. Wong, C. Lefevre, K. Nicholas, Y. Kuroki, M. Wakefield, K. Zenger, C. Wang, M. Ferguson-Smith, F. Nicholas, D. Hickford, H. Yu, K. Short, H. Siddle, S. Frankenberg, K. Chew, B. Menzies, J. Stringer, S. Suzuki, T. Hore, M. Delbridge, A. Mohammadi, N. Schneider, Y. Hu, W. O'Hara, S. Al Nadaf, C. Wu, Z.-P. Feng, B. Cocks, J. Wang, P. Flicek, S. Searle, S. Fairley, K. Beal, J. Herrero, D. Carone, Y. Suzuki, S. Sugano, A. Toyoda, Y. Sakaki, S. Kondo, Y. Nishida, S. Tatsumoto, I. Mandiou, A. Hsu, K. McColl, B. Lansdell, G. Weinstock, E. Kuczek, A. McGrath, P. Wilson, A. Men, M. Hazar-Rethinam, A. Hall, J. Davis, D. Wood, S. Williams, Y. Sundaravadanam, D. Muzny, S. Jhangiani, L. Lewis, M. Morgan, G. Okwuonu, S. Ruiz, J. Santibanez, L. Nazareth, A. Cree, G.

Fowler, C. Kovar, H. Dinh, V. Joshi, C. Jing, F. Lara, R. Thornton, L. Chen, J. Deng, Y. Liu, J. Shen, X.-Z. Song, J. Edson, C. Troon, D. Thomas, A. Stephens, L. Yapa, T. Levchenko, R. Gibbs, D. Cooper, T. Speed, A. Fujiyama, J. M Graves, R. O'Neill, A. Pask, S. Forrest & K. Worley (2011). Genome sequence of an Australian kangaroo, Macropus eugenii, provides insight into the evolution of mammalian reproduction and development. *Genome Biology* **12**, R81.

Rizzi, E., M. Lari, E. Gigli, G. De Bellis & D. Caramelli (2012). Ancient DNA studies: new perspectives on old samples. *Genetics Selection Evolution* **44**, 21-39.

Roberts, C. & S. Ingham (2008). Using ancient DNA analysis in palaeopathology: a critical analysis of published papers, with recommendations for future work. *International Journal of Osteoarchaeology* **18**, 600-613.

Russello, M. A., M. D. Waterhouse, P. D. Etter & E. A. Johnson (2015). From promise to practise: pairing non-invasive sampling with genomics in conservation. *PeerJ* **1**, 1-18.

Sangeetha, G. K. R., Blessy, B. M., Sudhamani, C. N., Bhoiva, N (2014) Mechanism of DNA binding and cleavage. *Biomedicine and Biotechnology* **2,** 1-9.

Sanger, F., S. Nicklen & A. R. Coulson (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National, Academy of Sciences of the United States of America* **74**, 5463-5467.

Sears, K. E. (2005). Role of development in the evolution of the scapula of the giant sthenurine kangaroos (Macropodidae: Sthenurinae). *Journal of Morphology* **265**, 226-236.

Shafer, A. B., J. B. Wolf, P. C. Alves, L. Bergstrom, M. W. Bruford, I. Brannstrom, G. Colling, L. Dalen, L. De Meester & R. Ekblom (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution* **30**, 78-87.

Shapiro, B., A. J. Drummond, A. Rambaut, M. C. Wilson, P. E. Matheus, A. V. Sher, O. G. Pybus, M. T. P. Gilbert, I. Barnes & J. Binladen (2004). Rise and fall of the Beringian steppe bison. *Science* **306**, 1561-1565.

Shapiro, B. & M. Hofreiter (2012). *Ancient DNA: methods and protocols,* New York, Humana Press.

Shapiro, B. & M. Hofreiter (2014). A paleogenomic perspective on evolution and gene function: New insights from Ancient DNA. *Science* **343**, 1-6.

Shendure, J. & J. Hanlee (2008). Next-generation DNA sequencing. *Nature Biotechnology* **26**, 1135-1145.

Shendure, J., G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra & G. M. Church (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728-1732.

Smith, I. C., A. T. Chamberlain, M. S. Riley, C. Stringer & M. Collins (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution* **45**, 203-217.

Steeves, T. E., R. N. Holdaway, M. L. Hale, E. McLay, I. A. W. McAllan, M. Christian, M. E. Hauber & M. Bunce (2010). Merging ancient and modern DNA: extinct seabird taxon rediscovered in the North Tasman Sea. *Biology Letters* **6**, 94-97.

Stefen, W., A. A. Burbidge, L. Hughes, R. Kitching, D. Lindenmayer, W. Musgrave, M. S. Smith & P. A. Werner (2009). *Australia's biodiversity and climate change*, CSIRO.

Stiller, M., G. Baryshnikov, H. Bocherens, A. G. d'Anglade, B. Hilpert, S. C. Munzel, R. Pinhasi, G. Rabeder, W. Rosendahl & E. Trinkaus (2010). Withering away - 25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution* **27**, 975-978.

Templeton, J. E. L., P. Brotherton, B. Llamas, J. Soubrier, W. Haak, A. Cooper & J. Austin (2013). DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investigative Genetics* **4**, 26-37.

Thomas, R. H., W. Schaffner, A. C. Wilson & S. Pääbo (1989). DNA phylogeny of the extinct marsupial wolf. *Nature London* **340**, 465-467.

Tillmar, A., Dell'Amico, J. Welander & Holmlund (2013). A universal method for species identification of mammals utilizing Next Generation Sequencing for the analysis of DNA mixtures. *PloS ONE* **8**, e83761.

Tringe S, G. & E. Rubin, M (2005). Metagenomics: DNA sequencing of environmental samples. *Nature Reviews Genetics* **6**, 805-814.

Valdiosera, C., N. Garcia, L. Dalen, S. Colin, R.-D. Kahlke, K. Liden, A. Angerbjorn, J.-L. Arsuaga & A. Gotherstrom (2006). Typing single polymorphic nucleotides in mitochondrial DNA as a way to access middle pleistocene DNA. *Biology Letters* **2**, 601-603.

Valiere, N. & P. Taberlet (2000). Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology* **9**, 2150-2152.

Watson, J. & F. Crick (1953). Molecular structure of Nucleic Acids. *Nature* **4356**, 2.

Westerman, M., B. P. Kear, K. Aplin, R. W. Meredith, C. Emerling & M. S. Springer (2012). Phylogenetic relationships of living and recently extinct bandicoots based on nuclear and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **62**, 97-108.

White, N. E., M. J. Phillips, M. T. P. Gilbert, A. Alfaro-Nunez, E. Willerslev, P. R. Mawson, P. Spencer & M. Bunce (2011). The evolutionary history of cockatoos (Aves: Psittaciformes: Cacatuidae). *Molecular Phylogenetics and Evolution* **59**, 615-622.

Willerslev, E., E. Cappellini, W. Boomsma, R. Nielsen, M. B. Hebsgaard, T. B. Brand, M. Hofreiter, M. Bunce, H. N. Poinar, D. Dahl-Jensen, S. Johnsen, J. P. Steffensen, O. Bennike, J.-L. Schwenninger, R. Nathan, S. Armitage, C.-J. de Hoog, V. Alfimov, M. Christl, J. Beer, R. Muscheler, J. Barker, M. Sharp, K. E. H. Penkman, J. Haile, P. Taberlet, M. T. P. Gilbert, A. Casoli, E. Campani & M. J. Collins (2007). Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. *Science* **317**, 111-114.

Willerslev, E. & A. Cooper (2005). Ancient DNA. *Proceedings of the Royal Society of London B: Biological Sciences* **272**, 3-16.

Willerslev, E., M. T. P. Gilbert, J. Binladen, S. Y. W. Ho, P. F. Campos, A. Ratan, L. P. Tomsho, R. R. Da Fonseca, A. Sher & T. V. Kuznetsova (2009). Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. *BMC Evolutionary Biology* **9**, e95.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky & A. Cooper (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* **300**, 791-795.

Woodward, S. R., N. J. Weyand & M. Bunnell (1994). DNA sequence from Cretaceous period bone fragments. *Science* **226**, 1229-1232.

Chapter Two

Using ancient DNA to gentically identify an extirpated island species

Figure 2.1.1 Black footed rock-wallaby (*Petrogale lateralis)* in natural habitat, rocky outcrops (Taken by Norbert Schuster).

2.1 PREFACE

*This chapter consists of a published manuscript titled '**The identity of the Depuch Island rock-wallaby revealed through ancient DNA**' [Australian Mammology 35 (1) 101-106]. The content of section 2.2 is the same as that of the published manuscript with only minor changes to incorporate thesis formatting and referencing. The published (pdf) version can be found in Appendix A.*

This chapter describes the use of conventional (or Sanger) sequencing technology to determine the genetic relationship of an extinct island rock-wallaby species to mainland taxa. Depuch Island, situated off the northwest coast of Western Australia, once contained a population of rock wallabies, which have become locally extirpated. Several morphologically similar mainland rock wallaby species exist in close proximity to the island, making identifications based on geography difficult. The information revealed through ancient DNA analysis was therefore useful in investigating which former species of rock wallaby was present on the island. In this study two mitochondrial genes are used: the cytochrome *b* gene contains well-defined regions for discriminating nucleotide polymorphisms between marsupial species, whilst the hypervariable control region is ideal for identifying genetic differences within species. Multifaceted studies of rock wallabies over the years have shown the complexity of wallaby taxonomy and evolutionary history, necessitating reliable species identification methods. It is therefore an important requirement for future conservation efforts, that we fully understand the past species that were present on the island before making restoration conservation decision. This is especially important as conservation managers are increasingly looking into islands as refugia, as mainland wallaby populations are experiencing severe environmental stresses (discussed further in section 2.2.2).

This chapter reports the successful amplification and sequencing of mtDNA from one of the only surviving Depuch Island museum specimens. Due to the degrading nature of ancient DNA, particularity that seen in Australia's warmer climates, strict procedures were undertaken to minimise exogenous DNA. Specific wallaby primers were designed to genetically characterise the island species, whilst Bayesian phylogenetic analyses were employed to model the DNA sequences together with

other published rock-wallaby sequences. The Sanger sequencing technology used in this study, which was the most common method for aDNA analysis at the time, is ideal for species identification using mtDNA genes and 'single source' samples, and contrasts with the HTS methods employed in Chapters Four and Five.

2.1.1 STATEMENT OF CONTRIBUTION

As lead author of the following manuscript, I conducted all laboratory work, including aDNA analysis and phylogenetic work. The Western Australian Museum kindly provided access to the fossil used for sampling in this study. I also wrote the first and final drafts of the manuscript with valuable comments, guidance and edits from Prof. Michael Bunce and Dr. James Haile. Dr. Peter Spencer also provided some comments and edits in the production of this manuscript.

Dalal Haouchar

2.2 THE IDENTITY OF THE DEPUCH ISLAND ROCK-WALLABY REVEALED THROUGH ANCIENT DNA

Dalal Haouchar[1], James Haile[1], Peter B. S. Spencer[2] and Michael Bunce[1].

[1]Ancient DNA Laboratory, School of Biological Sciences, Murdoch University, Perth, WA 6150, Australia.

[2]Wildlife Identification Laboratory, School of Biological Sciences, Murdoch University, Perth, WA 6150, Australia.

2.2.1 ABSTRACT

Ancient DNA is becoming increasingly recognised as a tool in conservation biology to audit past biodiversity. The widespread loss of Australian biodiversity, especially endemic mammal populations, is of critical concern. An extreme example occurred on Depuch Island, situated off the north-west coast of Western Australia, where an unidentified species of rock-wallaby (*Petrogale* sp.) became extinct as a result of predation by red foxes. Two potential candidate species, *Petrogale lateralis* and *P. rothschildi*, both have ranges adjacent to Depuch Island, making identification based on geography difficult. A museum bone (one of the only surviving Depuch Island specimens) was subjected to standard ancient DNA analyses and procedures. Mitochondrial DNA cytochrome *b* and hypervariable control region were targeted for species identification. Ancient DNA was successfully recovered from the bone: 200 base pairs (bp) of control region and 975 bp of the cytochrome *b* gene. Bayesian phylogenetic analyses were employed to model the Depuch Island rock-wallaby DNA sequences together with sequences of other rock-wallaby taxa from GenBank. Evidence suggests that of the two *Petrogale lateralis* subspecies proposed to have inhabited Depuch Island, *Petrogale lateralis lateralis* was identified as the most likely. The identification of the Depuch Island rock-wallaby population may assist in the reintroduction of an insurance population of *Petrogale lateralis lateralis*, which is becoming increasingly threatened on mainland Australia.

2.2.2 INTRODUCTION

The Australian mammal fauna has undergone several extinctions and population declines since European settlement (Burbidge *et al*., 2002; Johnson 2006). Among those species most severely affected have been the medium-sized species, more specifically marsupials that are in the 'critical weight range' (between 35 and 5500 g) (Johnson *et al.,* 2009). As a result, combinations of environmental pressures such as introduced feral predators, and demographic and genetic stochasticity have contributed to putting isolated populations of critical-weight marsupials, such as those that exist on islands (Frankham, 1997; Eldridge *et al.,* 1999) under extreme pressure.

Located off the north-west coast of Western Australia (Figure 2.2.1), Depuch Island (DI) was previously separated from the mainland by ~3 km (2 miles) of shallow water and mud flats (Ride, 1964). Whilst neighbouring islands in the Forestier's Archipelago are low and sandy, DI has undulating outcrops of dolerite and granite (McCarthy 1961; Ride 1964;), which provide ideal refuge sites and habitat for rock-wallabies. DI, known as Womalantha by Aboriginal people, was named in 1801 by Nicolas Baudin on board the *Le Géographe* and revisited 40 years later by Captain J. C. Wickham, commander of the *H.M.S. Beagle*, who conducted the first biological survey of the island (Ride, 1964). DI is well known to anthropologists as an important site of Aboriginal rock art and engravings (Ride, 1964) and was declared a sanctuary in 1958 for the protection of engravings and fauna (McCarthy, 1961). The island was previously inhabited by several bird, reptile and mammal species including a rock-wallaby (recorded as a small kangaroo by Pèron, 1807 and as *Petrogale lateralis* by Stokes, 1846), which became extinct as a consequence of predation by the introduced red fox (*Vulpes vulpes*) (Kinnear *et al.,* 1988, 1998). Rock-wallaby colonies were reported on DI at the mouth of the Balla Balla River and the animals were known to be present in 1964. Judging by the vast quantities of their droppings around the island, the rock-wallabies seemed plentiful at the time (Ride, 1964). However, 20 years later they were considered locally extinct (Hall and Kinnear, 1991) before any secure species identification could be made. Prior to any reintroduction attempts, a sound species and genetic characterisation of the extinct wallaby should be undertaken.

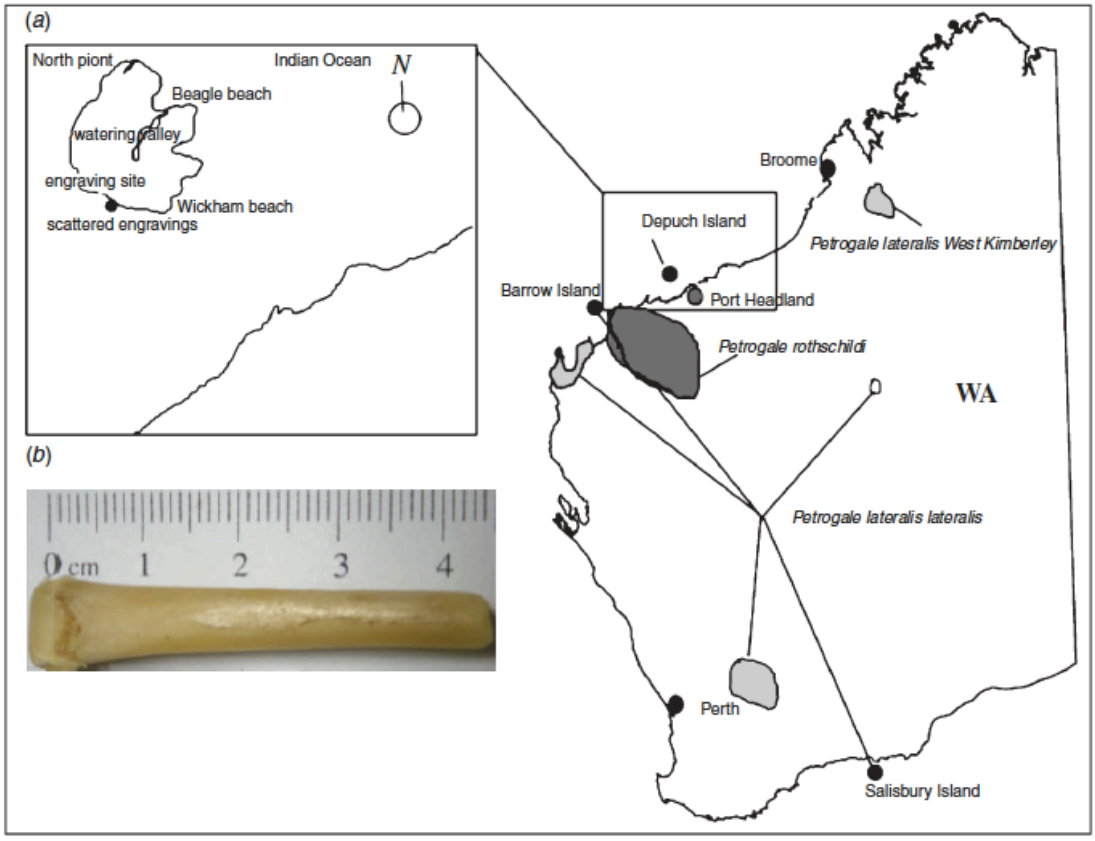


Figure 2.2.1 (a) Current species distribution of Petrogale rothschildi (dark grey), Petrogale lateralis lateralis (light grey) and Petrogale lateralis West Kimberley race (medium grey) in relation to Depuch Island. Inset shows an enlargement of Depuch Island, showing sites mentioned in the text. (b) Western Australian Museum fossil bone of fourth metatarsal (proximal end missing) (M5233; Depuch No. 48) from Depuch Island (collected on 15 June 1962) used in this study. Map (a) based on Mason *et al.,* (2011) and Eldridge *et al*., (1994).

At present, the taxonomic identity of the Depuch Island rock-wallaby is unresolved, with two candidate species inhabiting north-west Australia: the black-footed (black-flanked) rock-wallaby (*Petrogale lateralis*) and Rothschild's rock-wallaby (*Petrogale rothschildi*). The location of DI within the range of both species makes it difficult to predict which taxon might have formerly occupied this island (see Figure 2.2.1*a*). Ride (1964) set out to confirm Stokes' (Stokes, 1846) observation of an unknown species of rock-wallaby on DI, and stressed the importance of identifying the species for zoogeographical and taxonomic reasons. Ride (1964) noted that two species occurred in the vicinity of the island, *P. rothschildi* inhabiting islands of the Dampier Archipelago and *P. lateralis* being dispersed throughout the Western Australian mainland and also Barrow Island (Ride, 1964).

However, *Petrogale lateralis* is polytypic, consisting of five distinct subspecies/races – *P. l. lateralis*, *P. l. hacketti*, *P. l. pearsoni*, *P. l.* 'MacDonnell Ranges' and *P. l.* 'West Kimberley' (Sharman *et al.*, 1990; Eldridge *et al.*, 1994; Eldridge *et al.*, 1997; Pearson *et al.*, 1997; Eldridge *et al.*, 2001; Potter *et al.*, 2012) – some of which are readily distinguished by traditional morphology (e.g. skull/body morphology/size, coat colour and markings). However, such features are highly variable within rock-wallaby species and are therefore an unreliable means of classification (Eldridge *et al*., 1992; Eldridge *et al*., 1997; Eldridge *et al*., 1997). Prior to DNA studies one of the most convincing genetic species determiners had been chromosome and allozyme similarities (Sharman *et al.,* 1989; Eldridge *et al.,* 1999*a,* 1991*b,* 1994). However, the use of chromosomes to provide definitive species identification is limited by the technique's requirement for living tissue from which rapidly dividing cells can be cultured for analyses (Eldridge *et al.*, 1997; Loupis *et al.*, 2001). Now with DNA studies becoming a diagnostic tool for species identification, it is mandatory that this research be used in identifying which subspecies/races of *P. lateralis* likely occurred on DI. Given the location of *P. l. lateralis* and *P. l.* 'West Kimberley' near the island, either of these could be potential candidates. Therefore, the use of mitochondrial DNA (mtDNA) targeting cytochrome *b* and control region can be useful in telling these taxa apart.

Previously, chromosome analyses categorised *Petrogale* species into three distinct groups: the *xanthopus*, *brachyotis* and *lateralis–penicillata* groups (Eldridge and Close 1992, 1997). However, recent molecular analyses shows that the *xanthopus* group is not monophyletic and comprises independent lineages for *P. xanthopus*, *P. persephone* and *P. rothschildi*, despite their shared ancestral karyotypes (Potter *et al.,* 2012). Further analyses show that four distinct lineages were identified on the basis of mitochondrial and nuclear DNA: (1) the *brachyotis* group, (2) *Petrogale persephone*, (3) *Petrogale xanthopus* and (4) the *lateralis–penicillata* group (Potter *et al.,* 2012).

The aim of this study was to isolate ancient DNA from a museum specimen for species identification, namely targeting mtDNA. A marginally degraded subfossil museum bone (Figure 2.2.1*b*) was subjected to mitochondrial DNA (mtDNA) analysis of the cytochrome *b* gene and hypervariable control region. Until now no attempts had previously been made to clearly identify the endemic species of rock-

wallaby that previously inhabited the island, which is important as it will allow informed decisions to be made in any biodiversity restoration projects onto the now fox-free DI. Since this species is completely extinct on the island and there are only a few trace rock-wallaby museum specimens left, ancient DNA analysis is the most suitable technique and is highly recommended in order to handle and retrieve relevant data for species identification useful in this study.

2.2.3 MATERIALS AND METHODS

2.2.3.1 SAMPLE COLLECTION AND DNA EXTRACTION

Due to sampling restrictions, a single DI rock-wallaby subfossil bone (fourth metatarsal) was accessed from the Western Australian Museum (M5233; collected on 15 June 1962) and sampled for DNA (Figure 2.2.1*b*). DNA extraction procedures were carried out in a dedicated ancient DNA laboratory, minimising contamination from PCR amplicons and modern DNA. A sample of ~100 mg of bone powder was obtained using a Dremel tool (part no. 114; Germany) at high rotational speeds. Bone powder was collected in a 1.5-mL Eppendorf tube, weighed and stored for later digestion.

The digestion step included an overnight incubation at 55°C with rotation in a 1.5-mL digestion buffer containing 20 mM Tris pH 8.0 (Sigma, MO, USA), 10 mM dithiothreitol (Thermo Fisher Scientific, MA, USA), 1 mg mL^{-1} proteinase K powder (Amresco, OH, USA), 0.48 M EDTA (EDTA) (Invitrogen) and 1% Triton X-100 (Invitrogen). After digestion, centrifugation at 13 000 *g* was initiated for 1 min to pellet undigested material. The supernatant containing the DNA was concentrated to ~100 µL in a Vivaspin 500 column (MWCO 30000; Sartorius Stedim Biotech, Germany) at 13 000 *g*, and then combined with five volumes of PBi buffer (Qiagen, CA, USA). DNA was immobilised on silica spin columns (Qiagen) and washed with 700 µL of AW1 and AW2 wash buffers (Qiagen). Finally, the DNA was eluted from the silica in 50 µL of 10 mM Tris pH 8.0 (Sigma, MO, USA).

2.2.3.2 PRIMER DESIGN

Primers were designed in order to amplify small segments of DNA spanning the targeted mtDNA regions, ranging in size from 150 to 420 base pairs (bp) (see Table 2.2.1 for primer combinations). PCR amplification of 975 bp of cytochrome *b* gene and 200 bp of the hypervariable control region was targeted by designing specific primers for rock-wallabies using GENEIOUS 4.8 (Biomatters, New Zealand).

Table 2.2.1 Mitochondrial cytochrome *b* and control region primer sequences and amplification conditions used to genetically characterise the DI rock wallaby.

Primer name	Primer sequence 5′ → 3′	Annealing temperature (°C)	Amplicon size (bp)
DL1F	CCACAACACATCAACTYATTTG	53	150
DL1R	ATTCATTTTATGTATTACTAGAATTATGTA		
DL3F	TGTATTAAGACAGATATGTATAAAGT	53	250
DL2R	ATTCATTTTATGTATTACTAGAATTATGTA		
DL3F	TGTATTAAGACAGATATGTATAAAGT	53	280
DL3R	AGTCAGAGATTTGTTAGGTACG		
Cytb_WallabyF1	GACACCCTAACAGCCTTCTCATCAG	57	260
Cytb_WallabyR1	CGGTAGCTCCTCAGAATGATATTT		
Cytb_WallabyF2	AAATATCATTCTGAGGAGCTACCG	57	340
Cytb_WallabyR2	GAGAAGTTGTCTGGGTCGCC		
Cytb_WallabyF3	GGCGACCCAGACAACTTCTC	57	240
Cytb_WallabyR3	GGCTGTAAGGATTCAGAATAGGAT		

2.2.3.3 PCR OF CYTOCHROME B AND CONTROL REGION

Amplification of DNA product was carried out using real-time PCR to assess the amount of DNA preserved in the sample, a protocol advocated in many ancient DNA procedures (Cooper and Poinar 2000; Pruvost and Geigl 2004). The StepOne real-time PCR system (Applied Biosystems 3730 DNA analyser) was used with a final reaction concentration of: 1 × High Fidelity PCR Buffer (Invitrogen), 50 mM $MgSO_4$ (Invitrogen), 0.25 mM of each dNTPs (Austral Scientific), 8 μM of each primer, 0.25 U HIFI Taq polymerase (Invitrogen), 1 μL (10 mg mL^{-1}) bovine serum albumin (Fisher Biotech), 0.6 μL SybrGreen (Invitrogen cat no S7563, 1: 2000 dilution), ultrapure H_2O and 2 μL of DNA extract in a 25 μL reaction. PCR thermal cycling was initiated with a 5-min denaturation step at 94°C, followed by 50 cycles of 94°C

for 45 s, with a variable annealing temperature either 53°C or 57°C according to primer set used (see Table 2.2.1), followed by 68°C for 45 s, and a final extension at 72°C for 10 min. Quantitative PCR data were analysed using Applied Biosystems 3730 DNA analyser 2.0 software. The relative yield of DNA was assessed between each dilution sample according to the C_T values. Samples were visualised on a 2% w/v DNA-grade agarose gel electrophoresis (Bio-Rad) stained with ethidium bromide.

2.2.3.4 CLONING MTDNA

All DNA products were amplified at least twice, and control region PCR products that contained amplified fragments of the appropriate size were purified using a Qiaquick Purification kit (Qiagen) and cloned using pGEM-T vector system (Promega). Successfully cloned products were chosen after being screened for blue (which do not contain the pGEM-T vector) and white (which contain the pGEM-T vector) colonies on LB/ampicillin/IPTG/X-Gal plates. A random selection of ~10 white colonies and at least one blue colony was selected for size comparison and screened on a 2% agarose w/v gel. Amplicons were checked to ensure that products were of predicted size, purified finally and prepared for sequencing. This cloning step ensured the production of clean chromatographs facilitating unambiguous base calls, as well as allowing an assessment of DNA damage and polymerase error.

2.2.3.5 SEQUENCING OF CYTOCHROME B AND CONTROL REGION

Sanger sequencing was carried out at the commercial facility Macrogen (Seoul, South Korea) using BigDye ver. 3.1 (Applied Biosystems) chemistry on a cycling ABI3730XL capillary sequencer (Applied Biosystems). Sequences were analysed using GENEIOUS 5.4.3 (Biomatters, New Zealand) and deposited on GenBank under accession numbers JN898804 and JN898805.

2.2.3.6 PHYLOGENETIC ANALYSIS OF CYTOCHROME B AND CONTROL REGION SEQUENCE

Alignments of nucleotide sequences were carried out in GENEIOUS 5.4.3, with any ambiguities resolved by eye. Reference species used in this analysis were derived from published sequence data (Potter *et al.,* 2012). GenBank accession numbers for cytochrome *b* – *P. lateralis lateralis* (JQ042127-S972, Nangeen Hill, south-west Western Australia), *P. rothschildi* (JQ042134-S204, Rosemary Island, Dampier Archipelago) and *P. lateralis* West Kimberley race (JQ042130) – were used for comparison. Control region of *P. lateralis lateralis* (AF348675-S207, Ningaloo) was also used, along with *P. rothschildi* (not published) and *P. lateralis* West Kimberley race (AF348688) for comparison. To ensure validity of DNA sequences and to overcome ancient DNA damage, multiple-sequence datasets were created and an overall consensus was drawn for use in the final phylogenies. Topology analysis was conducted using the Bayesian phylogenetic program MrBayes (Huelsenbeck and Ronquist 2001), a plug-in application provided through GENEIOUS 5.4.3. The MrBayes analysis extended over 1 million iterations, and genealogies and model parameters were sampled every 1000 iterations, with 10% burn-in. An HKY85 substitution model was imposed and an invariance gamma model for among-site variation was chosen according to the simulations run in MrjModel Test (Nylander, 2004). Results were visualised and examined in FigTree ver. 1.2.2.

2.2.4 RESULTS

2.2.4.1 MTDNA ANALYSIS OF CYTOCHROME B AND CONTROL REGION

The cytochrome *b* sequence alignment of 975 bp of the DI rock-wallaby specimen to the two reference species, *P. l. lateralis* and *P. rothschildi*, consisted of 99% and 93% of identical sites, respectively. Among-site variation between the DI rock-wallaby and *P. l. lateralis* revealed 10 single-nucleotide polymorphisms (SNPs), nine transitions (7 C → T and 2 G → A) and one transversion (A → C). Between the DI rock-wallaby and *P. rothschildi* 69 SNPs were identified, 58 being transitions (48 C → T and 10 G → A) and 11 transversions (7 A → C and 4 A → T).

The control region (Domain 1) alignment of 200 bp of the DI rock-wallaby specimen and *P. l. lateralis* and *P. rothschildi* consisted of 92.3% and 63.2% identical sites,

respectively. Among-site variation revealed 15 SNPs between the DI rock-wallaby and *P. l. lateralis*, eight transitions (2 G → A and 6 C → T) and seven transversions (2 A → C, 4 A → T and 1 G → C). Among-site variation between the DI rock-wallaby and *P. rothschildi* consisted of 41 SNPs, 16 transitions (13 C → T and 3 G → A) and 25 transversions (3 A → C, 13 A → T, 3 G → C and 6 T → G).

2.2.4.2 PHYLOGENETIC ANALYSIS

Mitochondrial DNA sequences were positively identified on the basis of congruent phylogenies for both the cytochrome *b* gene and control region. This is a strong indication that sequences were not artefacts of nuclear copies. Modelling the concatenated consensus sequence data for both cytochrome *b* and control region was carried out using Bayesian phylogenetic methods. The *Thylogale* (pademelon) was chosen as the outgroup for this study. Phylogenetic analysis shows clear support that the unidentified rock-wallaby specimen collected from DI is most similar to *P. l. lateralis* at the mtDNA level (Figure 2.2.2). *P. l. lateralis* consistently forms a sister clade with the DI specimen (100% posterior probabilities) in both phylogenies to the exclusion of *P. rothschildi* and other closely neighbouring rock wallabies such as *P. lateralis* West Kimberley (Figure 2.2.2).

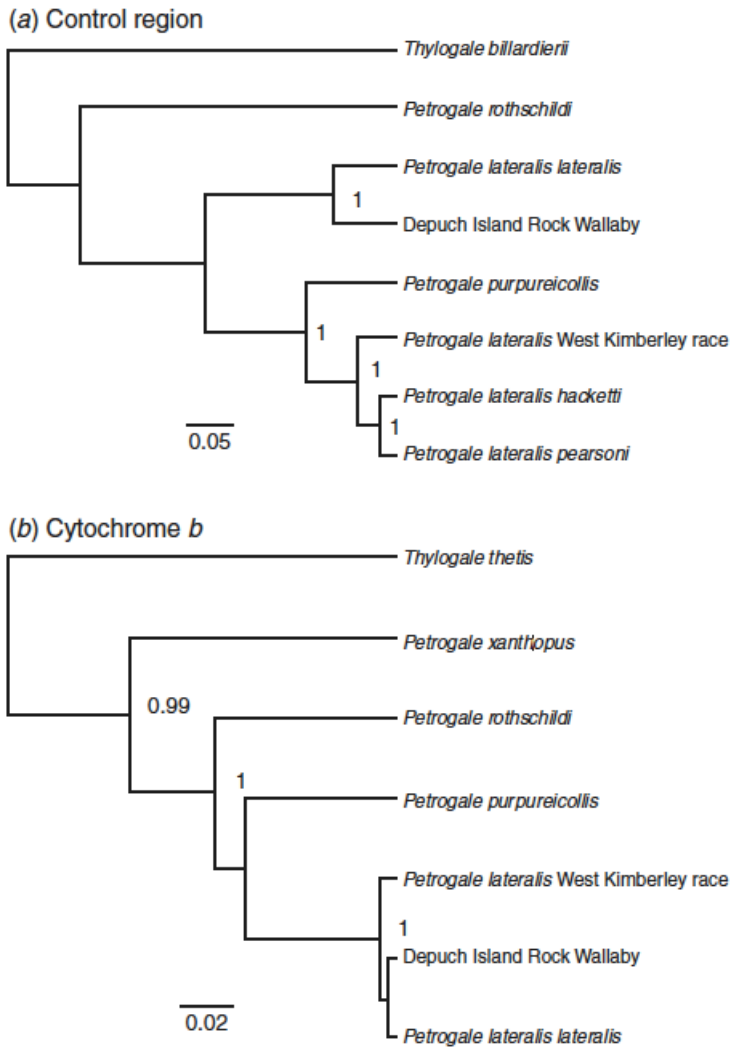


Figure 2.2.2 Bayesian phylogenetic tree showing the relationship of the Depuch Island rock-wallaby sequence with other available rock-wallaby sequence data: (**a**) based on 200 bp of control region and (**b**) based on 975 bp of cytochrome *b* sequence. Posterior probabilities greater than 90 are shown on nodes. The tree was built using a HKY85 model and invariant gamma was assumed and imposed with a relaxed molecular clock.

2.2.5 DISCUSSION

Mainland Australia has recorded the world's highest rate of recent mammal extinctions (Short and Smith 1994) and mainland *Petrogale* populations have not been immune to these changes. Having undergone severe population retractions, the remaining animals are highly vulnerable to introduced predators (Kinnear *et al.,* 1988, 1998), inbreeding depression and a multitude of interacting anthropogenic factors that further threaten population viability. Offshore islands have provided a refuge for many species, and currently harbor 14 species of Australian mammals

(Eldridge *et al.,* 1999) that are now extinct on the mainland. Such isolated islands can offer valuable insights into speciation and adaptive radiation and provide a refuge and conservation site for terrestrial mammal species (Maxwell *et al.,* 1996; Abbott, 2000). To date, species such as *Bettongia lesueur*, *Lagorchestus fasciatus*, *Perameles bougainville*, *Pseudomys fieldi*, amongst others, would now be extinct had they not been conserved on Western Australian islands (Burbidge *et al.,* 1997).

DI is currently isolated, with the land bridge connecting it to the mainland being permanently under water, creating a barrier to mammal dispersal to the island. Prior to this separation it was connected by ~3 km of mud flats, exposed at low tide, which enabled access to the island and therefore gene flow and migration of several species, including introduced predators such as the red fox. Rock-wallaby (*Petrogale* spp.) populations were noted to exist on the island at one time and on other offshore islands of Western Australia, such as Salisbury and Barrow Island (Hall and Kinnear 1991), which have remained fox free, mainly due to their isolation (Eldridge *et al.,* 2001). The species of rock-wallaby on DI, however, was never identified before it becoming locally extinct.

Our ancient DNA study has established that *P. l. lateralis* is most closely related to the DI rock-wallaby population and therefore the most probable rock-wallaby taxon that once inhabited DI. Having determined this makes future reintroduction attempts possible, with the exclusion of other potential rock-wallaby taxa like the *P. lateralis* 'West Kimberley' and *P. rothschildi* as potential candidates.

Previous attempts to identify unknown taxa using traditional morphological criteria have sometimes been challenging when dealing with morphologically cryptic species such as *Petrogale* (Loupis and Eldridge 2001). Other researchers have shown that although karyotypes can be useful in differentiating rock-wallaby taxa (Eldridge *et al.,* 1991a), a major limitation is the necessity of living tissue that can then be cultured (Eldridge and Spencer 1997). The advancement of molecular biology techniques and ancient DNA techniques means that even samples that have been subjected to severe degradation can be used as a valuable source of information to investigate past populations.

Petrogale l. lateralis is classified as Vulnerable by the *Environment Protection and Biodiversity Conservation Act 1999* and the *Western Australian Wildlife Conservation Act*. To ensure survival of this species in the face of the continuing decline of mainland populations (Hall and Kinnear 1991; Mason *et al*., 2011), we would advocate the relocation of a viable population of *P. l. lateralis* onto the now fox-free DI, which would provide an insurance population for the species, at least for the short term. Recent studies by Mason *et al.* (2011) and others (Frankham, 1997; Frankham *et al.,* 2002) highlight the importance of preserving fragmented mainland populations of rock-wallabies rather than translocating populations to islands, because of the increasingly low levels of genetic diversity. However, given that *P. l. lateralis* was the original inhabitant of DI, repopulating it will be a significant step in restoring the island's past biodiversity and one that is likely to have positive effects on the island's ecosystem. Additionally, a sound understanding of the genetic diversity of mainland populations is a necessary prerequisite to enhance the reintroduced population's survival in the new island environment (Mason *et al.,* 2011). The ultimate aim must be that, following successful reintroductions, such island sanctuaries can in the future be used as sources for returning species back onto the mainland, and, in doing so, restore at least some of the lost biodiversity of the recent past.

2.2.6 ACKNOWLEDGEMENTS

Thanks to Mark Eldridge for providing several critical reference sequences for this study to be carried out and to the Western Australian Museum (Dr Ric How) for providing the historical sample used in this study. This research was funded by the Australian Research Council grant 'An inventory of past biodiversity in Western Australia using ancient DNA' (FT0991741) awarded to MB.

2.2.7 REFERENCES

Abbott, I. (2000). Improving the conservation of threatened and rare mammal species through translocation to islands: case study Western Australia. *Biological Conservation* **93**, 195-201.

Burbidge, A. A., and Manly, B. F. J. (2002). Mammal extinctions on Australian islands: causes and conservation implications. *Journal of Biogeography* **29**, 465-473.

Burbidge, A. A., Williams, M. R., and Abbott, I. (1997). Mammals of Australian islands: factors influencing species richness. *Journal of Biogeography* **24**, 703-715.

Cooper, A., and Poinar, H. N. (2000). Ancient DNA: do it right or not at all. *Science* **289**, 1139.

Eldridge, M.D.B., and Close, R. L. (1997). Chromosomes and evolution in rock-wallabies, Petrogale (Marsupialia: Macropodidae). *Australian Mammalogy* **19**, 123-135.

Eldridge, M. D. B., and Close, R. L. (1992). Taxonomy of rock wallabies, Petrogale (Marsupialia, Macropodidae). I. A revision of the eastern *Petrogale* with the description of three new species. *Australian Journal of Zoology* **40**, 605-625.

Eldridge, M. D. B., Close, R. L., and Johnston, P. G. (1991a). Chromosomal rearrangements in rock-wallabies, *Petrogale* (Marsupialia, Macropodidae) .IV. G-Banding analysis of the *Petrogale lateralis* complex. *Australian Journal of Zoology* **39**, 621-627.

Eldridge, M. D. B., Johnston, P. G., and Close, R. L. (1991b). Chromosomal rearrangements in rock-wallabies, *Petrogale* (Marsupialia, Macropodidae). V. Chromosomal phylogeny of the *lateralis-penicillata* group. *Australian Journal of Zoology* **39**, 629-641.

Eldridge, M. D. B., King, J. M., Loupis, A. K., Spencer, P. B. S., Taylor, A. C., Pope, L. C., and Hall, G. P. (1999). Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby. *Conservation Biology* **13**, 531-541.

Eldridge, M. D. B., Kinnear, J. E., and Close, R. L. (1994). Identification of rock wallabies in the Calvert ranges, Little Sandy Desert, Western Australia, as *Petrogale lateralis lateralis*. *Australian Mammalogy* **17**, 129-131.

Eldridge, M. D. B., and Spencer, P. B. S. (1997). Sources and uses of genetic material in the study of *Petrogale* (rock wallabies) and other mammals. *Australian Mammalogy* **19**, 265-278.

Eldridge, M. D. B., Wilson, A. C. C., Metcalfe, C. J., Dollin, A. E., Bell, J. N., Johnson, P. M., Johnston, P. G., and Close, R. L. (2001). Taxonomy of rock-wallabies, *Petrogale* (Marsupialia: Macropodidae). III. Molecular data

confirms the species status of the purple-necked rock-wallaby (*Petrogale purpureicollis* Le Souef). *Australian Journal of Zoology* **49**, 323-344.

Foufopoulos, J., and Ives, A. R. (1999). Reptile extinctions on land-bridge islands: life-history attributes and vulnerability to extinction. *American Naturalist*, 1-25.

Frankham, R. (1997). Do island poulations have less genetic variation than mainland populations? *Heredity* **78,** 311-327.

Frankham, R., Ballou, J., and Briscoe, D. A. (2002). 'Introduction to Conservation Genetics'. Cambridge University Press: England.

Hall, G. P., and Kinnear, J. E. (1991). Recovery plan for the black-flanked rock-wallaby *Petrogale lateralis lateralis* (Gould) (Australian National Parks and Wildlife Service Endangered Species Program, Project 149). *Department of Conservation and Land Management, Western Australia.*

Huelsenbeck, J. P., and Ronquist, F. (2001). Mr Bayes: Baysian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.

Johnson, C. (2006). 'Australia's Mammal Extinctions: a 50000 year history'. Cambridge University Press: America.

Johnson, C. N., and Isaac, J. L. (2009). Body mass and extinction risk in Australian marsupials: The 'Critical Weight Range' revisited. *Austral Ecology* **34**, 35-40.

Kinnear, J. E., Onus, M. L., and Bromilow, R. N. (1988). Fox control and rock-wallaby population dynamics. *Australian Wildlife Research* **15**, 435-450.

Kinnear, J. E., Onus, M. L., and Sumner, N. R. (1998). Fox control and rock-wallaby population dynamics-II. An update. *Wildlife Research* **25**, 81-88.

Loupis, A. K., and Eldridge, M. D. B. (2001). Restriction fragment length polymorphism (RFLP) analysis of three nuclear genes in rock-wallabies (*Petrogale*: Marsupialia: Macropodidae): a search for genic markers to identify taxa within the *Petrogale lateralis-penicillata* group. *Australian Journal of Zoology* **49,** 27-36.

Mason, R., Browning, T., and Eldridge, M. (2011). Reduced MHC class II diversity in island compared to mainland populations of the black-footed rock-wallaby (*Petrogale lateralis lateralis*). *Conservation Genetics* **12**, 91-103.

Maxwell, S., Burbidge, A. A., and Morris, K. (1996). 'Action plan for the Australian Marsupials and Monotremes′ for the Australasian Marsupial and Monotreme Specialist Group, IUCN Species Survival Commission. *Wildlife Australia, Canberra*.

McCarthy, F. D. (1961). The rock-engravings of Depuch Island, north-west Australia. *Records of the Australian Museum* **25**, 121-148.

Nylander, J. A. A. (2004). MrModeltest v2. Program distributed by the author. *Evolutionary Biology Centre, Uppsala Univeristy*.

Pearson, D. J., and Kinnear, J. E. (1997). A review of the distribution status and conservation of rock-wallabies in Western Australia. *Australian Mammalogy* **19**, 137-152.

Peron, F. (1807). A voyage of discovery to the Southern Hemisphere performed by the order of the Emperor Napolean during the years, 1801, 1802,1803, and 1804. London: 129-132.

Potter, S., Cooper, S. J. B., Metcalfe, C. J., Taggart, D. A., and Eldridge, M. D. B. (2012). Phylogenetic relationships of rock-wallabies, *Petrogale* (Marsupialia: Macropodidae) and their biogeographic history within Australia. *Molecular Phylogenetics and Evolution* **62**, 640-652.

Pruvost, M., and Geigl, E. M. (2004). Real-time quantitative PCR to assess the authenticity of ancient DNA amplification. *Journal of Achaeological Science* **31,** 1191-1197.

Ride, W. D. L. (1964). Depuch Island. *The Western Australian Museum* **2**, 1-89.

Sharman, G., Close, R., and Maynes, G. (1990). Chromosome evolution, phylogeny and speciation of rock wallabies (*Petrogale*, Macropodidae). *Australian Journal of Zoology* **37**, 351-363.

Short, J., and Smith, A. (1994). Mammal decline and recovery in Australia. *Journal of Mammalogy* **75,** 288-297.

Stokes, J. L. (1846). Discoveries in Australia during the voyage of H.M.S. Beagle in the years 1837-38-39-40-41-42-43 London, (T. and W. Boone: London).

Woinarski, J. C. Z., A. A. Burbidge & P. L. Harrison (2015). Ongoing unraveling of a continental fauna: decline and extinction of Australian mammals since European settlement. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4531-4540.

2.3 CHAPTER SUMMARY

The content in this chapter demonstrates how the implementation of molecular tools can be applied to explore aspects of Australian mammal conservation. Well-preserved museum specimens, as described in this chapter, can provide valuable information and insight into contemporary management initiatives. More specifically, aDNA procedures were key to identifying the extinct species of rock wallaby on Depuch Island. As mainland populations shrink in size and distribution, the use of "mainland islands", sites at which the primary threat is excluded, typically by fencing, is a common approach to keeping species within their natural habitats, with out the need to translocate (Woinarski *et al.*, 2015). However, this is only successful if the threat is largely predation by feral cats and foxes and not deemed useful when the primary threat has been disease, poisoning by the introduced cane toads and intensive development. Therefore, translocation events become pertinent, given the primary threat is absent. Predator-free offshore islands, such as Depuch Island, continue to play a vital role in species conservation (see also Chapter Four, for another Australian island example). Coupled with information about a species genetic past diversity, which is now becoming more widely explored through aDNA analysis, is widening the scope of conservation genetics in Australian mammals, facilitating active management programs.

Whilst the genetic analysis from a well-preserved museum sample, as described above, can be highly informative, the interpretive power becomes greater when samples are collectively analysed from a number of populations across an entire species complex. Chapter Three explores another member of the Macropodidae family that has also experienced dramatic range retractions and population decline since European settlement, across the entire Australian continent. In Chapter Three, the mtDNA analysis of members of the genus *Bettongia* is conducted, with the inclusion of both modern genetic data and historical samples to fundamentally explore a different set of questions such as; former distribution; taxonomy, past biodiversity and evolutionary history.

Chapter Three

The radiation and evolutionary history of an endangered Australian marsupial (*Bettongia*) — using a combination of modern and ancient DNA

Figure 3.1.1 A juvenile species of bettong - the woylie (*Bettongia penicillata ogilbyi*). Photo taken by Sabrina Trocini.

3.1 PREFACE

*The work from this chapter has resulted in an accepted manuscript at the journal Biodiversity and Conservation titled: '**Ancient DNA reveals complexity in the evolutionary history and taxonomy of the endangered Australian brush-tailed bettongs (Bettongia: Marsupialia: Macropodidae: Potoroinae).'** Minor modifications have been made to this manuscript, in order to allow for in-thesis referencing.*

Well-preserved museum material (skins and fossils) with good provenance represent ideal substrates for aDNA studies. In Chapter Two the use of a single well-preserved bone, proved extremely valuable in aiding our understanding of an extirpated rock-wallaby species from an island location. In this chapter, an Australian genus (*Bettongia*) is investigated, and rather than being constrained to an offshore island, the study spans the entire Australian continent. Combining both modern and ancient DNA, this chapter explores the radiation, evolutionary history and taxonomy of bettongs, with a specific focus on brush-tailed bettongs. The integration of aDNA data presented in this study is crucial in understanding the past biodiversity and former connectivity of these endangered taxa. Increasing the number of historical specimens sampled allows for the investigation of maternal gene flow (using mtDNA) of an entire species complex across its former range, and the exploration of species inter-relationships past and present.

A recent study investigating the former connectivity and genetic decline of the woylie (*B. penicillata ogilbyi)* a brush-tailed bettong species in Western Australia, uncovered a specimen that was morphologically identified as *B. penicillata ogilbyi.* The mtDNA however, was significantly different from any modern or historical haplotype described to date in Western Australia (Pacioni *et al.,* 2015) and more closely matched the east coast bettong taxa. This finding raised questions about the taxonomic relationships of bettongs across the continent, and prompted the research in this chapter.

To investigate connectivity and species boundaries in the brush-tailed bettongs is in imperative to collate a meaningful sample size of modern and historical material. Ancient material can be gathered from a wide range of palaeontological sites that coincide with a species geographic range and also sourcing museum material (ideally integrated to capture past biodiversity), in order to collectively provide a more spatial and temporal comparative analysis. While both nuclear and mtDNA markers were used in Pacioni *et al.,* (2015) – see co-authorship in Appendix D, preservation restricted this study to the latter. In this chapter, two competing hypotheses are tested

in regards to the origin, connectivity and radiation of the brush-tailed bettongs. A better understanding of species boundaries is needed to ensure restoration initiatives (e.g. captive breeding) are being appropriately implemented.

3.1.1 STATEMENT OF CONTRIBUTION

As lead author of the following submitted manuscript, I conducted all laboratory work and wrote the first and final drafts of the manuscript. A/Prof Jeremy Austin provided the eastern bettong sequences data used in this study. The co-authors listed in the manuscript performed crucial roles in editing and providing support and assistance during the preparation of the manuscript. Dr Carlo Pacioni provided assistance with the data analysis and provided most of the modern data that was combined within this study. Dr Matthew Phillips, an expert in the field of phylogenetics and molecular dating approaches, provided critiques on the best approach in analysing the combined (modern and aDNA) bettong dataset, along with some phylogeographic assistance. Prof. Michael Bunce provided overall advise on experimental approaches, and reviewed numerous versions of the manuscript.

Dalal Haouchar

3.2 ANCIENT DNA REVEALS COMPLEXITY IN THE EVOLUTIONARY HISTORY AND TAXONOMY OF THE ENDANGERED AUSTRALIAN BRUSH-TAILED BETTONGS (*BETTONGIA*: MARSUPIALIA: MACROPODIDAE: POTOROINAE).

Dalal Haouchar[1, 2], Carlo Pacioni[1, 2], James Haile[2, 3], Matthew C. McDowell[4], Alexander Baynes[5], Matthew J. Phillips[6], Jeremy J. Austin[7], Lisa C. Pope[8], Michael Bunce[2].

[1]Ancient DNA laboratory, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia.
[2]Trace and Environmental DNA laboratory, Department of Environment and Agriculture, Curtin University, Bentley, WA 6845, Australia.
[3]PalaeoBARN, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, England.
[4]School of Biological Sciences, Flinders University of South Australia, Bedford Park, SA 5042, Australia.
[5]Western Australian Museum, Locked Bag 49, Welshpool DC, WA 6986, Australia.
[6]School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Gardens Point, QLD 4000, Australia.
[7]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Adelaide, SA 5000, Australia.
[8]School of Biological Sciences, The University of Queensland, St Lucia, QLD 4072, Australia.

3.2.1 ABSTRACT

The three surviving 'brush-tailed' bettong species — *Bettongia gaimardi* (Tasmania), *B. tropica* (Queensland) and *B. penicillata* (Western Australia), are all classified as threatened or endangered. These macropodids are prolific diggers and are recognised as important 'ecosystem engineers' that improve soil quality and increase seed germination success. However, a combination of introduced predators, habitat loss and disease has seen populations become increasingly fragmented and census numbers decline. Robust phylogenies are vital to conservation management, but the extent of

extirpation and fragmentation in brush-tailed bettongs is such that a phylogeny based upon modern samples alone may provide a misleading picture of former connectivity, genetic diversity and species boundaries. Using ancient DNA isolated from fossil bones and museum skins, we genotyped two mitochondrial DNA (mtDNA) genes: cytochrome *b* (266 bp) and control region (356 bp). These ancient DNA data were combined with a pre-existing modern DNA data set on the historically broadly distributed brush-tailed bettongs (~ 300 samples total), to investigate their phylogenetic relationships. Molecular dating estimates the most recent common ancestor of these bettongs occurred *c.* 2.5 Ma (million years ago), which suggests that increasing aridity likely shaped their modern-day distribution. Analyses of the concatenated mtDNA sequences of all brush-tailed bettongs generated five distinct and well-supported clades including: a highly divergent Nullarbor form (Clade I), *B. tropica* (Clade II), *B. penicillata* (Clades III and V), and *B. gaimardi* (Clade IV). The generated phylogeny does not reflect current taxonomy and the question remains outstanding of whether the brush-tailed bettongs consisted of several species, or a single widespread species. The use of nuclear DNA markers (single nucleotide polymorphisms and/or short tandem repeats) will be needed to better inform decisions about historical connectivity and the appropriateness of ongoing conservation measures such as translocations and captive breeding.

3.2.2 INTRODUCTION

Species in the rat-kangaroo genus *Bettongia* provide an ideal model for studying evolution and speciation within an Australian context, due to short generation times, high reproductive rates, and adaptation to a variety of habitats (Eldridge *et al.,* 2012). Since European settlement of Australia *c.* 1788, the distributions of *Bettongia* have been greatly reduced (Figure 3.2.1), with major range contractions coinciding with the introduction of exotic predators and competitors (Christensen 1980; Priddel and Wheeler 2004; Woinarski *et al.,* 2014), and the widespread clearing of native vegetation for agriculture (Sampson 1971; Strahan 1983; Christensen 1995; Pizzuto *et al.,* 2007; Pacioni *et al.,* 2011). The current taxonomy of *Bettongia* recognises five species that are known to have been extant at the time of European settlement of Australia (Wakefield 1967; Calaby and Richardson 1988; McDowell *et al.,* 2015). *Bettongia lesueur* (burrowing bettong) had a wide distribution across much of Australia apart from the forested eastern margin, the forested southwestern corner and the northern tropical savannahs. It is now extinct throughout that mainland range and only survives on three islands off northwestern Australia (Woinarski *et al.,* 2014). *Bettongia anhydra* (desert bettong) is known from only a single specimen collected as a live animal in central Australia (of which only the skull and dentaries were preserved), and is presumed to be extinct (McDowell *et al.,* 2015; Woinarski *et al.,* 2015). Originally described as a subspecies of *B.*

penicillata and later synonymised with *B. lesueur*, its species-level status and phylogenetic distinctiveness have only recently been demonstrated (McDowell *et al.,* 2015). The remaining three species comprise the "brush-tailed bettongs" (Wakefield 1967), which are the focus of this study (distributions are shown in Figure 3.2.1). *Bettongia gaimardi* (eastern bettong) is divided into two subspecies. A nominate subspecies, *B. g. gaimardi*, which was originally distributed through the southeastern Australian mainland between the Great Dividing Range and the coast, but is now extinct (Wakefield 1967; Woinarski *et al.,* 2014), and a second subspecies, *B. g. cuniculus*, which is restricted to central and eastern Tasmania, where its status is vulnerable (Woinarski *et al.,* 2014). *Bettongia tropica* (northern bettong) was the last species described (Wakefield 1967) and is restricted to sclerophyll forest abutting rain forest in the uplands of the wet tropics of northern Queensland (Laurance 1997; Pope *et al.,* 2000) where it is endangered (Woinarski *et al.,* 2014). *Bettongia penicillata* is classified as two subspecies. Museum records indicate that *B. p. penicillata* occurred in southeastern Australia inland of the Great Dividing Range and southeastern South Australia (Wakefield 1967). Anecdotal evidence (Burbidge *et al.,* 1988) extends its original range northward and westward as far as the Great Sandy Desert (Woinarski *et al.,* 2014) but this may involve confusion with other similar taxa, including *B. anhydra*. *Bettongia p. penicillata* is extinct (Woinarski *et al.,* 2014). The second subspecies, *B. p. ogilbyi* (Figure 3.1.1 - the woylie), occurs in southwestern Australia (Figure 3.2.1), where it now occupies a very much reduced range and is critically endangered (Woinarski *et al.,* 2014; Wayne *et al.,* 2015). Holocene fossils identified as *Bettongia penicillata* from caves in southeastern Western Australia, including the Nullarbor Plain (Baynes 1987), and western South Australia (McDowell 1997), appeared to indicate that in pre-European times *B. penicillata* had a continuous distribution between southeastern and southwestern Australia.

Taxonomy within the genus *Bettongia* is currently unstable due to conflicting morphological interpretations (Finlayson 1958; Wakefield 1967; Ride 1970; Sharman *et al.,* 1980). Early taxonomic accounts were primarily based on descriptions of morphological characteristics with emphasis on cranial structure and dental morphology (Finlayson 1958; Wakefield 1967; Strahan 1983; Rose 1986). The morphological differences between the three currently accepted species of brush-tailed bettongs and *B. lesueur* appear to be driven primarily by diet and habitat; *B. lesueur* feeds on vegetation such as leaves, flowers, seeds, fruits, roots, tubers and underground fungi (Menkhorst 2001), whilst the brush-tailed species primarily eat hypogeal fungi (McIllwee and Johnson 1998). As a result, the dental morphological differences and the driving forces for adaptation between species of brush-tailed bettong are rather subtle compared to *B. lesueur,* which has a less pronounced rostrum and more inflated tympanic bullae (Wakefield 1967; Menkhorst

2001; McDowell *et al.,* 2015). Similarly, *Bettongia pusilla*, the molars of which are much less bunodont than any other bettong and have a distinct protolophid and hypolophid (McNamara 1997), is morphologically distinct from the brush-tailed bettongs.

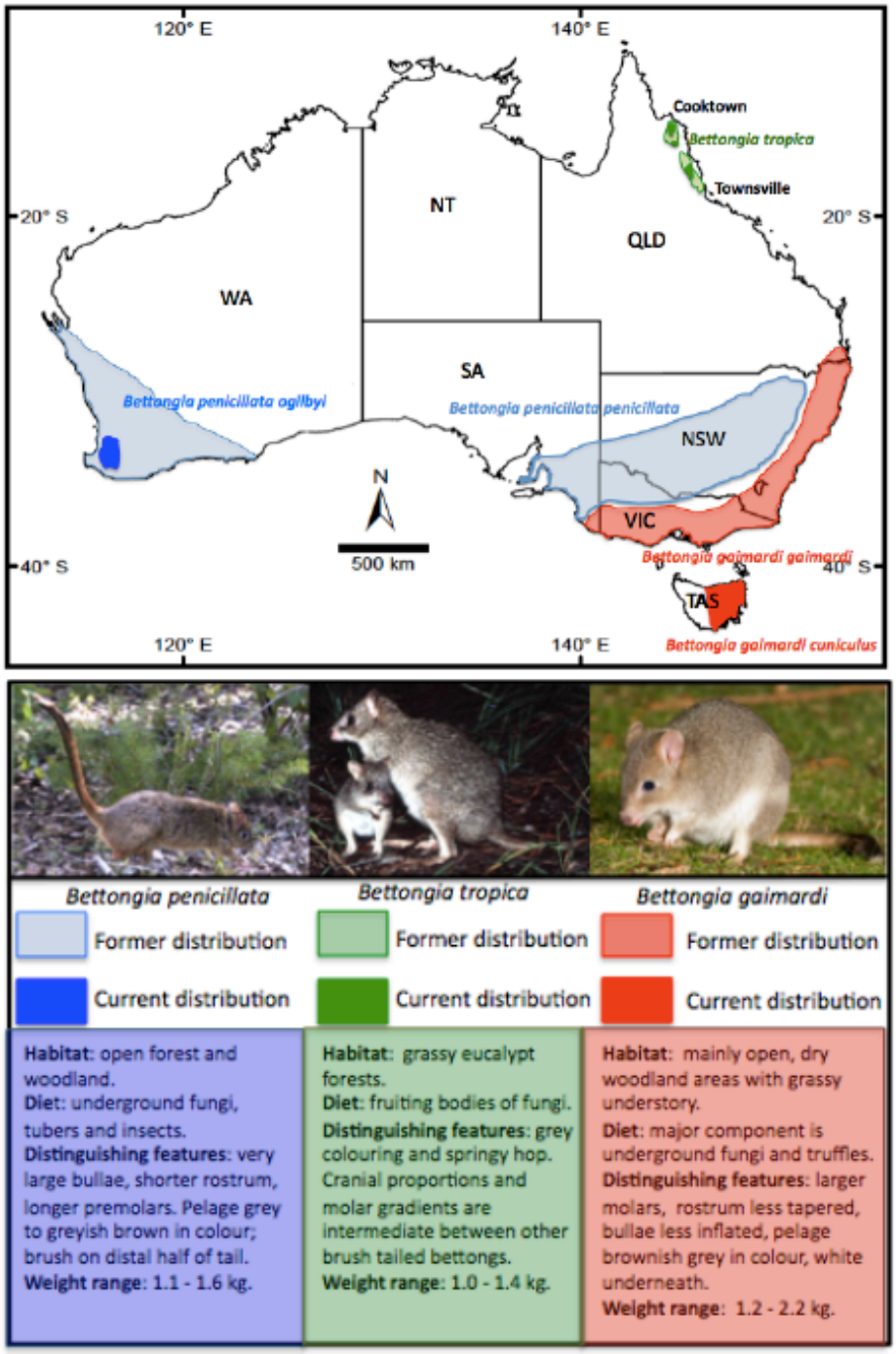


Figure 3.2.1 Former and current distributions of brush-tailed species of *Bettongia* throughout Australia. Only ranges based on fossils and/or skin samples are shown (ethnographic records omitted). Distributions and data were derived from Pope et al. (2000), Menkhorst (2001), Van Dyck and Strahan (2008) and mammals of Australia IUCN redlist website IUCN 2013. *Bettongia gaimardi* photograph was taken by JJ Harrison; *Bettongia penicillata* photograph by Sabrina Trocini; *Bettongia tropica* photograph by Karl Vernes. Concise data on *Bettongia* species included.

Genetic studies have identified *B. lesueur* as being distinct from the other *Bettongia* species, but other relationships are less clear. Preliminary work by Drummond (1933), later supported by Sharman (1961), found that *B. lesueur* has a total of 22 chromosomes, whereas *B. penicillata* has 28 chromosomes. In a recent study that investigated the genetic decline and former connectivity of the critically endangered *B. penicillata ogilbyi* (woylie) using aDNA (Pacioni *et al.,* 2015), a specimen identified on morphological characters as a small *B. penicillata* from a western Nullarbor cave was uncovered and yielded a DNA signature that was substantially divergent, with 'intermediate' affiliations to east-coast bettongs. This raised questions about the former connectivity amongst *Bettongia* species and populations across southern Australia and their taxonomic relationships.

Ancient DNA extracted from fossils can be used as a constructive tool to explore past biodiversity. The application of aDNA analyses is especially important for taxa, such as bettongs, that have been extirpated from the majority of their former ranges. In such circumstances the DNA isolated from small remnant extant populations may provide a misleading picture of species boundaries and former connectivity. Relevant aDNA data, particularly about a species' past connectivity and distribution limits, can inform conservation management and aid in both restoration initiatives and captive-breeding programmes. If the high level of mitochondrial genetic diversity loss observed in the woylie (conservatively estimated at 46-91%) (Pacioni *et al.,* 2015) is representative of the eastern Australian brush-tailed bettongs, then there is a pressing need to understand the conservation implications.

In this study, we use mtDNA loci (cytochrome *b* and control region) from Holocene fossil bone samples to: (i) obtain unique ancient DNA sequences, which, coupled with modern genetic data, we use to investigate the inter-relationships amongst extinct and extant members of *Bettongia;* (ii) explore the distribution and former connectivity amongst *Bettongia* species and populations across southern Australia; (iii) explore taxonomic relationships; and (iv) estimate divergence times between species and lineages. Collectively the aim of this work is to better understand the evolutionary history of brush-tailed bettongs with the goal of informing best-practice conservation management of these endangered taxa.

3.2.3 MATERIALS AND METHODS

3.2.3.1 SAMPLE COLLECTION

Genetic analyses were carried out on 88 ancient (fossil bone material) and historical samples (museum skins), including 18 *B. penicillata* and 54 *B. gaimardi* samples. Sixteen *B. tropica* samples and two sequences from GenBank (AY237236 and AY237237) were also included along with unique *B. p. ogilbyi* haplotypes from ~ 214 previously screened samples, consisting of 150 modern (Pacioni *et al.,* 2011) and 64 museum skin and fossil bone samples (Pacioni *et al.,* 2015) (Table 3.2.1). Together this represents a significant set of ancient, historical and modern bettong data and is the first time relationships between the brush-tailed bettongs have been critically evaluated using DNA.

Table 3.2.1 Samples of *Bettongia* used in this study, including the original source of sequences, geographic locations and names referred to in Figure 3.2.1. Sections highlighted represent historical / fossil samples. Samples labeled "this study" refer to the subset of unique haplotypes newly generated in this study from the initial dataset of 88 samples.

Species	Common name	Geographic location	Reference name to Figure 3.2.1	Source
B. lesueur	Burrowing bettong	Unknown locality	Bettongia lesueur	(Westerman *et al.*, 2004)
B. p. ogilbyi	Woylie	Upper Warren, WA	WA (haplotype F)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Tutanning, WA	WA (haplotype H)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Tutanning, WA	WA (haplotype I)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Upper Warren, WA	WA (haplotype K)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Upper Warren, WA	WA (haplotype L)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Upper Warren, WA	WA (haplotype N)	(Pacioni *et al.*, 2011)
B. p. ogilbyi	Woylie	Faure Island, WA	WA (MB808)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Tunnel cave, WA	WA (MB810)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Tunnel cave, WA	WA (MB812)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Witchcliffe Rock Shelter, WA	WA (MB816)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Witchcliffe Rock Shelter, WA	WA (MB818)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Contine, WA	WA (MB821)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Dwarladine Pool, WA	WA (MB823)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Jerramungup, WA	WA (MB827)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Mammoth cave, WA	WA (MB829)	(Pacioni *et al.,* 2015)
B. p. ogilbyi	Woylie	Dryandra, WA	WA (MB850)	(Pacioni *et al.,* 2015)
B. gaimardi	Eastern bettong	Michelago, NSW	NSW (5658A)	This study
B. gaimardi	Eastern bettong	Wombeyan Cave, NSW	NSW (5648A)	This study
B. gaimardi	Eastern bettong	Murdunna, TAS	TAS (5915A)	This study
B. gaimardi	Eastern bettong	Pipers River, TAS	TAS (5896A)	This study
B. gaimardi	Eastern bettong	Glenfern, TAS	TAS (5583A)	This study
B. gaimardi	Eastern bettong	Elizabeth Town, TAS	TAS (5900A)	This study
B. gaimardi	Eastern bettong	Randalls Bay, TAS	TAS (5931A)	This study
B. gaimardi	Eastern bettong	Upper Yarra, VIC	VIC (5595A)	This study
B. gaimardi	Eastern bettong	Unknown locality, VIC	VIC (5585A)	This study
B. penicillata	Brush-tailed bettong	Kangaroo Island, SA	SA (MB1510)	This study
B. penicillata	Brush-tailed bettong	Nullarbor Plain, WA	WA (MB1010)	This study
B. penicillata	Brush-tailed bettong	Nullarbor Plain, WA	WA (MB2010)	This study
B. penicillata	Brush-tailed bettong	Wylie scarp, WA	WA (MB1514)	This study
B. penicillata	Brush-tailed bettong	Venus Bay, SA	SA (MB1511)	This study
B. tropica	Northern bettong	QLD	QLD (08-369)	This study
B. tropica	Northern bettong	QLD	QLD (haplotype A)	(Pope *et al.*, 2000)
B. tropica	Northern bettong	QLD	QLD (haplotype C)	(Pope *et al.*, 2000)

The published sequences arising from 'this study' are published on GenBank (Accession numbers – XXXXX1 –XXXX15).

3.2.3.2 ANCIENT DNA EXTRACTION, PCR AMPLIFICATION AND SANGER SEQUENCING

All pre-PCR sample handling was performed in a dedicated aDNA facility at Murdoch University, WA, Australia, following stringent procedures to guard against contamination from DNA and PCR amplicons (Willerslev and Cooper 2005: see also Supplementary Information I (SI I) for more details). Fossil bones were cleaned by removing the surface, then approximately 0.2 g of the cleaned bone was fragmented using a Dremel tool (part no. 114: Germany) at high rotational speeds. All equipment was thoroughly cleaned to avoid cross contamination between samples as detailed in SI I. The bone powder from each sample was collected in a 1.5 ml Eppendorf tube, weighed and stored at 4°C for later digestion. The bone digest buffer and DNA extraction cleanup processes followed Haouchar *et al.* (2013: see also SI II). Briefly, a total of 1.5 ml of bone digest buffer was added to each tube containing the bone powder, followed by an over night digestion. The supernatant from the digest was concentrated in a Vivaspin 500 column and then processed with multiple wash steps, until finally being eluted from silica columns (SI II for more information). DNA extracts were screened using primer sets designed in this study to target *Bettongia* DNA sequences (Table 3.2.2). The coding region cytochrome *b* (Cyt *b*), and non-coding control region (CR), were chosen as they are highly informative genes, suitable for elucidating relationships at the species level (Patwardhan *et al.,* 2014).

Table 3.2.2 Mitochondrial DNA cytochrome *b* and control region primer information used in this study to genetically characterise *Bettongia.*

Primer name	Primer sequence 5' → 3'	Annealing temperature (°C)	Amplicon size (bp)	Source
Cytochrome *b* primer sets				
Woylie_cytb_32F	ATGAAAAACCATTGTTGTACTTC	53	220	This study
Woylie_cytb_249R	AGGGTGTCAGATGTATAGTGTAT			
Woylie_cytb_139F	ACCTTCCAACATTTGAGCCTGATG	55	250	This study
Woylie_cytb_388R	TGAGCCGTAGTAGATTCCTC			
Woylie_cytb_290F	AACTACGGCTGGTTAATCC	54	205	This study
Woylie_cytb_493R	TCAGAATGATATTTGTCCTCATGGCA			
Control region primer sets				
Woylie_cr_98F	GGTTTATGGRTTTAGATTCGTGAATC	55	145	This study
Woylie_cr_240R	CAAACCTATACCCATAGCATATCTCC			
Woylie_cr_156F	GCTTTGGATGTAATGTTTAAGGTAGAT	55	140	This study
Woylie_cr_296R	GCAAGYACAAATCAACCAAACTA			
Woylie_cr_1448F	TTGTACTTGCTTATATGCTTGGGGTGA	58	250	This study
Woylie_cr_1699R	GTGACTCGACCTTTTTCTACA			

A targeted qPCR (quantitative Polymerase Chain Reaction) assay was used to screen samples for mtDNA (also see SI IV – for more information on qPCR). Three concentrations of the final elution were tested – undiluted, 1/10 and 1/100. This approach was used in order to detect PCR inhibition and maximise input DNA; a common procedure that is recommended for use on degraded samples (Bunce *et al.,* 2012). Each qPCR reaction consisted of a total volume of 25 µl, containing; 12.5 µl ABI Power SYBR master mix (Applied Biosystems), 0.4 µM of forward and reverse primer, 8.5 µM H_2O and 2 µl DNA extract. Reaction conditions for the specific *Bettongia* primer sets were as follows; initial heat denaturation at 95°C for 5 mins, followed by 50 cycles of 95°C for 30 sec; with a variable annealing temperature depending on primer set used (see Table 3.2.2) for 30 sec; 72°C for 45 sec followed by a 1°C melt curve and a final extension at 72°C for 10 mins. For each qPCR assay, DNA extraction, negative PCR reagent and positive DNA template controls were included. The relative yield of DNA was assessed between each dilution sample according to the C_T values. Samples were visualised under UV light on a 3% w/v DNA-grade agarose gel electrophoresis, stained with ethidium bromide. Ancient DNA extracts that successfully yielded amplicons (as determined by initial qPCR screens), were prepared for sequencing. Firstly, a Qiaquick PCR purification kit (Qiagen) was used to purify amplicon, followed by Sanger sequencing in both directions at Macrogen (Seoul, South Korea) using BigDye version 3.1 (Applied Biosystems) on an ABI3730XL capillary sequencer (Applied Biosystems). DNA sequencing was, in some cases, repeated more than once, as a means of validating the sequences. Raw sequences were verified and aligned using GENEIOUS version 6.0.4 (Drummond *et al.,* 2011) native algorithm. Aligned sequences were then compared with previously published sequences available on GenBank (Table 3.2.1).

3.2.3.3 MODEL SELECTION AND PHYLOGENETIC ANALYSIS

Model comparisons within the Bayesian phylogenetic and phylogeographic analysis framework were performed utilising the stepping-stone (SS) and path sampling (PS) marginal likelihood estimators (Xie *et al.,* 2011; Baele *et al.,* 2012) in BEAST (Bayesian Evolutionary Analysis Sampling Trees - Drummond *et al.,* 2012). Two runs were performed for each model to check for convergence of Maximum Likelihood Estimates (MLE) and the two output files were then combined to estimate the final marginal likelihood. Finally, the screened models were ranked based on MLE. Initial tests for most appropriate models of DNA substitution were conducted with jModelTest version 2.1 (Posada 2008) using the Akaike Information Criterion (AIC). Subsequently, we compared, for both cytochrome *b* and control region, TN93, HKY and GTR (with and without *I* and *G*) in BEAST by comparing MLE as described above to confirm the accuracy of the selected

models (see results in Table SI IV). Once the most appropriate substitution model was determined, a final concatenated mtDNA alignment was produced in GENEIOUS, including the two mtDNA genes, imported into BEAUTI and executed in BEAST version 1.8.

A strict clock (Drummond *et al.,* 2006) was employed for the concatenated genes with the base frequencies estimated in BEAST. The concatenated alignment was then trialled over three different tree priors—Yule process (Yule 1925), birth-death process (Gernhard 2008) and birth-death incomplete sampling (Stadler 2009)—and compared with SS and PS marginal likelihood estimators (Xie *et al.,* 2011; Baele *et al.,* 2012). Two independent analyses of 10 million generations, sampled every 1000 generations, were run and TRACER version 1.5 (Rambaut and Drummond 2009) was used to check for sufficient mixing and to ensure effective sample sizes (ESS) for posterior estimates of more than 200. The first 10% of trees from each chain were discarded as burnin. Runs where chain convergence was not appropriate and ESS values were < 200 were repeated, increasing the number of generations. After the appropriate tree prior was selected, we proceeded to compare three different molecular clocks - strict, lognormal and exponential - with a similar approach as above. Finally, the data from the two runs from the supported molecular clock model were combined using LOGCOMBINER version 1.7.5 (Drummond and Rambaut 2007), the phylogenetic trees were summarised using TREEANNOTATOR version 1.7.4 (Drummond and Rambaut 2007; Drummond *et al.,* 2012) and the resultant Maximum Clade Credibility (MCC) tree was visualised using FIGTREE version 1.3.1. Once the main clades were identified, we computed the mean number of differences within (Tajima 1983) and between clades and the pairwise F_{ST} estimates (Nei and Li 1979) using Arlequin version 3.5 (Reynolds *et al.,* 1983). While it is computationally simpler to model only unique haplotypes, this approach has the potential to introduce bias. Therefore, we repeated the last analysis including 120 available samples but limited the modern woylie to 30 randomly selected sequences from Pacioni *et al.* (2011).

3.2.3.4 MOLECULAR DATING AND PHYLOGEOGRAPHY

There are no suitable fossils to calibrate the bettong radiation, with specimens either too young to be informative or phylogenetically disputed. In particular, the Miocene *Bettongia moyesi* may not even be a crown potoroine (Cooke *et al.,* 2015). Westerman *et al.* (2004) relied on external calibration. This option, however, would exclude the highly informative CR sequences, which cannot be confidently aligned between deeper divergences. We added a temporal component to the analysis by adding a prior (mean=8.5 million years, SD=1.785; with normal distribution) on the divergence time estimate for the root of *Bettongia* (of 8.3 - 1.5 million years, Westerman *et al.,*

2004). The analysis with the node calibration was executed in BEAST version 1.8.3. All other settings as well as post analyses (i.e. evaluating the runs in TRACER and generating MCC trees in TREEANNOTATOR) were the same as those mentioned in previous section. The tree obtained from BEAST was then used as an input file for the Generalised Mixed Yule Coalescent analysis (GYMC) (Fujisawa and Barraclough. 2013). This analysis is a likelihood-based approach that attempts to establish species delimitation. The analysis was conducted using a single threshold with the R package splits.

Data on source locations of samples (Table SI V), enabled us to investigate the phylogeographic radiation patterns of *Bettongia* using the continuous diffusion model, which has been demonstrated to have high statistical efficiency (Lemey *et al.,* 2010). This approach associates a location (in the form of coordinates in decimal degrees, Table SI V) to each sequence and allows the reconstruction of the spatial pattern of radiation while concurrently inferring the phylogeny. It does not assume that the initial location of the radiation process is sampled, as opposed to the discrete diffusion model (Lemey *et al.,* 2010), but allows the radiation to occur, continuously, in all possible directions.

As implemented in BEAST, the continuous diffusion model can use four different dispersal processes (see Lemey *et al.,* 2010 for details on statistical properties of this method and how it is implemented in BEAST). We compared MLE obtained with the analysis of the CR using the Homogeneous Brownian, Gamma, Cauchy and Lognormal diffusion models (Lemey *et al.,* 2010) and the most favourable model for the data set was then used for the phylogeographic analysis. To provide a spatial projection, we converted the MCC tree (with inferred posterior modal node locations and median node heights) into a keyhole markup language (KML) file generated by SPREAD version 1.0.6 (Bielejec *et al.,* 2011). The spatially continuous distribution was visualised in Google Earth.

Similarly as for the phylogenetic analysis, the final analyses were repeated with all samples available to us and a random sub set of 30 *B. penicillata ogilbyi* from Pacioni *et al.* (2011) were included. This analysis was run for 50 million generations (sampling every 5,000) and results were compared to those from the analysis with only unique sequences. This was done to control for the possible bias introduced by having a non-random selection of sequences and to (where possible) balance the sample size used for each species/clade. In using this framework, our intention was to evaluate two competing hypotheses. Hypothesis one: *Bettongia* populations colonised the continent and subsequently remained isolated (differentiation was therefore a function of local adaptation, in

part resulting from the degree of isolation). Hypothesis two: *Bettongia*'s evolutionary history followed a sequential geographic pattern, reflective of a progressive colonisation of the continent. In the case of hypothesis one, we would expect high dispersal rates in the deep phylogeny only, with further dispersal occurring only between localised, established lineages. Moreover, we would expect the diffusion rates to be relatively homogeneous through time, as opposed to hypothesis two, where we might observe "pulses" along the geographically directional radiation of the species that differentiated themselves during the process.

3.2.3.5 HIGH-THROUGHPUT SEQUENCING

In order to investigate whether using a new DNA isolation approach, targeting ultra short aDNA fragments (Dabney *et al.,* 2013), might be an effective approach for future bettong palaeogenetic studies, a pilot high-throughput sequencing (HTS) shotgun method was carried out on a single sample from "nailtail" cave in the Nullarbor region, that previously yielded mtDNA sequences. Single-stranded shotgun libraries were built on this sample, in a designated ultra-clean facility (TRACE) at Curtin University, WA, Australia, using the protocol as outlined in Gansauge and Meyer (2013) and purified using an Agencourt AMPure XP PCR purification kit (Beckman-Coulter) following manufacturer's instructions. The absolute concentration of the shotgun library was quantified via qPCR and run on an Illumina's MiSeq platform (150 cycle version 3) by following the manufacturer's instructions with minor modifications (see SI VII). After sequencing, the data were subject to bioinformatic analyses (see below). This sample was then prepared for mtDNA capture. The capture protocol was carried out using the MYselect Mycroarray kit combined with the protocol outlined in Li *et al.* (2013) with some modifications (see SI VIII for intergraded methods and detailed steps achieved for mtDNA capture and indexing). The captured library was sequenced on an Illumina MiSeq as previously described.

3.2.3.6 BIOINFORMATICS AND RAW SEQUENCING ANALYSIS

The FastQ files (raw data) generated from the Illumina MiSeq platform, were downloaded and imported into GENEIOUS version 7.1.8 for index separation and trimming (Murray *et al.* 2013). Sequences were thoroughly filtered (see SI IX) and taxon identification was achieved by aligning the filtered data sets against the NCBI's GenBank nucleotide reference database via BLASTn (SI IX). BLAST results were imported into MEGAN version 4.70.4 (ab.inf.uni-tuebingen.de/data/software/megan4; Huson *et al.,* 2007) so that they could be visualised (see SI IX

for parameters used). All metatherian DNA sequences were extracted from the MEGAN analysis and mapped to a reference genome (see SI IX for more details).

3.2.3.7 MAPPING TO A REFERENCE GENOME

Captured metatherian DNA sequences were extracted from MEGAN and imported back into GENEIOUS. Relevant information about the species identification, based on morphology, meant that appropriate mitochondrial genomes could be chosen as references and aligned to the captured mtDNA sequences. Alignments were performed in GENEIOUS, using the map to reference function, and increasing iterations up to 10 times to ensure the data were filtered appropriately, the final iteration was used for obtaining statistical data.

3.2.4 RESULTS

3.2.4.1 DNA SEQUENCING AND PHYLOGENETIC ANALYSIS

In this study, mtDNA sequences were obtained from 88 new (modern and historical) samples and analysed together with ~ 214 previously published bettong sequences (Pacioni *et al.,* 2011; Pacioni *et al.,* 2015) and two *B. tropica* Genbank sequences (AY237236 and AY237237). A total of 266 base pairs (bp) of cytochrome *b* and 346 bp of control region were sequenced, producing a concatenated mtDNA fragment of 612 bp. Only unique haplotypes were modelled to generate the phylogeny presented in Figure 3.2.2.

The Hasegawa, Kishino and Yano substitution model (Hasegawa *et al.,* 1985) and a *G* distribution were chosen for the cytochrome *b* analysis, whilst the control region data set required a *TN93* + *G* distribution. The comparison of tree priors revealed that BDIS was the most appropriate tree model for the data set, whilst the comparison of the molecular clock methods favoured a strict clock rather than any of the tested relaxed models.

The Bayesian phylogenetic analysis of the concatenated mtDNA alignment consistently produced five monophyletic clades (Figure 3.2.2, Clades I-V) that are supported by >95% posterior probabilities, with a clear phylogeographic signal. Clade I (individuals from the Nullarbor region); Clade II (Queensland populations, northeastern lineage, *B. tropica*); Clade III (Western Australian populations, southwestern lineage, *B. p ogilbyi*); Clade IV (Tasmanian, Victorian and New South Wales populations, southeastern lineage, *B. gaimardi*); and lastly, Clade V (Kangaroo Island

individual from South Australia, *B. penicillata*) (Figure 3.2.2). There is, however, a polytomy at the base of the brush-tailed bettongs, where there is very little consistency (i.e. low posterior support) for the branching topology between the clades (Figure 3.2.2). The taxonomic confusion in these bettongs is exemplified by the single Clade V sample (from KI) which was morphologically identified as *B. penicillata penicillata,* although it appears to be phylogenetically distinct from the other South Australian (Venus Bay) *B. penicillata* sample (Clade I). Clade V has 95% posterior support grouping it as sister to *B. gaimardi* (Clade IV, Figure 3.2.2). Based on the polytomy at the base of the brush-tailed bettongs (Figure 3.2.2) and ongoing uncertainties surrounding species boundaries (see later discussion) Clade V will be nominally referred to hereafter as belonging to *B. penicillata*, based solely on the original morphological identification.

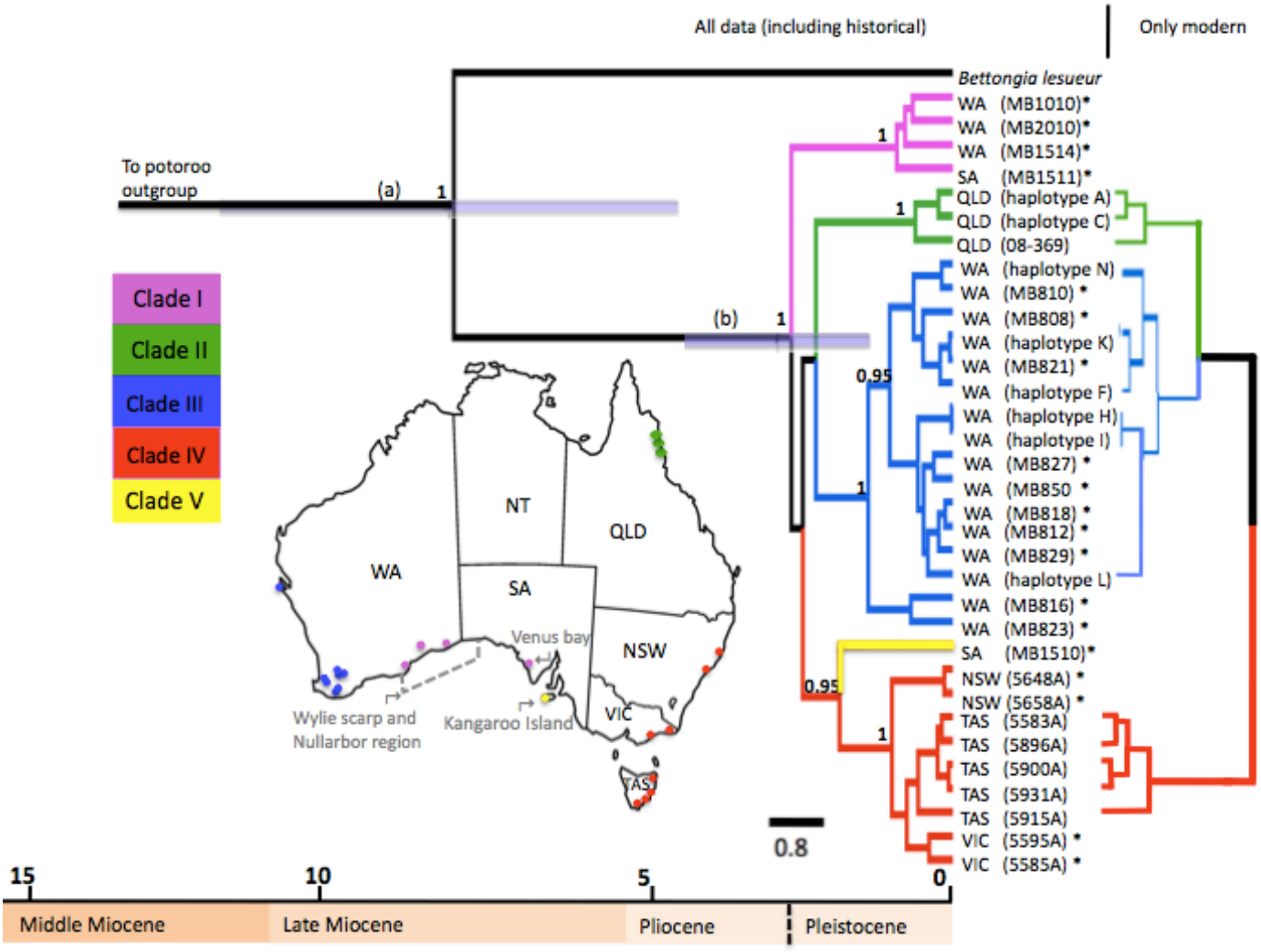


Figure 3.2.2 A concatenated mtDNA molecular phylogeny of cytochrome *b* (266 bp) and control region (346 bp) of brush-tailed *Bettongia* species. The phylogeny was constructed using Bayesian methods and the Maximum Clade Credibility tree is provided. Posterior probabilities >90% are shown on nodes. Tip labels marked with * indicate unique historical haplotypes (also see Table 3.2.1, for more information; highlighted sections refer to historical and fossil haplotypes / samples). Individual records are indicated on the map of Australia in reference to the colour codes on the tree, with five major clades represented. Small scale-bar represents nucleotide substitutions per site. Time scale represents millions of years. Phylogeny with all data, include only unique ancient haplotypes and unique modern haplotypes. Phylogeny with only modern unique haplotype data (on the right) is not to scale. The divergence date of the base of the genus *Bettongia* was calibrated from Westerman (2004), node (a), and the age of the brush-tailed species split was estimated to be 2.5 Ma, node (b).

The analysis with all samples available produced a similar tree topology with the exception that Clade V and *B. gaimardi* did not group as strongly (Figure SI II). The (GYMC) method identified six species in addition to *B. lesueur*. Five matched the five clades described above and the sixth was an internal clade to *B. penicillata* that included only two historical samples (MB816 and MB823, Figure SI III).

The mean number of differences between clades, excluding *B. lesueur*, ranged from 4.8-9.9 for cytochrome *b* and 21.7-44 for control region (Table 3.2.3). The calculated within-clade variability among Clades I-IV (i.e. where N > 1) for cytochrome *b* ranged from 2.5-4.67 (Table 3.2.3), while for the control region it was observed to be slightly higher within Clades II and III (range 9.33-18.5, Table 3.2.3).

Table 3.2.3 Pairwise distance and mean distances within clade, for bettong populations across two mtDNA genes. Clade numbers are in reference to those from Figure 3.2.2. Average differences and F_{ST} values are mentioned in the results section.

CYTOCHROME *b*	Clade I	Clade II	Clade III	Clade IV	Clade V	*B. lesueur*
Clade I	3.83	6.92	9.59	8.89	4.75	23.25
Clade II	0.39	4.67	9.58	9.04	5.67	24.33
Clade III	0.55	0.52	4.53	9.91	6.63	25.50
Clade IV	0.67	0.66	0.62	2.50	8.78	22.89
Clade V	0.19	0.18	0.32	0.72	0.00	21.00
B. lesueur	0.84	0.81	0.82	0.89	1	0.00
CONTROL REGION						
Clade I	18.50	43.17	42.84	29.36	42.25	87.75
Clade II	0.66	9.33	42.83	34.81	44.00	80.00
Clade III	0.57	0.62	18.38	31.72	40.75	85.88
Clade IV	0.53	0.69	0.51	11.44	21.67	69.78
Clade V	0.56	0.79	0.55	0.47	0.00	90.00
B. lesueur	0.79	0.88	0.79	0.84	1.00	0.00

Above diagonal: Average number of pairwise differences between clades (light grey shading)
Below diagonal: Pairwise FSTs between clades (no shading)
Diagonal elements: Average number of pairwise differences within clades (dark grey shading)

3.2.4.2 MOLECULAR DATING

Our temporal reconstruction using the concatenated mtDNA data set, including the calibration for the *Bettongia* radiation as reported by Westerman *et al.,* (2004), provided a divergence estimate for the last common ancestor of the extant species of *Bettongia* of approximately 7.7 Ma (95% Credibility Intervals (CI) 10.73 – 4.60 Ma; Figure 3.2.2) marking the separation of *B. lesueur.* The brush-tailed bettong radiation emerged later, around 2.5 Ma (CI 3.85 – 1.32 Ma; Figure 3.2.2).

3.2.4.3 PHYLOGEOGRAPHY OF *BETTONGIA*

A phylogeographic and spatial analysis for the *Bettongia* radiation was carried out to investigate the diffusion pattern of ancestral populations in the genus across the Australian continent. Based on the comparison of MLE, the Cauchy model of diffusion was favoured and it was implemented in the phylogeographic analysis. This demonstrated an initial, dispersal phase from the southern third of the continent followed by a more localised dispersal. No obvious directionality of the evolutionary process was identified (Figure SI IV) although the *B. gaimardi* radiation would appear to be slightly successive to the establishment of *B. penicillata* and the Nullarbor form. The analysis conducted with all available samples resulted in a similar pattern. However, given the larger number of samples, the uncertainty of the internal node locations was reduced and the successive radiation of *B. gaimardi* towards the south-east of Australia was more evident.

3.2.4.4 MITOCHONDRIAL DNA SHOTGUN LIBRARY AND CAPTURE

A shotgun library using a "nailtail" cave (historical) sample was built in order to check the practicality of finding nuclear markers. The results from this sample alone were successful, and a proportion of endogenous DNA could be amplified, producing segments of the mtDNA genome (pre-capture 1.122% of endogenous DNA sequences). Not surprisingly, the post-capture unique reads revealed the closest match when aligned to the *Bettongia penicillata* mtDNA (as opposed to other references, like *B. lesueur* — KJ868101, and *Aepyprymnus rufescens* — KJ868095) and resulted in a 5.4X coverage, of 89.6%, of the mtDNA genome (11,631 identical sites), with a pairwise identity of 96.2% and average fragment length of 54 bp (Figure 3.2.3).

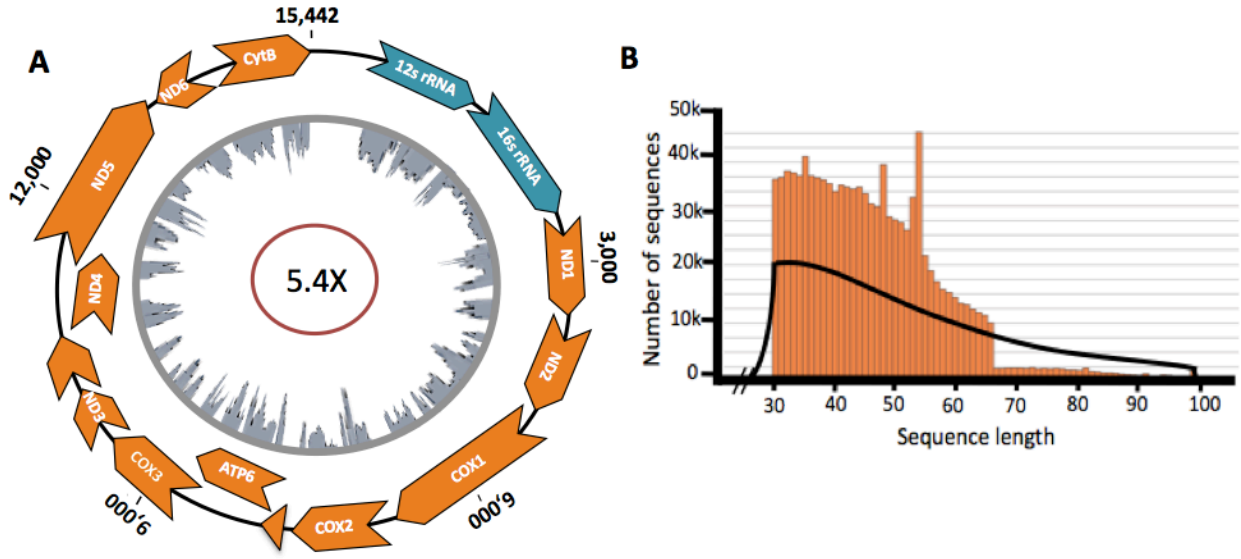


Figure 3.2.3 (A) Mitochondrial DNA sequence coverage (grey) of a historical bettong fossil from the Nullarbor Plain ("nailtail" cave), mapped to reference mtDNA genome (orange and blue arrows). (B) Length distribution of shotgun (black solid line) and mtDNA capture (orange bars). These data were generated using mtDNA capture baits and sequenced on an Illumina MiSeq platform.

3.2.5 DISCUSSION

European settlement has affected all Australian native mammals, some positively, but most negatively. Species of *Bettongia,* once among the most widespread marsupials on the Australian continent, are now some of the most severely restricted, persisting only in highly fragmented populations (Woinarski *et al.,* 2014). This study set out to explore the former species boundaries of members of the genus *Bettongia,* and their inter-relationships. Our study reveals that in order to capture the full extent of biodiversity loss and to assess the past evolutionary history of this genus, evidence from aDNA must be employed.

3.2.5.1 PHYLOGENY AND PHYLOGEOGRAPHY

The phylogeny generated by this study (Figure 3.2.2) has provided insights into the relationships within *Bettongia*, especially within the brush-tailed bettongs. When we explore the modern populations in isolation (Figure 3.2.2; see only modern representatives on right side of figure), three well-resolved mtDNA Clades (II, III and IV) are recovered that correspond to three currently accepted species, *B. penicillata, B. tropica* and *B. gaimardi*. The situation, however, becomes markedly more complex when the aDNA data are integrated into the phylogeny. The five mtDNA clades recovered from the combined ancient and modern samples have a clear phylogeographic signal (see map in Figure 3.2.2). The geographic affinities of the aDNA data, and congruence of the cytochrome *b* and control region phylogenies argue against influences of nuclear copies of mtDNA and contamination impacting on this data set. The presence of strong phylogeographic patterns in mtDNA is perhaps not surprising considering that bettongs display strong male-biased dispersal (Pacioni 2010; Pope *et al.,* 2012).

Within the extant Clades (II, III and IV), the inclusion of aDNA data clearly indicates loss of genetic diversity — this has been extensively explored within the Western Australian woylie (Pacioni *et al.,* 2015). Holocene fossil data show a single Clade (I) that extended across the Nullarbor region from at least the Wylie Scarp in Western Australia to Venus Bay in South Australia (Figure 3.2.2). The Nullarbor Plain has been considered an arid geographical barrier and suggested as the possible cause of separation between *B. p. ogilbyi* and *B. p. penicillata* over time. However, the shallow position of the base of Clade I and the lack of support for its internal nodes suggest that at least the Nullarbor form of *Bettongia* was adapted to arid conditions, and may have maintained gene flow across the Nullarbor.

When the genetic distinctiveness of Clade I was first demonstrated, one of the authors (A.B.) suggested that these small specimens might represent *Bettongia anhydra.* This was shown to be wrong when the detailed morphology and aDNA of the holotype were investigated, and *B. anhydra* was demonstrated to be a full species with affinities to *B. lesueur* (see Figure SI I) (McDowell *et al.,* 2015). Clade V is so far only represented by a single individual (MB1510) from Kelly Hill Cave, Kangaroo Island (Figure 3.2.2). This specimen was morphologically identified as *B. penicillata penicillata* (see results section 3.2.4.1 for a fuller explanation) across most tree topologies affiliates most closely with Clade IV (*B. gaimardi*), suggesting that the divergence between these forms is not as deep as it may seem from the phylogenetic tree constructed using modern data only.

The ability to spatially explore the past dynamics of species, using the continuous phylogeographic models implemented in the program BEAST (Lemey *et al.,* 2010), provides the opportunity to seek evidence of directionality in the bettongs' radiation. The key question here is whether the evolutionary history of the brush-tailed bettongs followed a geographic pattern: was the last common ancestor an eastern species that expanded west, or did one species spread out rapidly across the entire continent and populations then become progressively more isolated? Our analysis (together with the polytomy at the base of the brush-tail bettongs) suggests that prehistoric populations had a wide distribution in the southern third of the continent and that there is little signal for directionality in the radiation with the possible exception of the *B. gaimardi* deriving from south-central Australia bettongs. Increasing aridity throughout the Quaternary may have shaped the Holocene distribution of bettongs. Future studies targeting a wider geographic range of fossil samples, especially in central Australia, and the inclusion of nuclear markers (e.g. single nucleotide polymorphisms SNPs) may help establish a more robust evolutionary history of the taxa. However, poor DNA preservation (encountered in this and other studies) in these localities may limit the viability of these approaches.

3.2.5.2 MOLECULAR DATING AND DIVERGENCE EVENTS AMONG THE BETTONGS

Information on the timing of split-dates between clades can help elucidate the factors that drive speciation and provide a temporal context for taxonomic deliberations. Lack of informative bettong fossils (see methods) and the use of CR sequences that cannot be aligned with more distant macropods, necessitated the use of a calibration point taken from Westerman *et al.,* (2004). The *Bettongia* split from *Potorous* was dated in the late Oligocene (Westerman *et al.,* 2004). During this period, the sea withdrew from the Nullarbor region, exposing the surface to subaerial weathering

and erosion, conditions which continued for the following 10 Myr (Webb and James 2006; Hocknull *et al.,* 2007). Subsequent climate change and environmental modifications transformed the Australian landscape to what it is today (Hill *et al.,* 1999; Martin 2006; Hocknull *et al.,* 2007; Thomas 2010).

The split between *B. lesueur* and the brush-tailed bettongs is estimated to have occurred around the late Miocene about 7.7 Ma. Subsequent divergences within the genus (the base of Clades I-V) occurred in the late Pliocene or early Pleistocene about 2.5 Ma. During this time the Australian environment became drier and habitats became more fragmented (Fujioka *et al.,* 2005). These conditions continued until around 1 Ma (White 2001), probably reducing gene flow across Australia in many taxa (e.g. *Macropus fuliginosus,* Neaves *et al.,* 2009). However, the observed distribution of mitochondrial haplotypes suggests that in brush-tailed bettongs, matrilineal gene flow may have continued across southern Australia despite aridification of the Nullarbor region. In fact, while there is no evidence of maternal admixture between Clades I/III/IV/V during the last 2 Myr, differentiation of populations at the eastern and western ends of the Nullarbor region did not occur (i.e. these specimens group together in the strongly supported Clade I).

3.2.5.3 TAXONOMIC AND CONSERVATION IMPLICATIONS

The taxonomy of the genus *Bettongia,* like other members of the macropodid subfamily Potoroinae (e.g., *Potorous* and *Aepyprymnus*) has long been in a state of flux (Van Dyck and Strahan 2008). The morphological attributes initially used to separate members of this genus into species and subspecies were subtle. For example, from a molecular perspective, the data herein suggest there is little support (< 90% on the nodes – Figure 3.2.2) for the separation of the two subspecies of *B. gaimardi.* However, a mainland Australian subspecies (*B. gaimardi gaimardi*) and a Tasmanian subspecies (*B. g. cuniculus*) are currently recognised. Morphological attributes are also used to differentiate the critically endangered *B. penicillata ogilbyi* and extinct *B. penicillata penicillata* (Wakefield 1967), which have been the focus of most of the differences of opinion.

The management of conserved populations relies significantly on the ability to recognise a species as a unique biological unit in order to warrant protection and preservation (Fahrig and Merriam 1994). Holocene fossils can provide significant data regarding loss of genetic diversity and population fragmentation. The increased geographical sampling afforded by aDNA has clearly been instrumental in uncovering 'new' clades (i.e. Clades I and V) and several new haplotypes within Clades III and IV despite the paucity of fossil samples compared to modern data. The modern data,

interpreted in isolation, present a somewhat misleading picture of the distinctiveness of the brush-tailed species in the genus *Bettongia*. Our results suggest that the genetic differentiation we observe between species of brush-tailed bettongs results from localised drift. If so, the full spectrum of haplotypes could be recovered by sampling across their entire former distribution if the specimens existed. Results from the GMYC analysis (Figure SI III), illustrate the issue in delineating species boundaries where woylies (*B. penicillata ogilbyi*) assign as two distinct (but sympatric) species – where one putative species is comprised entirely from historical samples.

Our results clearly demonstrate that *B. penicillata* (Clades III and V), *B. gaimardi* (Clade IV), *B. tropica* (Clade II) and the Nullarbor form (Clade I) are genetically equidistant. However, we were unable to determine whether they should be considered species or subspecies. Generally, the longer populations are separated, the more substitutions they will accumulate, until eventually they become different enough to be considered different species (Bromham 2008). Notwithstanding this, recent recommendations favour the application of the biological species concept in conservation biology, or its extension: the differential fitness species concept (Frankham *et al.,* 2012). Therefore, the point at which two populations can no longer interbreed should clearly delineate divergent speciation (Mayr 2000). The status of the taxa that comprise genus *Potorous* has also been debated, complicating the conservation management decision-making process (Frankham *et al.,* 2012). Previous studies have suggested breeding potoroos from certain populations (i.e. southern mainland with northern mainland), to detect the occurrence of reproductive isolation (Johnston 1973). As for the *Potorous* species, brush-tailed bettongs may also be ideal for investigating interbreeding between currently recognised species from western and eastern populations. Such efforts are particularly important for restoring genetic diversity, especially in cases like the woylie in southwestern Australia, currently suffering major population decline that may be linked to disease (Wayne *et al.,* 2015).

3.2.5.4 POTENTIAL OF CONSERVATION GENOMICS AND PALAEOGENOMICS

Ultimately, many of the conservation and taxonomic questions raised above will be difficult to resolve definitively using partial mtDNA cytochrome *b* and control region sequences. To explore the viability of a palaeogenomic approach we conducted shotgun DNA sequencing (using single-stranded library build) and MyBait mtDNA capture on one of the Holocene fossil bones, which exhibited good DNA preservation (determined by qPCR). The successful reconstruction of a mtDNA genome (5.4X coverage, see Figure 3.2.3) and credible hits to macropodid nuclear genes (i.e. the published genome of *Macropus eugenii*), suggest that aDNA approaches may assist further

in understanding maternal and paternal gene flow across Australia. The average fragment length of aDNA retrieved from this fossil is 54 bp which largely rules out amplicon-based approaches to typing SNPs.

Conventional genetic studies mainly focus on a small number of markers in many individuals. More recently, coalescent approaches (i.e. the suite of probabilistic models that link genetic diversity to demographic changes) have demonstrated the power of multi-locus analyses, relaxing the need for large sample sizes (Wakely 2008). An area in which genomic data will be particularly useful is in attempts to estimate past demographics such as population size fluctuations (McMahon *et al.,* 2014). The genomic data from even a single individual can provide useful insights into an endangered species' past population dynamics (Shafer *et al.,* 2015). The natural extension to this project is to develop a set of SNPs (e.g. RADseq) from extant populations and then apply DNA hybridisation enrichment and capture techniques in order to isolate these targets (i.e. SNPs) in fossil samples.

3.2.6 CONCLUSIONS

Although the bettongs have had a complex taxonomic history, the inclusion of aDNA has enabled exploration of the evolution and interrelatedness of species within this group. Our results have uncovered a distinct 'new' clade that once occurred in central southern Australia, challenging the concept of aridity as a barrier to gene flow. Mitochondrial DNA is more readily accessible in degraded samples like those used in this study and it is often used to elucidate the evolutionary history of different species (Patwardhan *et al.,* 2014), but the information gathered is limited by its restriction to matriarchal lineages. This is an important factor because the brush-tailed members of the genus *Bettongia* exhibit strong sex-biased dispersal where female movement is restricted. Therefore, dispersal behaviour may have a substantial effect on the phylogeographic structure obtained from mtDNA, as has been recently demonstrated for *B. penicillata ogilbyi* (Pacioni *et al.,* 2015). Together with short generation times, historical loss of biodiversity, extinctions, fragmentation and possible hybridisation/introgression and differential dispersal, can result in a misleading interpretation of the ancient mtDNA data presented here. The solution may lie in including nuclear markers in these analyses. The results from our shotgun library and mtDNA enrichment suggest that a combination of genomics and palaeogenomics is a viable approach. Genome-level data will enable historical gene flow amongst the clades/species of brush-tailed bettongs to be formally tested.

Clearly there are major differences between how *Bettongia* species are defined using modern and aDNA samples. This conflict warrants further investigation, to refine the taxonomic revision of this highly endangered group. The four recognised taxonomic units, three currently classified as different species and one as a sub-species, are genetically equidistant in the mtDNA loci studied, suggesting that they represent either one highly diverse or four very similar species. Furthermore, the separation of the two *B. gaimardi* subspecies within Clade IV does not have strong support in our analysis, suggesting that a more thorough assessment of their status is warranted. Our work has assisted in clarifying the evolutionary history of bettongs across Australia and should act as a catalyst to help resolve lingering questions regarding gene flow and speciation, an understanding of which is crucial for future conservation and best-practice management of the surviving members of the genus *Bettongia*, all of which are endangered.

3.2.7 ACKNOWLEDGEMENTS

Thanks to the WA Museum, Museum of Victoria and the Australian National Wildlife Collection (CSIRO) for access to historical samples and assistance. Work was funded by Australian Research Council grants: FT0991741 (MB) DP120104435 (Gavin Prideaux and MB). We thank Gavin Prideaux (Flinders University) for helpful discussions and the Pawsey Supercomputer Centre, Western Australia, and CIPRES (http://www.phylo.org/sub_sections/portal/cite.php) for providing computational support.

3.2.8 REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman (1990). Basic local alingment search tool. *Journal of Molecular Biology* **215**, 403-410.

Baele, G., P. Lemey, T. Bedford, A. Rambaut, M. A. Suchard & A. V. Alekseyenko (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* **29**, 2157-2167.

Baynes, A. (1987). The original mammal fauna of the Nullarbor and southern peripheral regions: evidence from skeletal remains in superficial cave deposits, In: McKenzie, N.L., Robinson, A.C. (eds). A biological survey of the Nullarbor region South and Western Australia in 1984. South Australian Department of Environment and Planning, Adelaide, pp139-152.

Bielejec, F., A. Rambaut, M. A. Suchard & P. Lemey (2011). SPREAD: Spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27**, 2910-2912.

Bromham, L. (2008). *Reading the story in DNA; a beginner's guide to molecular evolution,* New York., Oxford University Press.

Bunce, M., C. L. Oskam & M. E. Allentoft (2012). Quantitative real-time PCR in aDNA research. In: Shapiro, B. Hofreiter, M (eds). Ancient DNA. Humana Press, New York, pp. 121-132.

Burbidge, A. A., K. A. Johnson, P. J. Fuller & R. I. Southgate (1988). Aboriginal knowledge of the mammals of the central deserts of Australia. *Australian Wildlife Research* **15**, 9-39.

Calaby, J. H. & B. J. Richardson (1988). Potoroidae. In: Walton, D.W., ed. Zoological Catalogue of Australia. Volume 5 Mammalia, Pp. 53-59. Australian Government Publishing Service, Canberra.

Christensen, P. (1995). Brush-tailed bettong *Bettongia penicillata* Gray, 1837. In: Strahan. R (ed) *The mammals of Australia.* Reed Books, Chatswood, pp. 292-293.

Christensen, P. E. S (1980). The biology of *Bettongia penicillata Gray, 1837, and Macropus eugenii (Desmarest, 1817)* in relation to fire. *Forests Department of Western Australia Bulletin No. 91, i-ix, 1-90*.

Cooke, B. N., K. J. Travouillon, M. Archer & S. J. Hand (2015). Ganguroo robustiter, sp. nov. (Macropodoidea, Marsupialia), a middle to early late Miocene basal macropodid from Riversleigh World Heritage Area, Australia. *Journal of Vertebrate Paleontology* **35**, e956879.

Dabney, J., M. Knapp, I. Glocke, M. T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. Garcia, S. Paabo, J. Arsuaga & M. Meyer (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15758-15763.

Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer & A. Wilson (2011). Geneious v 6.0.1. Available from http://www.geneious.com/. Accessed 15 Feb 2016.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips & A. Rambaut (2006). Relaxed phylogenetic dating with confidence. *PLoS Biology* **4**, 88.

Drummond, A. J. & A. Rambaut (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.

Drummond, A. J., M. A. Suchard, D. Xie & A. Rambaut (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969-1973.

Drummond, F. H. (1933). The male meiotic phase in five species of marsupials. *Quarterly Journal of Microscopical Science* **76**, 1-11.

Eldridge, M. D. B., M. P. Piggott & S. L. Hazlitt (2012). Population genetic studies of the Macropodoidea: a review. In Coulson, G and Eldridge M. D. B (eds) Macropods: the biology of kangaroos, wallabies and rat-kangaroos. CSIRO Publishing, Melbourne. pp 35-51.

Fahrig, L. & G. Merriam (1994). Conservation of fragmented populations. *Conservation Biology* **8**, 50-59.

Finlayson, H. (1958). On Central Australian mammals (with notice of related species from adjacent tracts). Part III — The Potoroinae. *Records of the South Australian Museum* **13**, 235-302.

Frankham, G. J., K. A. Handasyde & M. D. Eldridge (2012). Novel insights into the phylogenetic relationships of the endangered marsupial genus *Potorous*. *Molecular Phylogenetics and Evolution* **64**, 592-602.

Fujioka, T., J. Chappell, M. Honda, I. Yatsevich, K. Fifield & D. Fabel (2005). Global cooling initiated stony deserts in central Australia 2-4 Ma, dated by cosmogenic ^{21}Ne-^{10}Be. *Geology* **33**, 993-996.

Gansauge, M.-T. & M. Meyer (2013). Single-stranded DNA library preparation for the sequencing of ancinet or damaged DNA. *Nature Protocols* **8**, 737-748.

Gernhard, T. (2008). The conditioned reconstructed process. *Journal of Theoretical Biology* **253**, 769-778.

Haouchar, D., J. Haile, M. C. McDowell, D. C. Murray, N. E. White, R. J. N. Allcock, M. J. Phillips, G. J. Prideaux & M. Bunce (2013). Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late Pleistocene-Holocene cave deposit on Kangaroo Island, South Australia. *Quaternary Science Reviews* **84**, 56-64.

Hasegawa, M., H. Kishino & T.-a. Yano (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160-174.

Hill, R. S., E. M. Truswell, S. McLoughlin & M. E. Dettmann (1999). Evolution of the Australian flora: fossil evidence. In: Orchard, A. E (ed) Flora of Australia, vol. 1, 2nd ed. ABRS/CSIRO, Australia, pp. 251-320).

Hocknull, S. A., J.-x. Zhao, Y.-x. Feng & G. E. Webb (2007). Responses of Quaternary rainforest vertebrates to climte change in Australia. *Earth and Planetary Science Letters* **264**, 317-331.

Hunter, A. A., A. B. Macgregor, T. O. Szabo, C. A. Wellington & M. I. Bellgard (2012). Yabi: An online research environment for grid, high performance and cloud computing. *Source Code for Biology and Medicine* **7**, e1.

Huson, D. H., A. Auch, J. Qi & S. C. Schuster (2007). MEGAN analysis of metagenome data. *Genome Research* **17**, 377-386.

Johnston, P. G. (1973). Variation in island and mainland populations of *Potorous tridactylus* and *Macropus rufogriseus* (Marsupialia). Ph.D. Thesis. University of New South Wales, Sydney.

Laurance, W. F. (1997). A distributional survey and habitat model for the endangered northern bettong *Bettongia tropica* in tropical Queensland. *Biological Conservation* **82**, 47-60.

Lemey, P., A. Rambaut, J. J. Welch & M. A. Suchard (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution* **27**, 1877-1885.

Li, C., M. Hofreiter, N. Straube, S. Corrigan & G. J. P. Naylor (2013). Capturing protien-coding genes across highly divergent species. *BioTechniques* **54**, 321-326.

Martin, H. A. (2006). Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments* **66**, 533-563.

Mayr, E. (2000). The biological species concept. In Q. Wheeler and R. Meier (eds). *Species concepts and phylogenetic theory: a debate. Columbia University Press, New York*, 17-29.

McDowell, M. C. (1997). Taphonomy and palaeoenvironmental interpretation of a late Holocene deposit from Black's Point Sinkhole, Venus Bay, S.A. *Proceedings of the Linnean Society of New South Wales* **117**, 79-95.

McDowell, M. C., D. Haouchar, K. P. Aplin, M. Bunce, A. Baynes & G. J. Prideaux (2015). Morphological and molecular evidence supports specific recognition of the recently extinct *Bettongia anhydra* (Marsupialia: Macropodidae). *Journal of Mammalogy* **96**, 287-296.

McIllwee, A. P. & C. N. Johnson (1998). The contribution of fungus to the diets of three mycophagous marsupials in Eucalyptus forests, revealed by stable isotope analysis. *Functional Ecology* **12**, 223-231.

McMahon, B. J., E. C. Teeling & J. Hoglund (2014). How and why should we implement genomics into conservation? *Evolutionary Applications* **7**, 999-1007.

McNamara, J. A. (1997). Some smaller macropod fossils of South Australia. *Proceedings of the Linnean Society of New South Wales* **117**, 97-106.

Menkhorst, P. W. (2001). *A field guide to the mammals of Australia,* South Melbourne., Oxford University Press.

Murray, D. C., J. Haile, J. Dortch, N. White, D. Haouchar, M. I. Bellgard, R. Allcock, J, G. J. Prideaux & M. Bunce (2013). Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports* **3**, 3371.

Neaves, L. E., K. R. Zenger, R. I. T. Prince, M. D. B. Eldridge & D. W. Cooper (2009). Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, *Macropus fuliginosus*. *Molecular Ecology* **18**, 3363-3378.

Nei, M. & W. H. Li (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of National Academy of Sciences of the United States of America* **76**, 5269-5273.

Pacioni, C. (2010). The population and epidemiological dynamics associated with recent decline of woylies (*Bettongia penicillata*) in Australia. Ph.D. thesis. School of Veterinary and Life Sciences. Murdoch University, Western Australia.

Pacioni, C., H. Hunt, M. E. Allentoft, T. G. Vaughan, A. F. Wayne, A. Baynes, D. Haouchar, J. Dortch & M. Bunce (2015). Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. *Molecular Ecology* **24**, 5813-5828.

Pacioni, C., A. F. Wayne & P. B. S. Spencer (2011). Effects of habitat fragmentation on population structure and long-distance gene flow in an endangered marsupial: the woylie. *Journal of Zoology* **283**, 98-107.

Patwardhan, A., S. Ray & R. Amit (2014). Molecular markers in phylogenetic studies - A review. *Phylogenetics and Evolutionary Biology* **2**, 131-139.

Pizzuto, T. A., G. R. Finlayson, M. S. Crowther & C. R. Dickman (2007). Microhabitat use by the brush-tailed bettong (*Bettongia penicillata*) and burrowing bettong (*B. lesueur)* in semiarid New South Wales: implications for reintroduction programs. *Wildlife Research* **34**, 271-279.

Pope, L. C., A. Estoup & C. Moritz (2000). Phylogeography and population structure of an ecotonal marsupial, *Bettongia tropica*, determined using mtDNA and microsatellites. *Molecular Ecology* **9**, 2041-2053.

Pope, L. C., K. Vernes, A. W. Goldizen & C. N. Johnson (2012). Mating system and local dispersal patterns of an endangered potoroid, the northern bettong (*Bettongia tropica*). *Australian Journal of Zoology* **60**, 278-287.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Priddel, D. & R. Wheeler (2004). An experimental translocation of brush-tailed bettongs (*Bettongia penicillata)* to western New South Wales. *Wildlife Research* **31**, 421-432.

Rambaut, A. & A. Drummond (2009). Tracer version 1.5.0 <http://tree.bio.ed.ac.uk/software/tracer/>). Accessed 15 Feb 2016.

Reynolds, J., B. S. Weir & C. C. Cockerham (1983). Estimation for the coancestry coefficient: basis for short-term genetic distance. *Genetics* **105**, 767-779.

Ride, W. D. L. (1970). *A guide to the native mammals of Australia,* Oxford University Press, Melbourne.

Rose, R. (1986). The habitat, distribution and conservation status of the Tasmanian bettong, *Bettongia gaimardi* (Desmarest). *Australian Wildlife Research* **13**, 1-6.

Sampson, J. (1971). The biology of *Bettongia penicillata* Gray, 1837. Ph.D. thesis. University of Western Australia.

Shafer, A. B., J. B. Wolf, P. C. Alves, L. Bergstrom, M. W. Bruford, I. Brannstrom, G. Colling, L. Dalen, L. De Meester & R. Ekblom (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution* **30**, 78-87.

Sharman, G. B. (1961). The mitotic chromosomes of marsupials and their bearing on taxonomy and phylogeny. *Australian Journal of Zoology* **9**, 38-60.

Sharman, G. B., C. E. Murtagh, P. M. Johnson & C. M. Weaver (1980). The chromosomes of a rat-kangaroo attributable to *Bettongia tropica* (Marsupialia: Macropodidae). *Australian Journal of Zoology* **28**, 59-63.

Stadler, T. (2009). On incomplete sampling under birth-death models and connections to the sampling-based coalescent. *Journal of Theoretical Biology* **261**, 58-66.

Strahan, R. (1983). *Complete book of Australian mammals,* Sydney, Angus & Robertson.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437-460.

Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution* **13**, 283-285.

Thomas, C. D. (2010). Climate, climate change and range boundaries. *Diversity and Distributions* **16**, 488-495.

Van Dyck, S. & R. Strahan (2008). *The mammals of Australia*, Sydney, Reed New Holland.

Wakefield, N. A. (1967). Some taxonomic revision in the Australian marsupial genus *Bettongia* (Macropodidae), with description of a new species. *The Victorian Naturalist* **84**, 8-22.

Wakely, J. (2008). *Coalescent theory - an introduction.* Roberts and Company, Greenwood Village.

Wayne, A. F., M. A. Maxwell, C. G. Ward, C. V. Vellios, I. Wilson, J. C. Wayne & M. R. Williams (2015). Sudden and rapid decline of the abundant marsupial *Bettongia penicillata* in Australia. *Oryx* **49**, 175-185.

Webb, J. A. & J. M. James (2006). Karst evolution of the Nullarbor Plain, Australia. *Geological Society of America Special Papers* **404**, 65-78.

Westerman, M., S. Loke & M. S. Springer (2004). Molecular phylogenetic relationships of two extinct potoroid marsupials, *Potorous platyops* and *Caloprymnus campestris* (Potoroinae: Marsupialia). *Molecular Phylogenetics and Evolution* **31**, 476-485.

White, N. (2001). Karst in arid Australia. *National cave and karst management symposium*, pp. 109-113, Victoria, Australia.

Willerslev, E. & A. Cooper (2005). Ancient DNA. *Proceedings of the Royal Society of London B: Biological Sciences* **272**, 3-16.

Woinarski, J. C. Z., A. A. Burbidge & P. L. Harrison (2014). *The action plan for Australian mammals 2012*. CSIRO Publishing, Melbourne.

Woinarski, J. C. Z., A. A. Burbidge & P. L. Harrison (2015). Ongoing unraveling of a continental fauna: decline and extinction of Australian mammals since European settlement. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4531-4540.

Xie, W., P. O. Lewis, Y. Fan, L. Kuo & M.-H. Chen (2011). Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* **60**, 150-160.

Yule, G. U. (1925). A mathematical theory of evolution, based on the conclusions of dr. jc willis, frs. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **213**, 21-87.

3.2.9 SUPPLEMENTARY INFORMATION

SI I *Implementation of standard aDNA protocols*

Bones were kept in airtight containers at 4°C. Samples were prepared in an ancient DNA clean room preparation area. A full body suit, double gloves, mask, eyewear, boots, and a hood were worn as personal protective equipment, and to avoid contaminating the bone samples with human DNA. Before beginning, all surfaces (including tray, table, chair, electronic balance, and gloves) were cleaned with a solution of 10% bleach, followed by a solution of 70% ethanol (Willerslev *et al.*, 2005).

SI II *Bone digest protocol*

The digestion step included an overnight incubation in a digest oven, at 55°C with rotation in a 1.5 mL digestion buffer. The digest contained 20mM Tris pH 8.0 (Sigma, MO, USA), 10mM dithiothreitol (Thermo Fisher Scientific, MA, USA), 1 mgmL–1 proteinase K powder (Amresco, OH, USA), 0.48 M EDTA (EDTA) (Invitrogen) and 1% Triton X-100 (Invitrogen). After digestion, centrifugation at 13 000g was initiated for 1 min to pellet undigested material. The supernatant containing the DNA was concentrated to ~100 mL in a Vivaspin 500 column (MWCO30000; Sartorius Stedim Biotech, Germany) at 13 000g, and then combined with five volumes of PBi buffer (Qiagen, CA, USA). DNA was immobilised on silica spin columns (Qiagen) and washed with 700 mL of AW1 and AW2 wash buffers (Qiagen). Finally, the DNA was eluted from the silica in 50 mL of 10mM Tris pH 8.0 (Sigma, MO, USA).

SI III *Amplicon sequencing*

A unique index was used to eliminate the possibility of sequence contamination arising from previously amplified DNA. Amplicons were purified using an Agencourt AMPure XP PCR purification kit (Beckman-Coulter) at a ratio of 1.2 ul of AMPure XP SPRI beads to 1 µl of PCR product and eluted in 40 µl EB (QIAGEN). Of this eluant, 5 µl of genomic DNA from the sample was combined with 0.5 µl of 6x loading dye and run on a 3% agarose gel electrophoresis (3.3 g agarose, 110 mL TAE buffer, 8 ul GelRed). The gel was visualized and photographed using a BioRad transilluminator. The amplicons (from this study) were then pooled (i.e. pooled along with samples from other research) in approximate equimolar amounts based on the intensity of the bands on the gel - these results however do not provide any information about the quality and quantity of

endogenous DNA, for this, we needed to assess the amplification of endogenous DNA through qPCR and subsequent sequencing. The total concentration of the sequencing library was quantified via qPCR (SI IV), in order to quantify the number of template molecules in the final sequencing reaction. CT-values were recorded and compared to the standards in order to calculate the number of copies in each dilution, and determine the volume of library to input into the sequencing reaction (SI IV). Unidirectional sequencing was performed on Illumina's MiSeq platform by following the manufacturer's instructions for the MiSeq 300 V2 Nano kit using a custom sequencing primer.

Table SI I
Primers used for the initial qPCR test pre-amplicon sequencing, and further used for amplicon sequencing as the gene specific primers. The architecture of the fusion tag indexed primers are as follows: 5' - MiSeq P5 – 5' Sequencing Adapter – Index – Forward Gene Specific Primer– 3' (forward), and 5' – MiSeq P7 – Index – Reverse Gene Specific Primer -3'(reverse).

Name	Sequence (5'-3')	Target taxa	Gene	Amplicon size (bp)	Annealing Temp (°C)	Reference
16sMam1F	CGGTTGGGGTGACCTCGGA	Marsupial (universal)	16S	150	54	(Taylor, 1996)
16sMam1R	GCTGTTATCCCTAGGGTAACT	Marsupial (universal)	16S		54	(Taylor, 1996)
12sMarsMini (Forward)	TAGTTAGACCTACACATGCAAGTT	Marsupial (universal)	12S	85	55	This study
12sMarsMini (Reverse)	CCTGATACCCGCTCCTRTTR	Marsupial (universal)	12S		55	This study

***SI IV** PCR reaction and thermocycling conditions for quantifying the sequencing library*

All assays in this study were carried out qPCR. This procedure provides a way to assess the relative quantity and quality of the amplifiable DNA from each single extraction by measuring the accumalation of DNA through fluorescence as it amplifies exponentially in the PCR reaction. The The amount of starting template molecules in an extract relative to another can be estimated by comparison of the CT-values, the cycle at which DNA is amplified beyond a certain threshold of detection. The CT-value is inversely related to the number of starting template molecules in an extract—the lower the CT-value, the greater the number of starting template molecules. This means of analysis was closely observed in each qPCR reaction, and was useful in detecting whether PCR inhibition was a factor to unsuccessful samples, due to the presence of contaminants.

The sequencing library was diluted 1/10, 1/100, 1/1000, 1/5000, 1/25000, 1/125000, 1/625000 in EB buffer (QIAGEN). Each dilution was qPCR amplified in duplicate along side a standard of known concentration (10^8, 10^6, 10^5, 10^4, 10^3, 10^2 molecules) using primers complementary to the sequencing adapters, in order to quantify the number of template molecules in the final sequencing library. PCR 'no-template controls' were included. The PCR reaction contained reagents in final concentrations of: 1X ABI Power SYBR Master Mix, 0.4 µM IDT forward primer P5, 0.4 µM IDT reverse primer P7 (Table AI IV), and 2 µl of library in a total reaction of volume of 25 µl (including 8.5 µl GIBCO HPLC-grade water). Thermocycling conditions were: 95°C for 5 min, 40 cycles of 95°C for 30 sec, 60°C for 45 sec, followed by a melt curve of 95°C for 15 sec, 60°C for 1 min, and 95°C for 15 sec. CT values were recorded and compared to the standard in order to calculate the number of copies in each dilution, and determine the volume of library to input into the sequencing reaction.

Table SI II Primers used for quantifying the final sequence library by qPCR

Name	Sequence (5'-3')	Target	Annealing Temp (°C)
P5	AATGATACGGCGACCACCGAGATCTACAC	Forward qPCR quant primer	60
P7	CAAGCAGAAGACGGCATACGAGAT	Reverse qPCR quant primer	60

SI V *Modifications to the single-stranded library building method by (Gansauge & Meyer 2013)*

Modifications to the adapters used are listed in Table SI III. An extraction control, template (water) control, and CL104 positive control were also included in the library building process. At step 1, reactions were performed in 0.2 ml 8-well PCR strip tubes. 12 ul of DNA extract was used, and Afu UDG was replaced by Ultra-pure water. At step 7, ligation products were stored overnight at -20°C. At step 13, tubes were incubated in a rotating hybridisation oven for 2 min at 65°C as opposed to a thermal shaker. At step 13, tubes were transferred to a thermal shaker pre-cooled to 15°C as opposed to a thermocycler. Steps 14, 15, 18, 19, 23, and 25 were performed in a thermal shaker. Step 22 was performed in a rotating hybridisation oven. At step 25, the supernatant was stored in a 1.5 mL Lo-Bind Eppendorf tube at -20°C. After step 28, the PCR products were run on a 2% agarose gel electrophoresis in order to confirm the library building process worked.

Table SI III
Ligation adapters and sequencing primers used for shot-gun sequencing (modifications from Gansauge *et al.*, 2013) The architecture of the fusion-tag indexed primers is: 5' - MiSeq P5 – RD1–Index—5' SS adapter -3' (forward), and 5' - MiSeq P7—RD2—Index—3' SS adapter - 3' (reverse). Unique indexes are represented by NNNNNNN (any base) below.

Name	Sequence (5'-3')	Function	Annealing Temp (°C)
CL53	ACACGACGCTCTTC-ddC	Double-stranded adapter, strand 1	RT
CL78	[Phosphate]AGATCGGAAG[C9Spacer]3[TEG-biotin]	Single-stranded adapter	60
CL105_106_Std	ACACTCTTTCCCTACACGACGCTCTTCCTCGTCGTTTGGTATGGCTTCTA TCGUATCGATCGATCGACGATCAAGGCGAGTTACATGAAGATCGGAA GAGCACACGTCTGAACTCCAGTCAC	Synthetic qPCR standard	-
P5-RD1-Index-Fwd adapter	AATGATACGGCGACCACCGAGATCTACAC-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-NNNNNNN ACACTCTTTCCCTACACGACGCTCTT	Forward indexing primer	60
P7-RD2-Index-Rev adapter	CAAGCAGAAGACGGCATACGAGAT-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-NNNNNNN-GTGACTGGAGTTCAGACGTGT	Reverse indexing primer	60
RD1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	Standard sequencing primer	-

SI VI *PCR reaction and thermo cycling conditions for the shot-gun library fusion tag/indexing PCR*

At step 30 (Gansauge & Meyer 2013), the libraries were amplified in quintuplicate with unique fusion-tag indexing primers (Table SI III). The PCR reaction contained reagents in final concentrations of: 1X ABI Power SYBR Master Mix, 0.4 µM forward indexing primer, 0.4 µM reverse indexing primer, 1 µl of neat library, made up to a total of 25 µl final volume with HPLC-grade water. Thermocycling conditions were: 95°C for 2 min, 26 cycles of 95°C for 15 sec, 60°C for 30 sec, 72°C for 1 min. Replicate reactions were combined and purified using the AMPure XP PCR purification kit (Beckman-Coulter). Step 33 was not performed. Libraries were quantified via qPCR as per SI IV. Libraries were pooled in equal amounts and the final shot-gun sequencing library was quantified once again via qPCR as per SI IV.

SI VII *Modifications to the MiSeq 150 version 3 kit*

At step 34 (Gansauge & Meyer 2013), a standard sequencing primer was used (Table SI II). Once the approximate amount of input library template molecules were targeted, it was prepared for sequencing using a MiSeq 150 version 3 kit. The target concentration and volume of the library was combined with 12 µl of EB buffer (QIAGEN), and 2 µl of 1M molecular biology-grade NaOH and incubated for 5 minutes at 25°C, then placed on ice. 10 µl of this mixture was then added to 990 µl of HT1 buffer, and placed on ice. 550 µl of this mixture was then combined with 50 µl of 20 pM denatured PhiX, and placed on ice. 600 µl of this mixture was added to the reagent cartridge in slot 17. The options selected when creating the sample sheet were: Other/ FastQ only/ Sample preparation kit = TruSeq LT/ No index reads/ Single end/ 175 cycles/ No custom primer/ No trimming.

SI VIII *Mitochondrial DNA capture and indexing*

The capture protocol was carried out using a combination of methods; the MYselect Mycroarray kit combined with the protocol outlined in Li *et al.* (2013) with some modifications. These modifications included the following; once the hybridised libraries (the biotinylated target-bait complex) were bound to the streptavidin-coated beads, and washed accordingly so that unbound material and bound non-target DNA can be removed (Li *et al.*, 2013), a post-hybridisation indexing PCR (off-beads amplification) was carried out (Li *et al.*, 2013). This off-beads amplification ensures that the captured target DNA is amplified off the target-bait-bead complex during the indexing PCR, therefore the beads are still present during the amplification step (Li *et al.*, 2013).
The indexing PCR was carried out in duplicates for 18 cycles (Li *et al.*, 2013). The replicates of each sample were combined and the PCR products were cleaned using the AMPure XP kit. A dilution series was carried out and then quanted with set standards (as described in SI IV). The qPCR CT-values were recorded and used to calculate the number copies in the reaction and sequenced on the Illumina MiSeq platform (150 version 3 kit). The raw data produced was sorted as detailed below.

SI IX *Bioinformatics and raw sequence analysis*

The raw data, from the FastQ file generated from the Illumina MiSeq platform were downloaded and imported into Geneious version 7.1.8 for index separation and trimming (Murray *et al.*, 2013). Sequences that did not perfectly match the expected index combinations were discarded from

further analyses. Adapter sequences, forward and reverse primers were also removed. Abundance filtering was not performed on shot-gun data, as PCR amplification was stopped during the linear phase resulting in few clonal copies. Sequences shorter then 30 bases were discarded and only unique reads were accepted. Taxonomic identification was achieved by aligning the filtered data se ts against the NCBI's GenBank nucleotide reference database via BLASTn (-F No, -e 0.01, -m Pairwise, -v 20, -b 20) (Altschul *et al.*, 1990), with searches executed in YABI (ccg.murdoch.edu.au/yabi; Hunter *et al.*, 2012). BLAST results were imported into MEGAN version 4.70.4 (ab.inf.uni-tuebingen.de/data/software/megan4; Huson *et al.*, 2007) so that they can be visualised. The LCA parameters used in MEGAN were: min support of 1, min score of 35, top per cent of 10, min complexity 0.44, win score 0. Identifications were based on the percent sequence similarity of the query to the reference across > 80% of the query, in the BLAST hits of the terminal nodes of the MEGAN tree. All sequences were categorised into the following, metatherian, human and bacterial DNA, other and no hits. All metatherian DNA sequences were extracted from the MEGAN analysis and mapped to a reference genome.

SI X *Modifications to BEAUTi xml block, to implement phylogeography*

We added an additional statistic to the analysis by editing the xml. This statistic, the treeDispersionStatistic, keeps track of the rate of diffusion by measuring the distance covered along each branch (based on the spatial coordinates inferred at the parent and descendent node of each branch), summing this distance for the complete tree and dividing this by the tree length (Lemey *et al.*, 2010).

Table SI IV. Final marginal likelihood values using path sampling and stepping stone. The ranks of the substitution models listed for each gene, generated using BEAST version 1.7.4.

Cytochrome *b*		
Model	Path sampling	Stepping stone
HKY+G	-1172.793075	-1173.002507
TN93+I+G	-1176.132128	-1176.377385
TN93+G	-1176.71563	-1176.948257
HKY+I+G	-1177.880425	-1178.115091
GTR+I+G	-1186.811414	-1187.067436
GTR+G	-1188.600046	-1188.881962
TN93	-1207.304111	-1207.497056
HKY	-1213.630519	-1213.833047
GTR	-1220.028565	-1220.276293

Control region		
Model	Path sampling	Stepping stone
GTR+G	-2399.292242	-2399.657896
TN93+G	-2399.297299	-2399.65638
HKY+G	-2400.336337	-2400.698
HKY+I+G	-2401.226781	-2401.577936
GTR+I+G	-2404.108107	-2403.727282
TN93+I+G	-2411.161992	-2411.491306
TN93	-2522.161155	-2522.552839
HKY	-2523.398957	-2523.745517
GTR	-2529.64984	-2530.001208

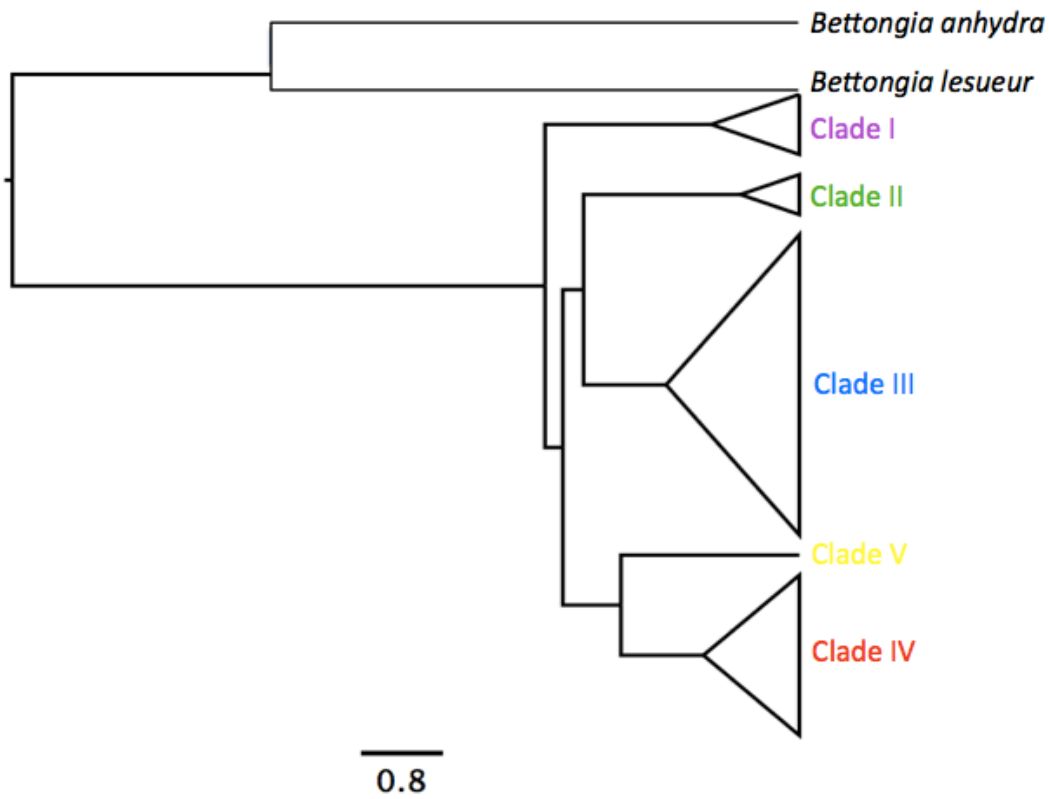


Figure SI I. Simplified tree of Figure 3.2.2 (found in manuscript) with collapsed tips for each brush-tailed bettong Clade (in colour), and the additional *B. anhydra* specimen, and its relationship to other *Bettongia* taxa, using partial cytochrome *b* gene. The phylogeny is constructed using Bayesian methods (BEAST program – HKY + G model) and the Maximum Clade Credibility tree is provided. Small scale bar represents nucleotide substitutions per site.

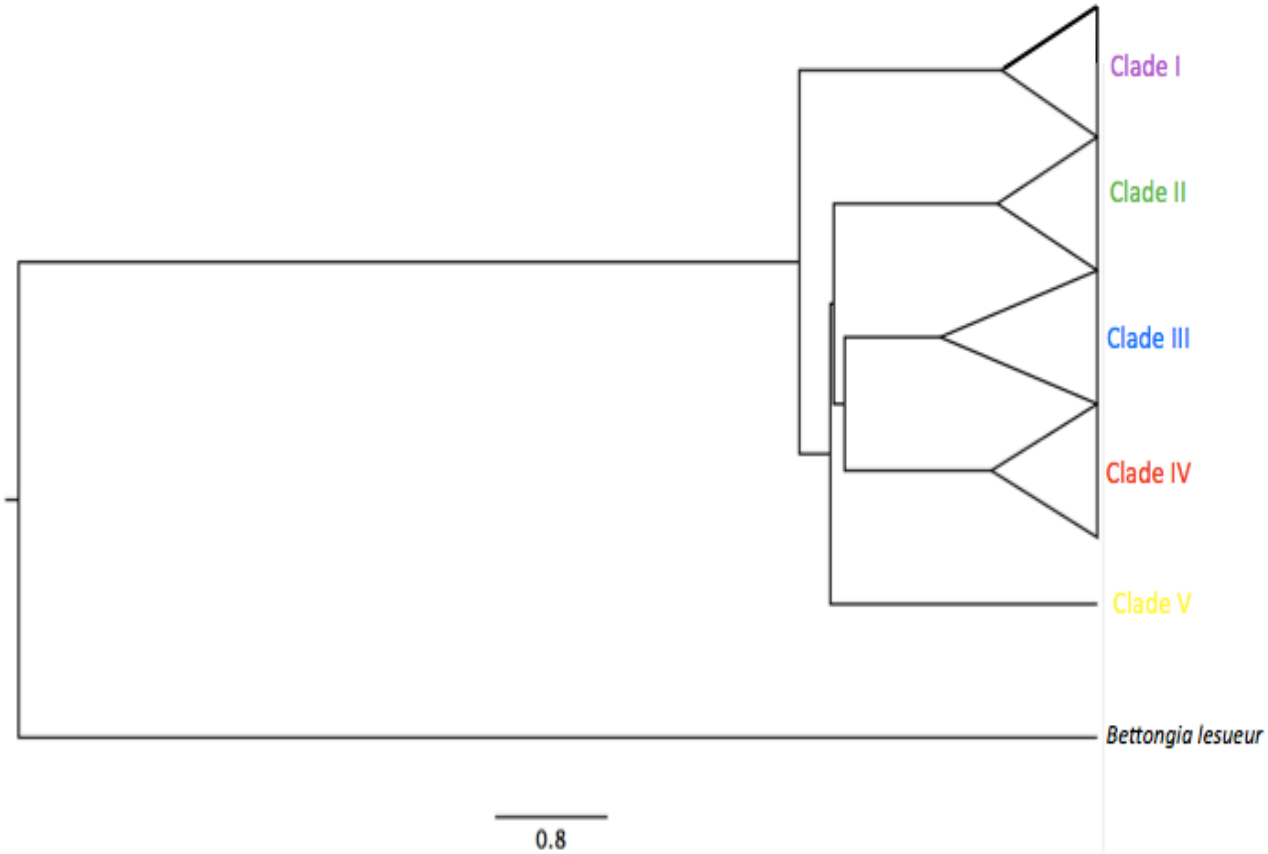


Figure SI II. Repeated phylogenetic analysis of the results depicted in Figure 3.2.2 within the main text. All sequences were used (120 sequences in total) in order to assess possible differences with the results of using only unique haplotypes (Figure 3.2.2). Tip labels have been collapsed to simplify tree. The MCC tree is provided.

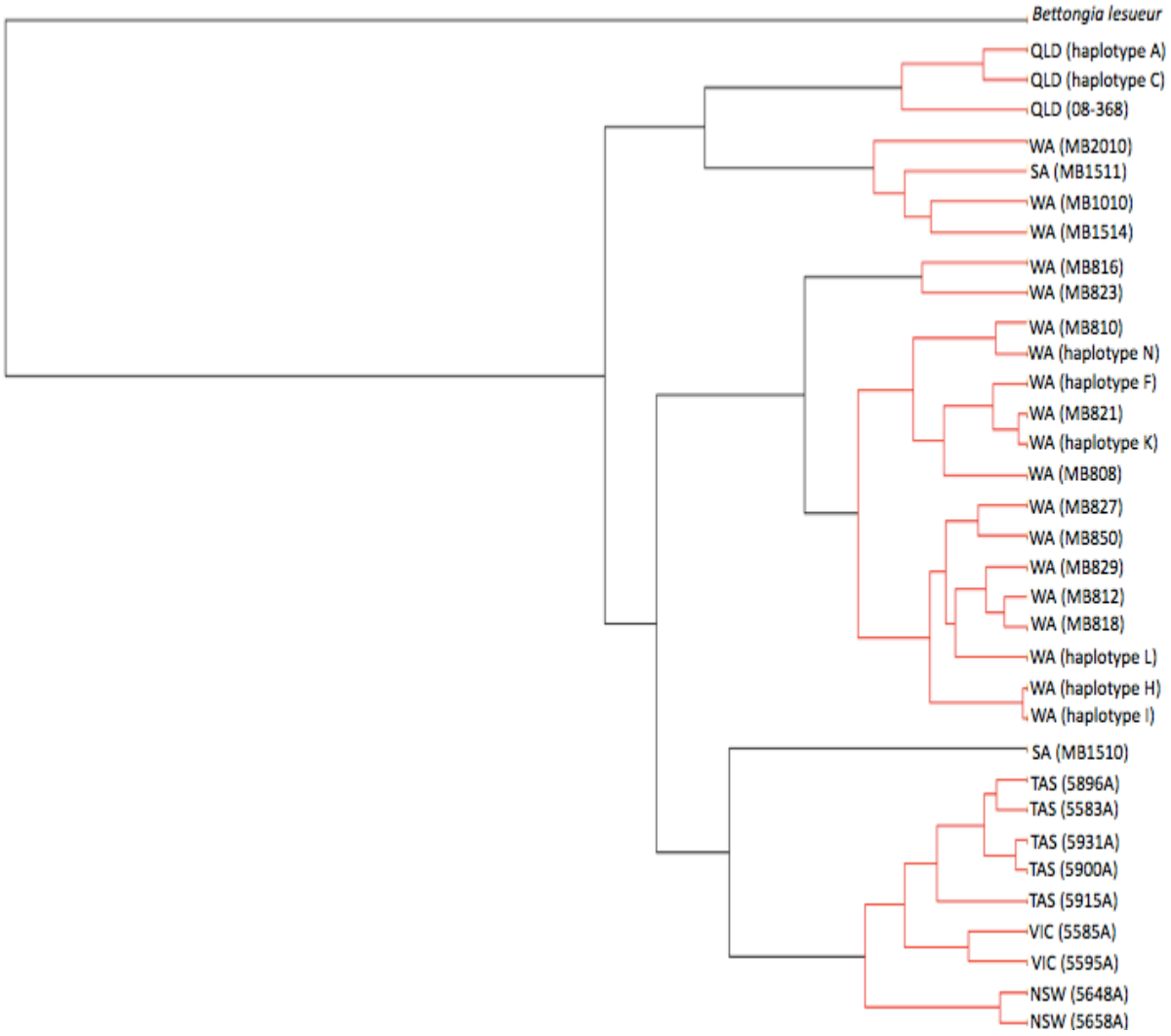


Figure SI III. Species delimitation tree constructed using the tree output from Figure 3.2.2 (in the main text) as an input file for the Generalised Mixed Yule Coalescent analysis. The analysis was conducted using a single threshold with the R package splits. The branches in red mark each inter-species tips up to most internal node.

Table SI V
Specified latitude and longitude information for all samples used in this study, which were used to carry out the phylogeography analysis.

Traits	Lat	Long
Bettongia_tropica_A	-19.256	146.818
Bettongia_tropica_C	-23.001	143.001
Bettongia_tropica_08-369	-15.466	145.283
5583A_gaimardi_TAS	-42.001	147.001
5585A_gaimardi_VIC	-37.001	144.471
5595A_gaimardi_VIC	-37.28	145.471
5648A_gaimardi_NSW	-31.001	147.001
5658A_gaimardi_NSW	-31.015	146.001
5896A_gaimardi_TAS	-41.001	146.001
5900A_gaimardi_TAS	-42.215	146.371
5915A_gaimardi_TAS	-41.987	147.462
5931A_gaimardi_TAS	-42.001	146.241
WA_MB2010	-30.767	128.971
WA_MB1010	-30.767	128.971
WA_Venus_bay	-33.216	134.666
WA_Wylie_scarp	-33.272	123.919
B_p_ogilbyi_Hapl_F	-32.331	117.875
B_p_ogilbyi_Hapl_H	-31.231	117.095
B_p_ogilbyi_Hapl_I	-30.331	116.875
B_p_ogilbyi_Hapl_K	-32.341	116.738
B_p_ogilbyi_Hapl_L	-31.762	117.285
B_p_ogilbyi_Hapl_N	-32.352	118.875
Kangaroo_Island_SA_MB1510	-35.833	137.251
MB816_WRS_WA	-32.331	117.875
MB823_Dwarladinepool_WA	-32.783	116.966
MB810_Tunnel_cave_WA	-34.042	115.008
MB827_Jerramungup_WA	-33.941	118.921
MB808_Faure_island	-25.843	113.894
MB818_WRS_WA	-32.33	117.875
MB821_Contine_WA	-32.799	117.045
MB829_Mammoth_cave_WA	-34.042	115.008
MB850_Dryandra_WA	-32.783	116.966
MB812_Tunnel_cave_WA	-34.042	115.008
Bettongia_lesueur (estimate)	-20.798	115.406

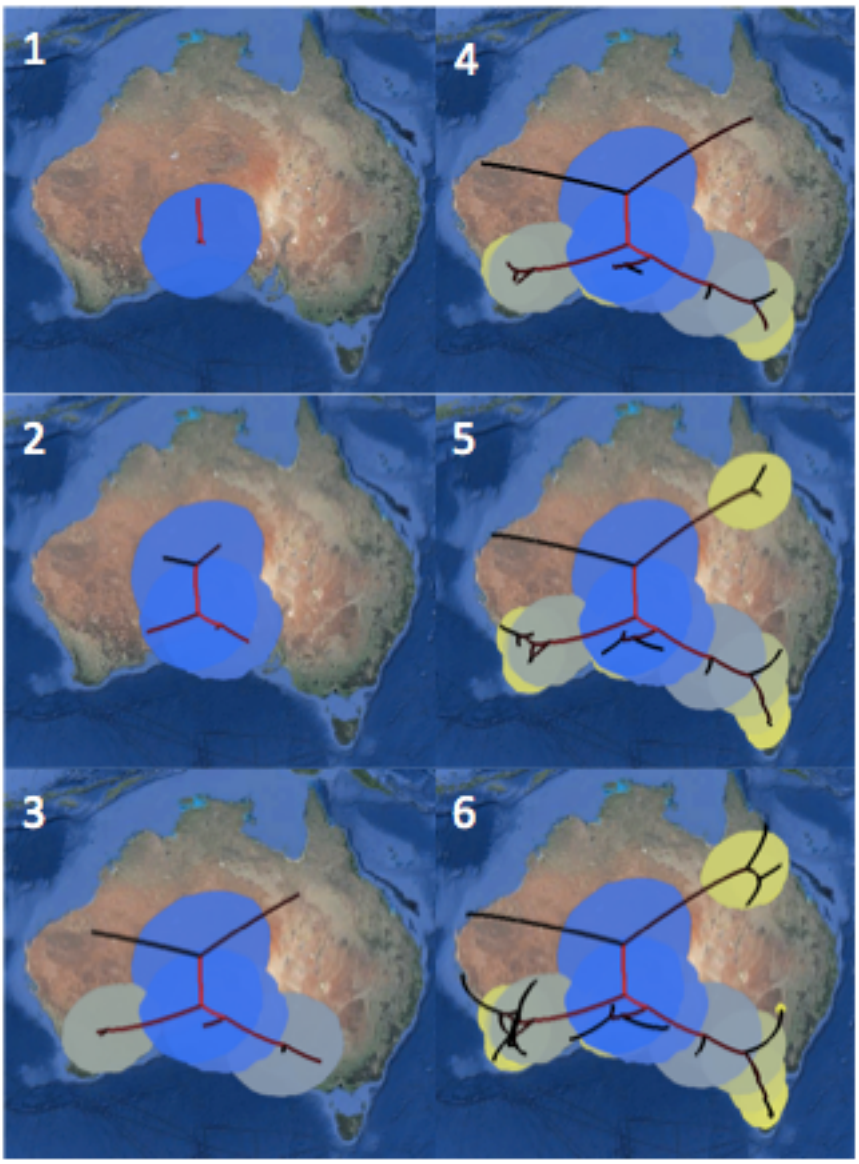


Figure SI IV. Reconstructed continuous spatial diffusion pathways of *Bettongia*. Branch colour mapping relates the node height values of the MCC tree to the colours in between the specified maximal and minimal boundaries resulting in a continuous colour gradient (red to black). Polygons colour mapping are mapped according to the relative time of the dispersal patterns (blue to yellow), the sequence of events follows numbers in order 1 through to 6. This visualisation was generated using Google Earth.

3.3 CHAPTER SUMMARY

In Chapter Three, we set out to investigate some outstanding questions that were raised as a result of two previous bettong studies – namely, the *Bettongia penicillata ogilbyi* (woylie) study of (Pacioni *et al.,* 2015 – see Appendix D) and the *Bettongia anhydra* study of (McDowell *et al.*, 2015 - see Appendix C). During collation of data for these studies it became apparent that there were inconsistencies within the taxonomic framework of bettongs. In this chapter, newly sequenced aDNA from historical samples and, modern data from all bettong taxa were collated for the first time. Our results uncovered a distinct 'new' clade that once lived in central southern Australia, challenging the concept that the arid Nullarbor zone acted as a barrier to gene flow in brush-tailed bettongs. The restriction of this study to mtDNA loci meant that information gathered is limited to the study of matrilines. A solution to this limitation is the integration of nuclear genes; preliminary data in this chapter might suggest this approach is viable on a select number of specimens. In comparison to mtDNA markers, nuclear genes offer high rates of duplication and translocation, allowing for genetic questions to be explored at a population wide level (see Chapter Six - for more discussion of this future direction).

Recent developments in extraction techniques and DNA sequencing technologies, has opened up new opportunities for the exploration of old and degraded samples in Australia. In Chapter Three, we briefly utilised some of these HTS techniques (see Chapter Five for a more in-depth study), in order to assess the practicality of attaining nuclear DNA from one of the historical samples from the Nullarbor caves location. A study integrating ancient mtDNA and nuDNA (from future studies and already published data i.e. Pacioni *et al.*, 2015) will be pivotal in providing more robust evidence towards the notion of whether the brush-tailed bettongs represent distinct *bona fide* species or if they are fragmented populations of a single biological species. This information can aid current conservation strategies and offer greater insights into translocation programs currently aimed at preserving genetic diversity loss in brush-tailed bettongs populations.

Focusing our efforts back onto a single location within Australia, the following chapter (Chapter Four), thoroughly assesses a palaeontological cave site and holistically explores the faunal and floral assemblage. This chapter is a move away from a single species, and rather explores a collection of species identified in a palaeo-ecosystem found within a cave site. The conservation rationale for this is that aDNA may assist in the rehabilitation of islands and that preserved DNA may complement the fossil record in making informed decisions regarding reintroduction.

CHAPTER FOUR

The first genetic study to explore fauna and flora species from historical bones and ancient sediments from Kelly Hill Cave, Kangaroo Island

Figure 4.1.1 Top left, the entrance to Kelly Hill Cave. Bottom left excavation of the site at Kelly Hill Cave (K1), from which all samples in this study were taken. Photo on the right, shows surface scatter found at the excavation site. (Photos taken by DH).

4.1 PREFACE

This chapter consists of a published manuscript titled ***'Thorough assessment of DNA preservation from bone and sediments excavated from a late Pleistocene—Holocene cave deposit on Kangaroo Island, South Australia'*** *[Quaternary Science Reviews 84 (2014) 56-64]. The content of section 4.2 is the same as that of the published manuscript with only minor changes to incorporate thesis formatting and referencing. The published version can be found in Appendix A.*

The exploration of islands has long provided useful information about evolutionary processes and change. Islands are also a major focus as refugia for endangered species to the ever-increasing pressures from anthropogenic and invasive species. Previously in Chapter Two, a single (rock-wallaby) species from an island was examined for identifications purposes. In this chapter an island will again provide the location, this time presenting the results of an assessment of aDNA preservation from a cave site (Figure 4.1.1) on Kangaroo Island. In this chapter a high-throughput sequencing (HTS) approach was used and is capable of producing a large amount of highly informative genetic data, than the previous Sanger sequencing techniques (used in Chapter's Two and Three). The data generated from HTS platforms, allows thorough characterisation of DNA damage, contamination and polymorphisms, a necessity for accurate ancient molecular analysis. Due to the degradable nature of the fossilised material used in this study, the applications of HTS proved to be an advantageous approach for several reasons: (1) the samples were variable in age, (some >20 kya); (2) DNA is likely to be highly degraded and; (3) samples contained a mixture of sequences from different species, and also expected to be of high risk of human contamination.

Chapter Four provides the first genetic study from Kelly Hill Cave (KHC) on Kangaroo Island (KI), South Australia and focuses on both fossil bones and sediments of various ages. Bones and sediment samples taken from all (15) layers of the site were assessed for aDNA preservation and analysed using two HTS platforms. The aim of this chapter was to: (1) assess the preservation of aDNA from late Pleistocene-Holocene fossils in KHC; (2) determine whether the morphological identities assigned to the fossil bones match the molecular identities; (3) identify those bones which are morphologically indistinct using a molecular approach to

attain a species level identification and complement previous morphological assignments; (4) assess the preservation of aDNA in sediments from all 15 layers of the site and identify any floral sequences to family and/or genus level of taxa in the stratagraphic layers; (5) and lastly provide insight on how identities of the extinct taxa recovered, may provide as useful information for conservation strategies for future research.

4.1.2 STATEMENT OF CONTRIBUTION

As lead author on this manuscript, I collected the bone samples and sediment cores during the field trip (in 2011) to Kelly Hill Cave, Kangaroo Island, with the supervision of Prof. Michael Bunce and Dr James Haile. All samples were catalogued at Flinders University and prepared for DNA work in the ancient DNA facility at Murdoch University. I carried out all ancient DNA molecular analyses, bioinformatics and phylogenetic work. Edits to the final draft, comments and feedback were contributed from all co-authors. More specifically, interpretations and discussions required for the palaeontological aspect of the paper were provided by Dr Gavin Prideaux and Dr Matthew McDowell (Flinders University, Adelaide). Prof. Michael Bunce, further provided advice and guidance throughout the production of this manuscript.

Dalal Haouchar

4.2 THOROUGH ASSESSMENT OF DNA PRESERVATION FROM FOSSIL BONE AND SEDIMENTS EXCAVATED FROM A LATE PLEISTOCENE–HOLOCENE CAVE DEPOSIT ON KANGAROO ISLAND, SOUTH AUSTRALIA.

Dalal Haouchar[1], James Haile[1], Matthew C. McDowell[2], Dáithí C. Murray[1], Nicole J. White[1], Richard JN Allcock[3], Matthew J Phillips[4], Gavin J. Prideaux[2] and Michael Bunce[1, 5].

[1]Australian Wildlife Forensic Services and Ancient DNA Laboratory, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA, 6150.

[2]School of Biological Sciences, Flinders University, Bedford Park, SA 5042, Australia.

[3]LotteryWest State Biomedical Facility: Genomics, School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, WA, 6009, Australia.

[4]School of Earth, Environment and Biological Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia.

[5]Trace and Environmental DNA laboratory, Department of Environment and Agriculture, Curtin University, Perth, WA 6845, Australia.

4.2.1 ABSTRACT

Fossils and sediments preserved in caves are an excellent source of information for investigating impacts of past environmental changes on biodiversity. Until recently, studies have relied on morphology-based palaeontological approaches, but recent advances in molecular analytical methods offer excellent potential for extracting a greater array of biological information from these sites. This study presents a thorough assessment of DNA preservation from late Pleistocene–Holocene

vertebrate fossils and sediments from Kelly Hill Cave Kangaroo Island, South Australia. Using a combination of extraction techniques and sequencing technologies, ancient DNA was characterised from over 70 bones and 20 sediment samples from 15 stratigraphic layers ranging in age from >20 ka to ~6.8 ka. A combination of primers targeting marsupial and placental mammals, reptiles and two universal plant primers were used to reveal genetic biodiversity for comparison with the mainland and with the morphological fossil record for Kelly Hill Cave. We demonstrate that Kelly Hill Cave has excellent long-term DNA preservation, back to at least 20 ka. This contrasts with the majority of Australian cave sites thus far explored for ancient DNA preservation, and highlights the great promise Kangaroo Island caves hold for yielding the hitherto-elusive DNA of extinct Australian Pleistocene species.

4.2.2 INTRODUCTION

Islands have long provided a natural laboratory for the study of evolutionary processes because evolutionary changes on them are often magnified, simplified and therefore more readily interpretable (e.g., Darwin *et al.*, 1858; MacArthur *et al.*, 1967; Losos *et al.*, 2010). The study of genetic variation on islands also has a long history (Lomolino *et al.*, 1989; Van der Geer *et al.*, 2010). However, ancient DNA (aDNA) analyses applied to stratified, dated faunal successions can add a temporal context, allowing the ebb and flow of genes, species and communities to be assessed, particularly in combination with more traditional analyses of vertebrate and plant macrofossils and pollen. A necessary prerequisite for aDNA research is adequate biomolecule preservation. Cave systems represent an ideal environment for palaeontological investigations as they often contain relatively complete and undisturbed stratigraphic deposits that harbour several environmental proxies (White, 2007; Butzer, 2008) that have been subjected to minimal temperature and humidity fluctuations; conditions that favour DNA persistence (Stone, 2000). Such caves represent archives of well-preserved Quaternary vertebrate assemblages (Prideaux *et al.*, 2007; Prideaux *et al.*, 2010), with the ability to preserve invaluable repositories of past biodiversity. All samples (bones and sediments) analysed in this study were obtained directly from Kelly Hill Cave (KHC), Kangaroo Island (KI) with the aim of conducting a thorough assessment of DNA preservation in KHC to determine

whether genetic data could enhance temporal information about faunal change on KI. Moreover, as part of this study the preservation of plant DNA obtained directly from sediments was assessed (Haile *et al.*, 2007; Willerslev *et al.*, 2007) with the aim to provide palaeovegetation data to complement fossil data.

In this study a combination of techniques such as a relatively new bulk-bone method (Murray *et al.*, 2013) and high-throughput sequencing (HTS) technology was used in order to capture aDNA from a variety of samples collected from KHC. Also we show how the addition of a palaeontological molecular perspective may complement existing morphology based studies allowing identification of osteologically absent and cryptic species, and the investigation of genetic change over time. These results are overlayed upon the palaeogeographic history of KI, which provides a model context for studying mainland–island interactions. This is pertinent to KHC, because it contains an excellent Late Quaternary vertebrate fossil assemblage (Pledge, 1979) that records the response of Australian native fauna to both the Last Glacial Maximum (LGM) and isolation of KI caused by rising sea levels at 8.9 ka (McDowell, 2013).

4.2.2.1 KANGAROO ISLAND

Kangaroo Island lies at the entrance to Gulf St Vincent in South Australia (Figure 4.2.1), and is the third-largest land-bridge island in Australia (4,405 km^2), with a length of 145 km and width of 55 km at its widest point (Abbott, 1974; Hope *et al.*, 1977; Lampert, 1981; Twidale *et al.*, 2002) (Figure 4.2.1). It is geologically continuous with the adjacent Fleurieu Peninsula, but was isolated by glacial erosion during the Late Carboniferous and Early Permian (Belperio *et al.*, 1999). Today it is separated from Yorke Peninsula by Investigator Strait, a 50-km stretch of 30–35 m deep water (Figure 4.2.1). During the late Pleistocene, sea levels were 120 m lower than at present (Chappell *et al.*, 1986; Yokoyama *et al.*, 2001) and KI was connected to the mainland. Global warming during the early Holocene caused a rapid rise in sea level, isolating KI from the mainland at 8.9 ka (Belperio *et al.*, 1999; Bradley, 1999; Cutler *et al.*, 2003). Prior to its isolation, gene flow was presumably continuous between the mainland KI which supported a species-rich fauna (Abbott, 1974; McDowell, 2013). Once isolated, the newly marooned organisms would have become more susceptible to genetic drift; island flora and fauna tend to be diverse

after initial isolation but subsequently suffer elevated selection pressures, loss of genetic diversity and elevated rates of extinction (Diamond, 1972; Foufopoulos *et al.*, 1999; Stiller *et al.*, 2010).

Despite the loss of genetic diversity caused by island living, many species that have become extinct or endangered on the mainland find refuge on islands due to relaxed competition and reduced predation pressures (Lister, 2004). KI retains the largest proportion of uncleared native vegetation of any southern Australian agricultural district. In addition KI remains free of rabbits and foxes that have had such a catastrophic impact on mainland biota (Robinson *et al.*, 1999) enhancing its conservation importance. Biodiversity management can be complimented by aDNA analyses and assessments of the fossil record, providing parameters such as population sizes, levels of gene flow and population relatedness (Ramakrishnan *et al.*, 2005; Leonard, 2008; De Bruyn *et al.*, 2011).

4.2.2.2 STUDY SITE

The KHC complex on KI (35.83° S, 137.33° E) is ideally suited to explore biomolecular preservation as it has an already well-studied and well-dated palaeontological record that spans the terminal late Pleistocene to the middle Holocene (McDowell, 2013). In addition numerous surveys of the island's modern flora and fauna have also been made (Robinson *et al.*, 1999).

KHC is the focus of an ongoing palaeontological project that investigates how climate change and isolation due to sea level rise has affected the fauna of KI (McDowell *et al.*, 2013). To date, a fauna rich in mammals, birds, reptiles, frogs and land-snails has emerged from an excavation 4 m^2 x 1.5 m deep, and includes the remains of several species not previously recorded on KI. This site was selected based on the presence of fossil bones on the cave floor surface, depth of sediment, likelihood of stratigraphic integrity and likelihood of encountering speleothems that can be U/Th dated (McDowell, 2013; McDowell *et al.*, 2013).

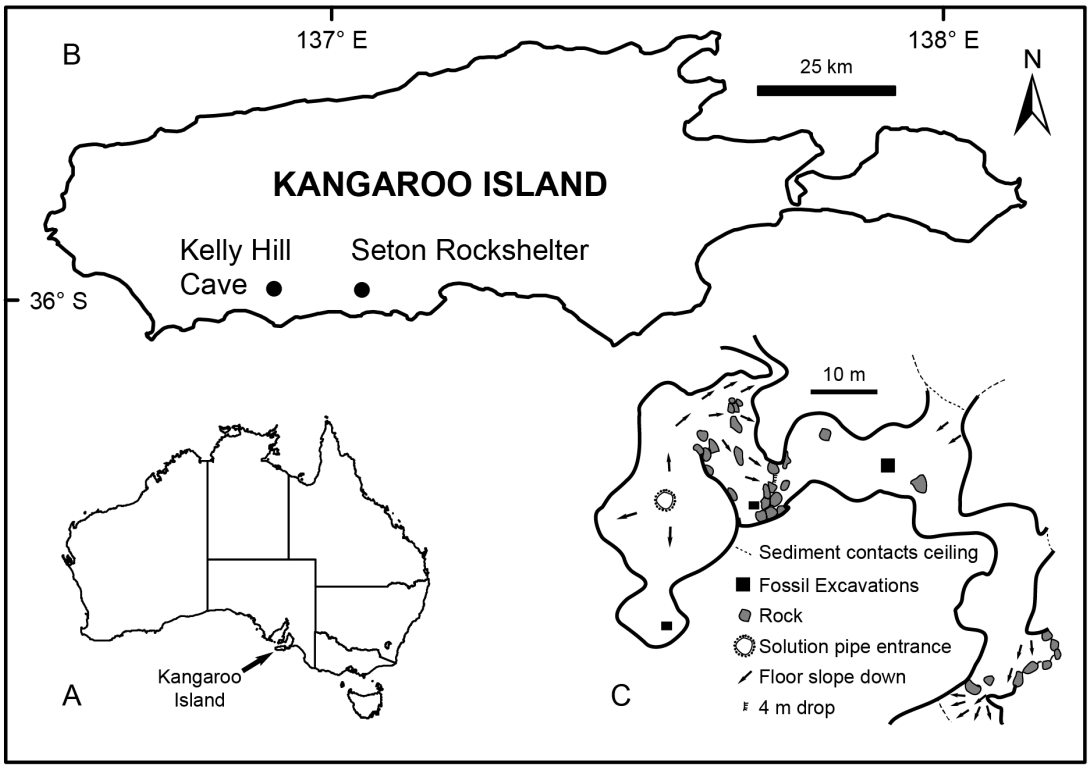


Figure 4.2.1 A. Location of Kangaroo Island relative to the Australian mainland. B. Location of Kelly Hill Cave and Seton Rockshelter, Kangaroo Island. C. Relevant map section of Kelly Hill Cave showing the location of modern solution pipe entrances, fossil excavations and the blocked palaeo-entrance through which excavated sediments and bones entered the cave (McDowell, 2013).

4.2.3 MATERIALS AND METHODS

We used HTS technologies (Roche GS-Junior and Ion Torrent, Personal Genome Platforms (PGM)) to target chloroplast (plastid) DNA (cpDNA) and mitochondrial DNA (mtDNA) from sediment samples and assessed aDNA preservation of bones collected from several sedimentary layers ranging in age from >20 to 6 ka (McDowell *et al.*, 2013). This study uses a combination of techniques to recover aDNA from multiple samples including a novel bone-grind technique, hitherto applied only to two cave sites in Western Australia (Murray *et al.*, 2013). To test the veracity or to refine morphology-based identifications, aDNA was extracted from 70 complete macropodid and reptile postcranial bones. Some specimens had been identified morphologically to species level, while others were only identified to family level (McDowell, 2013).

The layers span the terminal Late Pleistocene to the middle Holocene (>20 ka - 6 ka). Bones from layers 3 - 11 contained adequate collagen to be AMS radiocarbon dated (McDowell *et al.*, 2013). Skeletal remains from layers 12 - 15 lacked viable collagen and what little remained was too degraded to be radiocarbon dated (McDowell *et al.*, 2013). Standard aDNA protocols specifies that pre-PCR procedures are conducted in a dedicated aDNA clean room, with subsequent post-PCR work carried out in a separate laboratory in order to minimise contamination (Pääbo, 1989; Cooper *et al.*, 2000; Fulton, 2012). DNA extractions and amplifications were conducted at Murdoch University, whilst Sanger sequencing was performed at a commercial facility (South Korea) and HTS was carried out at Murdoch University (Roche, 454 GS-Junior) and the Lotterywest State Biomedical Facility Genomics Node at Royal Perth Hospital (Ion Torrent, PGM). A more detailed version of the molecular methods can be found in the Supplementary data (section 4.2.8 – SI III).

4.2.3.1 BACKGROUND TO SEDIMENTS AND SAMPLE COLLECTION

Numerous sediment core samples were collected from all exposed layers of the KHC excavation using sterile equipment and protective clothing to preserve their genetic integrity. Prior to sediment collection approximately 5 cm of surface soil was removed from the wall prior to coring to minimise possible contamination. A 50 mL falcon tube was pushed into the pit wall to core out the sediment of each distinctive layer. Each tube was sealed, labelled appropriately and stored at 4ºC to be used in subsequent aDNA work. Previous sedimentary analyses of all 15 layers has been conducted to assess grain composition, size, colour, petrography and geochemistry (McDowell *et al.*, 2013). Sub-samples from a total of 20 sediment cores were then subjected to aDNA analysis, targeting plant cpDNA and vertebrate mtDNA.

4.2.3.2 DNA EXTRACTION METHODS

All samples were extracted using methodologies designed for optimal aDNA recovery. Therefore, different extraction protocols were applied to bones and sediments and are described independently.

4.2.3.3 DNA EXTRACTION OF BONES

Fossil remains were ground to powder using a Dremel tool (part no. 114: Germany) set to high rotational speeds using DNA extraction method described by (Haouchar *et al.*, 2012). Briefly, ~0.2g of bone powder was collected in a 1.5 mL Eppendorf tube, weighed and stored at 4° C for later digestion. A number of bones were individually extracted (i.e. one bone per extraction digest), although where bones were morphologically uninformative and unidentified, a bulk-bone method was applied. This entailed grinding up to six discrete bones together and extracting DNA from the resulting powder for molecular analysis (Murray *et al.*, 2013). Silica-based DNA extractions were performed (Haouchar *et al.*, 2012) and all extracts quantitatively screened using primer sets targeting marsupials 12S and 16S rRNA markers (see Table 4.2.1 for primer combinations). Quantitative PCR assays were carried out using SYBR-green qPCR (Bunce *et al.*, 2012) as described in the Supplementary data (section 4.2.8 – SI III).

4.2.3.4 DNA EXTRACTION FROM SEDIMENTS

All sediments were processed using the Sergey Bulat extraction method optimised for small amounts of material including controls (Haile, 2012). Briefly, ~2g of sediment were processed in the Bulat buffer and purified over silica columns (see Haile *et al.,* 2012). Like the bone, DNA extractions were screened using qPCR that employed two generic primer sets for plants; *trn*Lg/h and *rbc*L. The *trn*Lg/h assay amplifies short sections of the *trn*L intron (Taberlet *et al.*, 1991; Taberlet *et al.*, 2007) and the *rbc*L primers target a coding segment of the plastid *rbc*L gene (Table 4.2.1) (Gielly *et al.*, 1994; Chiang *et al.*, 1998; Kress *et al.*, 2007). Concentrations for the DNA digest, qPCR set up and cycling conditions are further described in the Supplementary data.

Table 4.2.1 Mitochondrial 12s rRNA/16s rRNA and Cytochrome *b* primer sequences and conditions used in this study to genetically characterise marsupial/mammal fossils and plants from sediments from KHC.

Primer name	Primer sequence 5' → 3'	Annealing temperature (°C)	Amplicon size (bp)	Primer info
Primers for mammal/marsupial fossils				
12s_Macro_40F	GAYCTACACATGCAAGTTTCCGC	53	175	This study
12s_Macro_240R	CGGTGGCTGGCACGAGATTTAC			
12s_Macro_725F	GGAAAGYAATGGGCTACATTTTCTAA	60	115	This study
12s_Macro_843R	GCCTATTTCAATTAAGCTCTCTATTCT			
12s_Mars_520F	GGTCATAGCATTAACCCAAATTAACAG	55	170	This study
12s_Mars_690R	CTAATCCCAGTTTGTCTCTTAGCT			
16s_Mam1_F	CGGTTGGGGTGACCTCGGA	54	150	(Taylor,
16s_Mam1_R	GCTGTTATCCCTAGGGTAACT			1996)
Cytb_Macropod_250F	CACGCTAACGGAGCATCCATATTC	56	160	This study
Cytb_Macropod_450R	GCCGATGTAGGGGATAGCGG			
Cytb_Macropod_400F	TACCGTGAGGACAAATATCATTCTGA	56	160	This study
Cytb_Macropod_600R	GAGCCTGTTTCGTGTAGGAATAG			
Primers for plants from sediment				
*trn*Lg	GGGCAATCCTGAGCCAA	54	90	(Taberlet
*trn*Lh	TTGAGTCTCTGCACCTATC			*et al.*, 2007)
*rbc*L_F	GGCAGCATTCCGAGTAACTCCTC	53	100	(Chiang *et*
*rbc*L_*R*	CGTCCTTTGTAACGATCAAG			*al.*, 1998)
Primers for snake fossils				
Cytb_Snake_55F1	CTCCACCTGATGAAACTTCGG	54	145	This study
Cytb_Snake_220R1	ATATGGATGCGCCGATTGCG			

4.2.3.5 SEQUENCING OF DNA FROM BONES AND SEDIMENTS

All bone extracts were screened for DNA amplification. Any extracts, which successfully yielded amplicons of the target size were purified (see Supplementary data, section 4.2.8 – SI III) and sequenced. For single-source bone samples Sanger sequencing was employed (using ABI, BigDye chemistry at Macrogen), for mixed samples and sediments a next generation HTS approach was used.

DNA extracts chosen for HTS were assigned a unique six base pair (bp) DNA tag (specifically a Multiplex Identifier-tag, MID-tag) and built into primers as fusion tags. All fusion-tagged PCRs were carried out in 25 µl reactions (see SI I and II – for master mix) and imaged by qPCR (Bunce *et al.*, 2012). The general cycling conditions and amplicon purification can also be found in Supplementary data (SI

III). After pooling the amplicons, the library was then quantified with calibrated standards using qPCR (Bunce *et al.*, 2012) to determine appropriate templating ratios for HTS. Emulsion PCR and GS Junior 454 Sequencing were performed as per Roche GS Junior protocols for amplicon sequencing (http://www.454.com). Sediment DNA was prepared for GS junior 454 and Ion Torrent PGM sequencing using both *trn*Lg/h and *rbc*L primer sets (Table 4.2.1). Extracts were quantified in the same manner as the 454. All qPCRs were generated in triplicate and pooled accordingly. Amplicon pools were cleaned using Agencourt AMPure XP PCR Purification Kit (Beckman Coulter Genomics, NSW, Aus) and separated by gel electrophoresis. GS Junior 454 Sequencing and emulsion PCR were performed as per Roche GS Junior protocols. Ion Torrent emulsion PCR was performed on One Touch 2 system and sequencing was performed on an Ion Torrent (PGM).

4.2.3.6 DATA ANALYSIS AND SEQUENCE IDENTIFICATION

All sequence reads generated from the two platforms (454 GS Junior and Ion Torrent PGM) for both bone and sediment extracts were filtered in a similar manner. First, sequences were sorted into sample batches based on the unique MID-tags using the program Geneious v5.6.5 (Drummond *et al.*, 2011). Tags and primers were trimmed from the sequences allowing for no mismatch in length or base composition. All sequence results that were seemingly less than the expected amplicon size depending on the primer set used, were removed from the analysis. Once sequences were searched to contain all relevant information, i.e., they all retained the unique MID-tag, forward and reverse primer and adaptor primer, they were trimmed and searched against the NCBI GenBank nucleotide database (Benson *et al.*, 2006) using BLASTn version 2.2.23 (Altschul *et al.*, 1990) to identify reads. BLASTn datafiles were analysed in the program MEtaGenome Analyzer v4 (MEGAN) (Huson *et al.*, 2007) (see Supplementary data, SI IV for a detailed description of the analyses).

4.2.4 RESULTS AND DISCUSSION

4.2.4.1 OVERVIEW OF SEQUENCE DATA OF BONES FROM KHC

Approximately 350,000 sequence reads were generated from multiple 454 (GS Junior) and Ion Torrent (PGM) runs. DNA from a range of taxa including *Macropus*, *Onychogalea*, *Potorous*, *Bettongia*, *Dasyurus*, *Rattus* and *Notechis* was amplified

using an array of primers suitable to target the gene and organisms of interest (Table 4.2.1). The following results sections summarise the DNA analyses from KHC bones (Figure 4.2.2). Initially, a number of bones that had previously been identified morphologically at family level were screened to investigate DNA preservation at the site, and confirm the taxonomic identity. Approximately 70% of the 19 bones that were randomly selected (dispersed evenly throughout the 15 layers) from the site yielded DNA, including those from the deeper layers 14 and 15 enabling amplification of between 100 and 250bp sequences at a time. After this initial screening, another 45 bone fragments were subject to aDNA methods including an inventory of fragments from macropodid and murid limbs and vertebrae, and snake vertebrae.

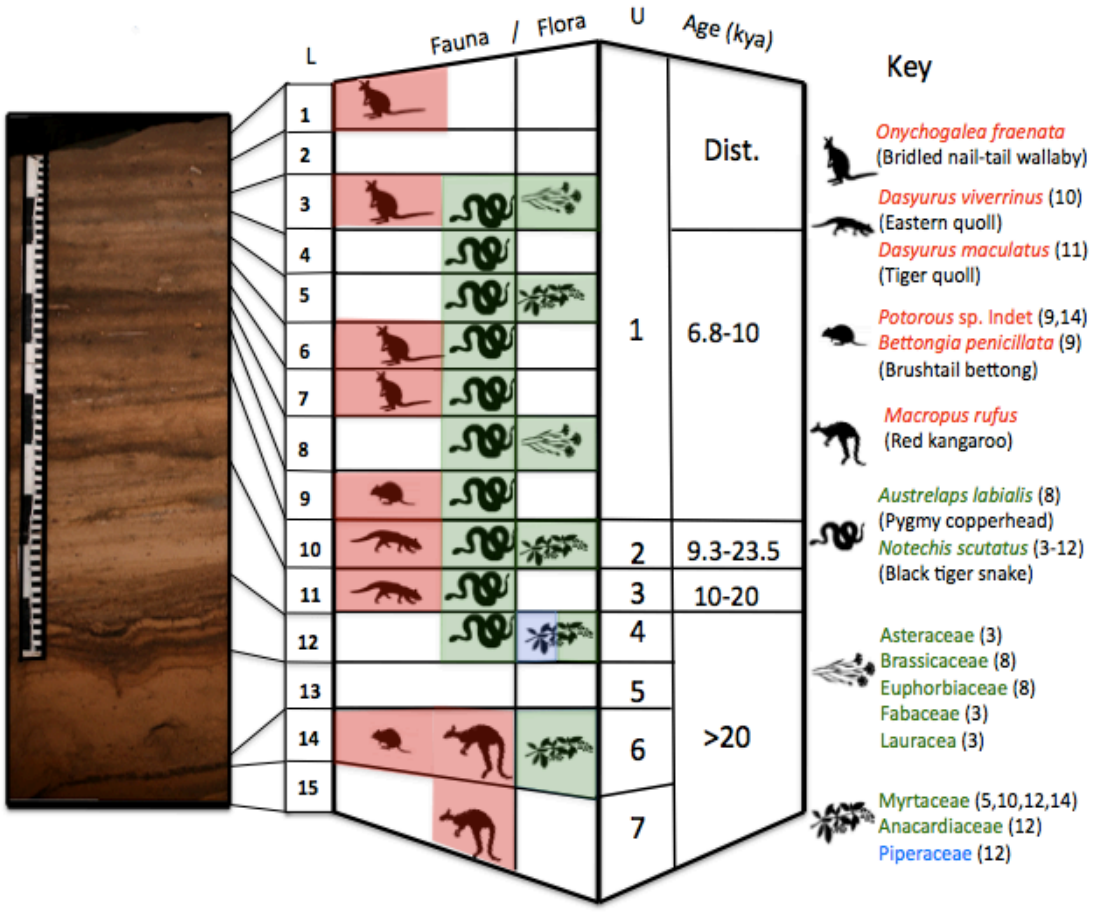


Figure 4.2.2 A snapshot of the fauna and flora recovered throughout the layers of the excavation site in KHC. Units are indicated (U: 1-7) as well as Layers (L: 1-15). Age (ka) indicates the time over which each accumulated based on both U-Th and radiocarbon ages. Dating results suggest that a depositional hiatus occurred between accumulation of Units 3 and 2 and that Unit 2 has been reworked. The first three layers represent a disturbance (Dist.) followed by twelve undisturbed layers (4-15). Scale bar is 1.0 m in length (McDowell et al., 2013). Red shading indicates fauna and flora extirpated from the island; green shading indicates species still present on KI today; blue shading shows species not previously documented on KI. Numbers in the key represent the layers each species was found, when multiple species are found in one layer.

4.2.4.2 RED KANGAROO

Two postcranial bones identified as 'Macropodinae genus and species indeterminate' (sp. indet.) from layers 14 and 15 which are >20 ka (Figure 4.2.2), yielded two sequences, one being 98% and the other 99% similar to red kangaroo (*Macropus rufus*), respectively. Morphologically diagnostic specimens of this species were not detected by McDowell (2013), but the species has previously been recorded on KI on the basis of three tooth fragments retrieved from Seton Rockshelter, nearby archaeological assemblage (Hope *et al.*, 1977; Figure 4.2.1). The presence of *M. rufus* in layers 14 and 15 likely represents a time when KI was connected to the mainland and predated the LGM. Red kangaroos are currently widely distributed through the drier regions of mainland Australia but typically occur in arid and semi-arid regions (Van Dyck *et al.*, 2008; Jackson *et al.*, 2010). This finding is significant as it demonstrates that KHC is capable of long-term DNA preservation and provides additional impetus for further excavations to be carried out at this site. *M. rufus* probably became extirpated on KI as more wooded vegetation returned to KI.

4.2.4.3 BRIDLED NAIL-TAIL WALLABY

Bridled nail-tail wallaby (*Onychogalea fraenata*) sequences were obtained from a combination of single-bone analysis and from bulk-bone sampling methods. A total of four haplotypes were observed with genetic differences distinct from the mainland individuals (71% to *O. unguifera* and 98% to *O. fraenata*). GenBank lacks reference sequences of the Crescent Nail-tail Wallaby (*O. lunata*) and although the geographic ranges of *O. fraenata* and *O. lunata* overlapped on the mainland until the late 19th century, only *O. fraenata* has been identified from KI based on morphological evidence (McDowell, 2013). *O. fraenata* is now highly endangered (Gordon *et al.*, 1980; McKnight, 2008) and persists only in managed populations (Van Dyck *et al.*, 2008; Kingsley *et al.*, 2012). *O. lunata* is extinct and *O. unguifera* (Northern Nail-tail wallaby) is widespread across northern Australia. Prior to European settlement *O. fraenata* and *O. lunata* were common throughout southern and eastern Australia (Van Dyck *et al.*, 2008). The critical status of *O. fraenata,* and its susceptibility to foxes, stock grazing and habitat destruction has resulted in a massive decline over the last 100 years. Pending further research, KI might be considered as a potential reintroduction site for *O. fraenata.*

4.2.4.4 POTOROO

DNA sequences with a potoroo affinity were obtained using both a single-bone extraction and bulk-bone methods. Overall, three haplotypes were observed which varied considerably from the reference mainland potoroo sequences acquired from GenBank. Two bones identified as Macropodinae sp. Indet. were individually extracted and analysed. A number of bones (5 in particular) identified as *Potorous platyops* from various layers collected (including surface scatter, layers 5 and 7) were subjected to the bulk bone method. An unresolved polytomy (Figure 4.2.3) is revealed from the three samples (two single bone extractions and one bulk sample) with low posterior support values for the four haplotypes observed on KI, making species identification problematic. The cluster does however show that the species belongs in *Potorous*, since the percentage similarities between the mainland and island sequences show little discrepancy. Potoroo bone 1 (Figure 4.2.3) and potoroo bone 2 share a similar match of 92.8% and 91.9% to *P. tridactylus* whilst haplotype 1 and 2 have a 98.3% and 98.5% similar match to *P. gilbertii*. As the species of *Potorous* is not clearly identified using this dataset, more bone fragments from KHC should be sequenced to identify whether the variability in this sequence cluster represents elevated genetic drift due to island isolation or potential for a new *Potorous* species, endemic to KI.

Prior to European settlement, potoroos were widely distributed across the continent. However, the combined pressure of habitat loss and introduced predators and competitors (Frankham *et al.*, 2012) resulted in dramatic range reductions for most potoroos. *P. platyops* is extinct, *P. tridactylus* is listed as vulnerable, *P. longipes* is endangered and *P. gilbertii* is critically endangered. Since the time of European settlement *P. platyops* was only recorded as living in the south-west of Western Australia and was likely already extinct on KI at that time (Robinson *et al.*, 1999). Fossil remains of *P. tridactylus* have been recovered from KHC (McDowell, 2013), but it appears to have been extirpated well before European settlement.

Another bone recovered from KHC analysed using the single bone extraction method, showed a genetic similarity of 99.3% to sequence from mainland specimens of the brush-tailed bettong or woylie (*Bettongia penicillata*) (Figure 4.2.3; *Bettongia*

bone 1). Bettongs have also suffered dramatic range reductions due to anthropogenic effects and introduced predators throughout Australia (Claridge *et al.*, 2007). The now extinct subspecies *Bettongia penicillata penicillata* was once plentiful on the South Australian mainland, whilst the burrowing bettong or boodie (*B. lesueur*) and *B. penicillata* have been identified from KI fossil bones (Hope *et al.*, 1977; Robinson *et al.*, 1999; McDowell, 2013). Both are now extirpated from the island. Attempts to reintroduce the burrowing bettong (*B. lesueur*) back onto KI have met with limited success (Short *et al.*, 1992; Robinson *et al.*, 1999).

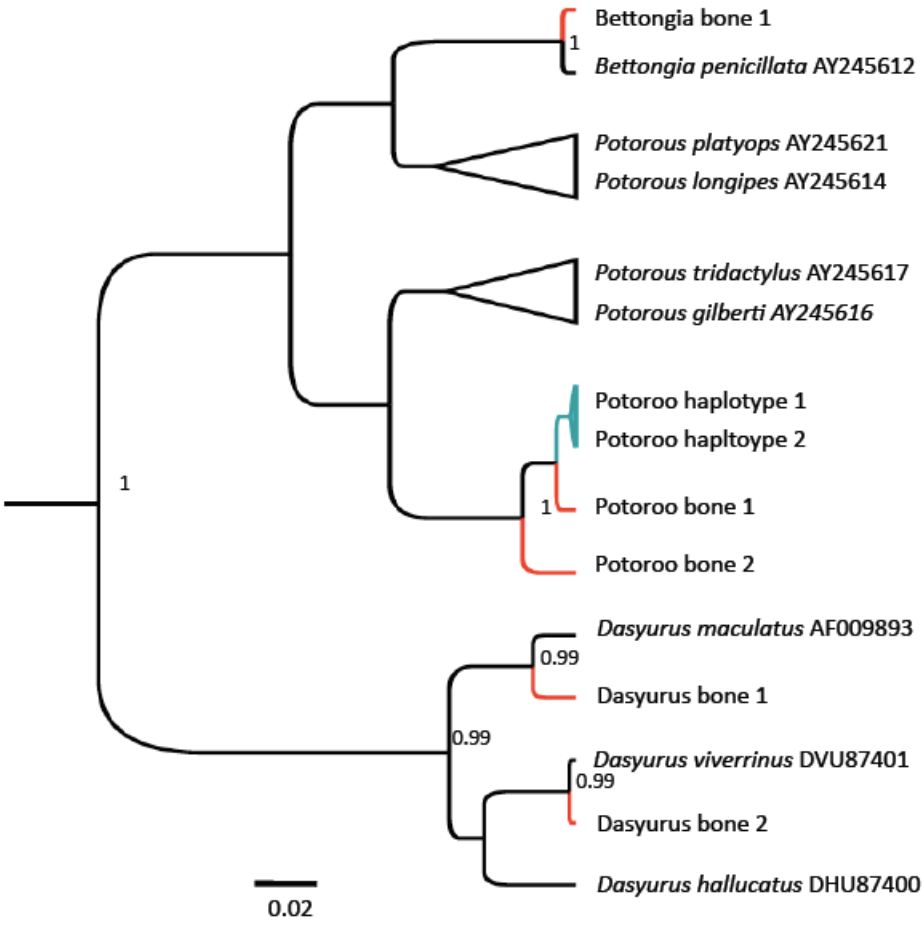


Figure 4.2.3 Bayesian 12s rRNA phylogeny from 180 bp alignment showing the closest genetic match of the fossils from KHC to the mainland reference sequences retrieved from GenBank (accession numbers shown). Sequence data illustrates the relationships between Bettongia, Potoroo and Dasyurus. Blue coloured nodes are the result from bulk bone sampling; whilst red coloured nodes are single bones, single extraction samples. The tree was built using a HKY model and Yule tree prior with invariant gamma sites and imposed with a relaxed molecular clock. Values on node show >90% posterior probabilities. The scale represents the number of nucleotide substitutions per site.

4.2.4.5 DASYURIDAE

Sequences of quoll species (*Dasyurus*) were obtained using a single-bone extraction method. The results show that one bone yielded a DNA sequence 99% similar to tiger quoll, *D. maculatus* (Figure 4.2.3; *Dasyurus* bone 1). Whilst DNA extracted from a small right femur provisionally identified by one of us (MCM) as cf. *Phascogale*, was found to be 99.5% similar genetically to mainland eastern quoll, *D. viverrinus* (Figure 4.2.3; *Dasyurus* bone 2). These findings support physical fossil evidence that both species were present on KI. *D. viverrinus* appears to have been lost from KI during the mid-Holocene (McDowell, 2013) but *D. maculatus* persisted until it was extirpated by Europeans. During the 19th century, quolls were reported as being extant on KI until *ca*. 1886 (Robinson *et al.*, 1999), although no specimens are known to have been lodged in a museum. The presence of *D. maculatus* in a 200-year-old European fur-trapper accumulation at Bales Bay (Walshe, In press) strongly points to this as the persistent species. Rehabilitation efforts and applications of genetic management to conserve quolls on the mainland (Jones *et al.*, 2003) have taken place to reintroduce certain species back to their former ranges (Firestone *et al.*, 1999; Firestone *et al.*, 2000). Direct evidence of fossils (McDowell, 2013), and now *Dasyurus* mtDNA haplotypes from KI suggests this top predator may be a candidate for reintroduction.

4.2.4.6 MURIDAE

A number of rodents have been collected from KI, for example, *Pseudomys shortridgei* (heath rat), *Rattus lutreolus* (swamp rat) and *R. fuscipes* (bush rat). In this study crania morphologically identified as *R. fuscipes,* were genetically assessed using the bulk-bone method to confirm their species identity. The results show that all three groups were 98–99% similar to *R. fuscipes,* confirming the morphological identifications. Single nucleotide polymorphisms (SNPs) occur between samples and layers and probably between the different individuals sampled. However to prove this, further analyses of all bones sampled should be reanalysed individually using a single-bone extraction method in order to compare the SNPs throughout the layers and further compare this to the genetic population on the mainland.

R. fuscipes is common on KI and the mainland as well as 13 other continental islands off the coast of South Australia (Hinten *et al.*, 2003). As a result of elevated sea levels following the LGM (Barrows *et al.*, 2002; Petherick *et al.*, 2008), genetically isolated populations developed on several of the newly formed islands (Schmitt, 1978; Seddon *et al.*, 2002; Hinten *et al.*, 2003). In distinct contrast to continental islands off Australia, a combination of mtDNA and microsatellite analyses (Hinten *et al.*, 2003) indicate that the KI population of *R. fuscipes* has a greater level of genetic diversity than the adjacent mainland.

R. fuscipes has experienced population bottlenecking on the mainland (Hinten *et al.*, 2003). Combined pressures imposed by feral cats, foxes and habitat fragmentation may be further contributing to these dwindling populations (Hinten *et al.*, 2003). Given that KI has remained free from foxes, it serves as an ideal place in which to preserve the already diverse and self-maintaining *R. fuscipes* population and to host insurance populations.

4.2.4.7 ELAPIDAE

The pygmy copperhead (*Austrelaps labialis*) and eastern black tiger snake (*Notechis ater*, although now believed to be synonymised with *Notechis scutatus*) were common on KI (Robinson *et al.*, 1999; Houston *et al.*, 2002). Individuals of these species have 200–400 vertebrae that vary morphologically along the vertebral column making them difficult to identify to species when found as fossils. However, because each snake has so many vertebrae they are common in Australian palaeontological excavations. Elapid vertebrae consist largely of thick cortical bone and typically preserve high-quality DNA that can be specifically identified using molecular methods (Polly *et al.*, 2004).

The bulk-bone method (Murray *et al.*, 2013) allowed rapid initial identification of the snake species present, and then single bone extractions verified the layers in which the species occurred (see Figure 4.2.2). Elapid DNA were amplified from strata that have been radiometrically dated to >20 ka (layer 11: McDowell *et al.*, 2013) and yielded 100% and 99.5% similar matches to the *Notechis ater* and *Austrelaps labialis* respectively (the only two species recognised from excavated crania). Little genetic difference was observed within species over time, nor when compared to the modern

mainland reference sequences. The pygmy copperhead, which is widespread and common on KI, has a depauperate mainland range and is found primarily restricted to the southern Mount Lofty ranges and Fleurieu Peninsula. Habitat destruction is probably the main force driving the decline of this species (Robinson *et al.*, 1999; Houston *et al.*, 2002).

4.2.4.8 OVERVIEW OF PLANT DATA FROM SEDIMENTS

A total of eight plant families were detected (Table 4.2.2) from an initial assessment of ancient DNA preservation in sediment (*sed*aDNA). DNA was screened for 12 sediment sub-samples ranging from >20 ka to approximately 9 ka. Sediment DNA was also tested for mtDNA although no animal DNA was detected. A total of six sediment samples (Figure 4.2.2 and Table 4.2.2) successfully yielded DNA using a two-locus approach targeting the chloroplast *trn*L and *rbc*L genes resulting in approximately 100,000 sequences from several HTS runs. Sequences were identified using BLAST (Altschul *et al.*, 1990), and results interpreted with MEGAN (Huson *et al.*, 2007) (Table 4.2.2). As with previous *sed*aDNA, plant literature using these chloroplast loci (Parducci *et al.*, 2005; Jorgensen *et al.*, 2011) the taxonomic resolution is reliant upon comparative database coverage, and within certain families is constrained by the degree of interspecific variation. Nevertheless, the results provide some interesting insights. There are nearly 900 native plant species recorded on KI which currently include 40 endemic species (Holiday *et al.*, 2003). Families which are well represented on KI are those characteristic of the higher rainfall areas of South Australia in general such as Myrtaceae (*e.g.*, eucalyptus), Cyperaceae (*e.g.,* sedges), Liliaceae (*e.g.,* lilies), Ericaceae (formerly – Epacridaceae *e.g.,* heath shrubs and herbs) and Apiaceae (formerly – Umbelliferae *e.g.,* heath shrubs).

The majority of species present in the layers were found on KI and/or the adjacent mainland. However, some of the deeper layers yielded taxa not found on KI or mainland South Australia. For example layer 12 (dated at >20 ka) yielded *Piper* sp. indet. (Piperaceae – pepper family) (Table 4.2.2) from *sed*aDNA, but has never been recorded on the island before. Closely related species appear to be common in the eastern states of Australia, particularly all along the coast of Queensland and New South Wales (Atlas of living Australia; http://www.ala.org.au/). Species of the Myrtaceae family, which are considerably common throughout KI, was found in four

sedimentary layers, and included a number of taxa that could be identified to genus (Table 4.2.2). *Eucalyptus sp., Leptospermum sp.* and *Syzygium sp.*, were present in the deeper excavation layers and with the exception of *Syzygium sp.,* which remains common on the mainland, all can be found on KI today.

Table 4.2.2 Summary of plants achieved using High-Throughput DNA sequencing and number of sequences from sediment cores taken from six layers in KHC, KI. Eight families were detected using a two-locus approach targeting the chloroplast *trn*L and *rbc*L genes, resulting in over 100,000 sequences. Families indicated in bold and genus is italicized. Approximate layer ages are indicated in Figure 4.2.2. Plant locations were sourced from FloraBase [http://florabase.dec.wa.gov.au/], eFloraSA [http://flora.sa.gov.au] and Atlas of living Australia [http://www.ala.org.au/].

Plant taxa identification	**Sedimentary layer and sampling depth from surface (KHC, KI)**					
	Layer 3 (-10cm):	Layer 5 (-18cm):	Layer 8 (-59cm):	Layer 10 (-80cm):	Layer 12 (-98cm)	Layer 14 (-130cm)
Anacardiaceae					n=100	
Anacardium sp.					n=36	
Asteraceae	n=200[#§]					
Brassicaceae			n=11400[#§]			
Brassica sp.			n=10[§]			
Euphorbiaceae			n=260[#§]			
Hevea sp.			n=160			
Fabaceae	n=4000[#§]					
Lauraceae	n=350 [#§]					
Cinnamomum sp.	n=7[§]					
Myrtaceae		n=26500[#§]		n=10500[#§]	n=20700 [#§]	n=14200[#§]
Myrcia sp.		n=7			n=24	n=5
Eucalyptus sp				n=13[#§]	n=22[#§]	n=6[#§]
Leptospremum sp.		n=765[#§]		n=250[#§]	n=875[#§]	n=375[#§]
Syzygium sp.						n=8[§]
Piperaceae					n=1000	
Piper sp.					n=263	

Key: # - Presence documented on Kangaroo Island; § - presence documented on South Australia mainland.
n = number of sequences in each family/genera detected using HTS.

The study of past plant species distribution (palynology), which principally relies on pollen, has a long and venerable history. However, until the discovery of aDNA persistence in sediments (Willerslev *et al.*, 2003), the absence of macro-fossil remains (seeds, buds or vegetative tissue) and micro-fossils (pollen) was a serious limitation (Jorgensen *et al.*, 2011). Ancient sediments, in particular permafrost deposits, have proven to be an excellent archive for the long-term preservation of

environmental ancient DNA (*sed*aDNA) (Willerslev *et al.*, 2003), making them useful for palaeo-reconstructions. However, *sed*aDNA preservation in less favourable climates like that of Australia can be somewhat challenging and so far have only been reproducible in a minority of plant aDNA studies. To date no record has been made on the preservation of plant macro- and micro-remains in the sediments of KHC, therefore any information that can be gathered of past floral assemblages will provide a significant contribution to our understanding of paleovegetation and climate changes.

4.2.5 CONCLUSION

This is the first aDNA assessment of vertebrate fossils and plant DNA from KI in South Australia. Outcomes presented here add significant value to the late Pleistocene–mid-Holocene palaeontological record of KI. We demonstrate the utility of aDNA techniques when applied to KI cave deposits and build upon results of morphological studies by confirming existing identifications and revealing additional species not preserved as diagnostic fossil specimens. The excellent preservation of both animal and plant DNA extracted from KHC bones and sediments are fundamental in providing such valuable information about the past biodiversity of KI. This study provides a detailed molecular record of animal and plant species that once lived in the KHC region, including some species that have been extirpated from KI. We reveal an array of taxa ranging from marsupials and reptiles to shrubs and trees, some of which were deposited >20 ka. Our results also highlight the potential of the ongoing excavation that will delve even deeper (> 15 layers) into the past and may provide insights into the extinction of megafaunal species previously recorded on KI (Hope *et al.*, 1977) and at the KHC site (McDowell, 2013).

4.2.6 ACKNOWLEDGEMENTS

Thanks to Grant Gully for modifications of Figure 4.2.1. This study was funded by ARC DP120104435 to GJP and MB. MB is funded by an Australian Research Council Future Fellowship (FT0991741).

4.2.7 REFERENCES

Abbott, I. (1974). The avifauna of Kangaroo Island and causes of its impoverishment. *Emu* **74**, 124-134.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman (1990). Basic local alingment search tool. *Journal of Molecular Biology* **215**, 403-410.

Barrows, T. T., J. O. Stone, L. K. Fifiled & R. G. Cresswell (2002). The timing of the last glacial maximum in Australia. *Quaternary Science Reviews* **21**, 159-173.

Belperio, A. & R. Flint (1999). Kangeroo Island biological survey: Geomorphology and geology. *A Biological Survey of Kangaroo Island, South Australia 1989 & 1990*. (ed. by A. C. In Robinson, and Armstrong, D. M.). Department of Environment, Heritage, and Aboriginal Affairs, South Australia, p. 19-31.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell & D. L. Wheeler (2006). GenBank. *Nucleic Acids Research* **34,** 16-20.

Bradley, R. S. (1999). Paleoclimatology. Reconstructing Climates of the Quaternary, 2nd ed. International Geophysics Series. San Diego, London, Boston, New York, Sydney, Tokyo, Toronto: Academic Press. *Geological Magazine* **138**, 501-502.

Bunce, M., C. L. Oskam & M. E. Allentoft (2012). Quantitative real-time PCR in aDNA research. *Ancient DNA*. (ed. by B. Shapiro and M. Hofreiter), pp. 121-132. Springer.

Butzer, K. W. (2008). Challenges for a cross-disciplinary geoarchaeology: The intersection between environmental history and geomorphology. *Geomorphology* **101**, 402-411.

Chappell, J. & N. Shackleton (1986). Oxygen isotopes and sea level. *Nature* **324**, 137-140.

Chiang, T. Y., B. A. Schaal & C. I. Peng (1998). Universal primers for amplification and sequencing a noncoding spacer between the atpB and rbcL genes of chloroplast DNA. *Botanical Bulletin of Academia Sinica* **39**, 245 -250.

Claridge, A., J. Seebeck & R. Rose (2007). *Potoroos and the musky rat-kangaroo,* Collingwood, Vic, Australia, CSIRO Publishing.

Cooper, A. & H. N. Poinar (2000). Ancient DNA: do it right or not at all. *Science* **289**, 1139.

Cutler, K., R. Edwards, F. Taylor, H. Cheng, J. Adkins, C. Gallup, P. Cutler, G. Burr & A. Bloom (2003). Rapid sea-level fall and deep-ocean temperature change since the last interglacial period. *Earth and Planetary Science Letters* **206**, 253-271.

Darwin, C. R. & A. Wallace, R (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Proceedings of the Linnean Society of London* **3**, 46-50.

De Bruyn, M., A. R. Hoelzel, G. R. Carvalho & M. Hofreiter (2011). Faunal histories from Holocene ancient DNA. *Trends in Ecology & Evolution* **26**, 405-413.

Diamond, J. M. (1972). Biogeographic kinetics: estimation of relaxation times for avifaunas of southwest Pacific islands. *Proceedings of the National Academy of Sciences* **69**, 3199-3203.

Drummond, A. J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer & A. Wilson (2011). Geneious v 6.0.1. Available from http://www.geneious.com/.

Firestone, K. B., M. S. Elphinstone, W. B. Sherwin & B. A. Houlden (1999). Phylogeographical population structure of tiger quolls Dasyurus maculatus (Dasyuridae: Marsupialia), an endangered carnivorous marsupial. *Molecular Ecology* **8**, 1613-1625.

Firestone, K. B., B. A. Houlden, W. B. Sherwin & E. Geffen (2000). Variability and differentiation of microsatellites in the genus Dasyurus and conservation implications for the large Australian carnivorous marsupials. *Conservation Genetics* **1**, 115-133.

Foufopoulos, J. & A. R. Ives (1999). Reptile extinctions on land-bridge islands: life-history attributes and vulnerability to extinction. *American Naturalist* **153**, 1-25.

Frankham, G. J., K. A. Handasyde & M. D. B. Eldridge (2012). Novel insights into the phylogentic relationships of the endangered marsupial genus *Potorous* *Molecular Phylogenetics and Evolution* **64**, 592-602. http://www.journals.elsevier.com/molecular-phylogenetics-and-evolution/.

Fulton, T. L. (2012). Setting up an Ancient DNA laboratory. *Ancient DNA: Methods and Protocols*. (ed. by B. Shapiro and M. Hofreiter), pp. 1 - 11. Springer Science.

Gielly, L. & P. Taberlet (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus rbcL sequences. *Molecular Biology and Evolution* **11**, 769-777.

Gordon, G. & B. C. Lawrie (1980). The rediscovery of the bridled nailt-tail wallaby, onychogalea fraenata (Gould) in Queensland. *Australian Wildlife Research* **7**, 339-345.

Haile, J. (2012). Ancient DNA Extraction from Soils and Sediments. *Ancient DNA*. (ed. by B. Shapiro and M. Hofreiter), pp. 57-63. Humana Press.

Haile, J., R. Holdaway, K. Oliver, M. Bunce, M. T. P. Gilbert, R. Nielsen, K. Munch, S. Y. W. Ho, B. Shapiro & E. Willerslev (2007). Ancient DNA chronology

within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor? *Molecular Biology and Evolution* **24**, 982-989.

Haouchar, D., J. Haile, P. B. S. Spencer & M. Bunce (2012). The identity of the Depuch Island rock-wallaby revealed through ancient DNA. *Australian Mammalogy* **35**, 101-106.

Hinten, G., F. Harriss, M. Rossetto & P. Braverstock (2003). Genetic variation and island biogreography: microsatellite and mitochondrial DNA variation in island populations of the Australian bush rat, Rattus fuscipes greyii. *Conservation Genetics* **4**, 759-778.

Holiday, I., B. Overton & D. Overton (2003). *Kangaroo Island's native plants,* Adelaide, South Australia, Finsbury Printing.

Hope, J., R. Lampert, E. Edmondson, M. Smith & G. Van Tets (1977). Late Pleistocene faunal remains from Seton rock shelter, Kangaroo Island, South Australia. *Journal of Biogeography* **4**, 363-385.

Houston, T. F. & M. J. Tyler (2002). Reptiles and Amphibians *Natural Histroy of Kangaroo Island.* (ed. by M. Davies, C. Twidale and M. J. Tyler). Royal Society of South Australia, Adelaide.

Hunter, A. A., A. B. Macgregor, T. O. Szabo, C. A. Wellington & M. I. Bellgard (2012). Yabi: An online research environment for grid, high performance and cloud computing. *Source Code for Biology and Medicine* **7**, e1.

Huson, D. H., A. Auch, J. Qi & S. C. Schuster (2007). MEGAN analysis of metagenome data. *Genome Research* **17**, 377-386.

Jackson, S. & K. Vernes (2010). *Kangaroo: Portrait of an extraordinary marsupial,* NSW, Allen and Unwin.

Jones, M. E., M. Oakwood, C. A. Belcher, K. Morris, A. J. Murray, P. Woolley, K. B. Firestone, B. Johnson & S. Burnett (2003). Carnivore concerns: problems, issues and solutions for conserving Australasia's marsupial carnivores. *Predators with pouches: the biology of carnivorous marsupials*), pp. 422-434. Csiro Publishing.

Jorgensen, T., J. Haile, P. Moller, A. Andreev, S. Boessenkool, M. Rasmussen, F. Kienast, E. Coissac, P. Taberlet & C. Brochmann (2011). A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology* **21**, 1989-2003.

Kingsley, L., A. Goldizen & D. O. Fisher (2012). Establishment of an Endangered species on a private nature refuge: what can we learn from reintroductions of the bridled nailtail wallaby *Onychogalea fraenata*? *Oryx* **46**, 240-248.

Kress, W. J. & D. L. Erickson (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One* **2**, e508.

Kress, W. J. & D. L. Erickson (2007). A two-locus global DNA barcode for land plants: the coding rbcl gene complements the non-coding trnH-pspA spacer region. *PLoS One* **2**, 508-515.

Lampert, R. J. (1981). The great Kartan mystery. Terra Australis 5). Australian National University, Canberra.

Leonard, J. A. (2008). Ancient DNA applications for wildlife conservation. *Molecular Ecology* **17**, 4186-4196.

Lister, A. M. (2004). The impact of Quaternary Ice Ages on mammalian evolution. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 221-241.

Lomolino, M. V., J. H. Brown & R. Davis (1989). Island biogeograhy of montane forest mammals in the American Southwest. *Ecology* **70**, 180-194.

Losos, J. B. & R. E. Ricklefs (2010). *The Theory of Island Biogeography revisted*, Princeton University Press, Princeton, NJ.

MacArthur, R. H. & E. O. Wilson (1967). *The Theory of Island Biogeography*, Princeton University Press, Princeton, NJ.

McDowell, M. C. (2013). Late Quaternary fuanal responses to environmental change and isolation on a large Australian land-bridge island. *School of Biological Sciences*). Flinders University, South Australia.

McDowell, M. C., E. A. Bestland, F. Bertuch, L. K. Ayliffe, J. C. Hellstrom, G. E. Jacobsen & G. J. Prideaux (2013). Chronology, stratigraphy and palaeoenvironmental interpretation of a Late Pleistocene to mid-Holocene cave accumulation on Kangaroo Island, South Australia. *Boreas* **42**, 974-994.

McKnight, M. (2008). *Onychogalea fraenata*. *In IUCN 2013. IUCN Red List of Threatened Species. Version 2013. 1*.

Murray, D. C., J. Haile, J. Dortch, N. White, D. Haouchar, M. I. Bellgard, R. Allcock, J, G. J. Prideaux & M. Bunce (2013). Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports.* **3**, 3371.

Pääbo, S. (1989). Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 1939-1943.

Parducci, L., Y. Suyama, M. Lascoux & K. D. Bennett (2005). Ancient DNA from pollen: a genetic record of plant population history. *Molecular Ecology* **14**, 2873-2882.

Petherick, L., H. McGowan & P. Moss (2008). Climate variability during the Last Glacial Maximum in eastern Australia: evidence of two stadials? *Journal of Quaternary Science* **23**, 787-802.

Pledge, N. S. (1979). Fossil vertebrates. *Natural history of kangaroo island* (ed. by M. J. Tyler, C. Twidale and J. K. Ling). Royal Society of Southern Australia, Adelaide.

Polly, P. D. & J. J. Head (2004). Maximum-likelihood identification of fossils: taxonomic identification of Quaternary marmots (Rodentia, Mammalia) and identification of vertebral position in the pipesnake Cylindrophis (Serpentes, Reptilia). *Morphometrics: applications in biology and paleontology*), p. 197. Springer.

Prideaux, G. J., G. A. Gully, A. M. C. Couzens, L. K. Ayliffe, N. R. Jankowski, Z. Jacobs, R. G. Roberts, J. C. Hellstrom, M. K. Gagan & L. M. Hatcher (2010). Timing and dynamics of Late Pleistocene mammal extinctions in southwestern Australia. *Proceedings of the National Academy of Sciences* **107**, 22157-22162.

Prideaux, G. J., J. A. Long, L. K. Ayliffe, J. C. Hellstrom, B. Pillans, W. E. Boles, M. N. Hutchinson, R. G. Roberts, M. L. Cupper & L. J. Arnold (2007). An arid-adapted middle Pleistocene vertebrate fauna from south-central Australia. *Nature* **445**, 422-425.

Ramakrishnan, U., E. A. Hadly & J. L. Mountain (2005). Detecting past population bottlenecks using temporal genetic data. *Molecular Ecology* **14**, 2915-2922.

Robinson, A. C. & D. M. Armstrong (1999). *A Biological Survey of Kangaroo Island South Australia in November 1989 and 1990,* Adelaide Department for Environment, Heritage and Aboriginal Affairs.

Schmitt, L. H. (1978). Genetic variation in isolated populations of the Australian bush-rat, Rattus fuscipes. *Evolution* **38**, 1-14.

Seddon, J. & P. Baverstock (2002). Variation on islands: major histocompatibility complex (Mhc) polymorphism in populations of the Australian bush rat. *Molecular Ecology* **8**, 2071-2079.

Short, J., S. D. Bradshaw, J. Giles, R. I. T. Prince & G. R. Wilson (1992). Reintroduction of macropods (Marsupialia: Macropodoidea) in Australia - a review. *Biological Conservation* **62**, 189-204.

Stiller, M., G. Baryshnikov, H. Bocherens, A. G. d'Anglade, B. Hilpert, S. C. Munzel, R. Pinhasi, G. Rabeder, W. Rosendahl & E. Trinkaus (2010). Withering away - 25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution* **27**, 975-978.

Stone, A. C. (2000). Ancient DNA from skeletal remains. *Biological anthropology of the human skeleton*. (ed. by M. A. Katzenberg and S. R. Saunders), pp. 351–371. New York Wiley-Liss.

Taberlet, P., E. Coissac, F. Pompanon, L. Gielly, C. Miquel, A. Valentini, T. Vermat, G. Corthier, C. Brochmann & E. Willerslev (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* **35**, 14-21.

Taberlet, P., L. Gielly, G. Pautou & J. Bouvet (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution* **13**, 283-285.

Twidale, C. & J. A. Bourne (2002). *The land surface. In Tyler, M. J., Twidale, C. R., Ling, J. K. (eds): Natural history of kangaroo island*, Royal Society of South Australia, Adelaide.

Van der Geer, A., G. Lyras, J. De Vos & M. Dermitzakis (2010). *Evolution of Island mammals: adaptation and extinction of placental mammals on islands,* London, John Wiley & Sons.

Van Dyck, S. & R. Strahan (2008). *The mammals of Australia*, Sydney, Reed New Holland.

Walshe, K. (In press). Archaeological evidence for a sealer's and wallaby hunter's skinning site on Kangaroo Island, South Australia. *Journal of Island and Coastal Archaeology*.

White, W. B. (2007). Cave sediments and paleoclimate. *Journal of Cave and Karst Studies* **69**, 76-93.

Willerslev, E., E. Cappellini, W. Boomsma, R. Nielsen, M. B. Hebsgaard, T. B. Brand, M. Hofreiter, M. Bunce, H. N. Poinar, D. Dahl-Jensen, S. Johnsen, J. P. Steffensen, O. Bennike, J.-L. Schwenninger, R. Nathan, S. Armitage, C.-J. de Hoog, V. Alfimov, M. Christl, J. Beer, R. Muscheler, J. Barker, M. Sharp, K. E. H. Penkman, J. Haile, P. Taberlet, M. T. P. Gilbert, A. Casoli, E. Campani & M. J. Collins (2007). Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. *Science* **317**, 111-114.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky & A. Cooper (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* **300**, 791-795.

Yokoyama, Y., P. De Deckker, K. Lambeck, P. Johnston & L. Fifield (2001). Sea-level at the Last Glacial Maximum: evidence from northwestern Australia to constrain ice volumes for oxygen isotope stage 2. *Palaeogeography, Palaeoclimatology, Palaeoecology* **165**, 281-297.

4.2.8 SUPPLEMENTARY INFORMATION

SI I *DNA extraction of bone*

Fossil remains were ground to powder using a Dremel tool (part no. 114: Germany) set to high rotational speeds using DNA extraction method described by (Haouchar *et al.*, 2012). Approximately 0.2g of bone powder was collected in a 1.5 mL Eppendorf tube, weighed and stored (4° C) for later digestion. A number of bones were individually extracted (i.e. one bone per extraction digest), although where bones were morphologically uninformative and unidentified, a bulk-bone method was applied. This entailed grinding up to six discrete bones together and extracting DNA from the resulting powder for molecular analysis (Murray *et al.*, 2013).

Silica-based DNA extractions were performed following (Haouchar *et al.*, 2012) all extracts were quantitatively screened using primer sets targeting marsupials 12S and 16S rRNA markers (see Table 4.2.1 for primer combinations). Quantitative PCR assays were carried out using Sybr-green qPCR (Bunce *et al.*, 2012) and made up to a total volume of 25 µl, containing 12.5 µl ABI Power SYBR master mix (Applied Biosystems), 0.4 µM of forward and reverse primer, 8.5 µM H_2O and 2 µl DNA extract. Reaction conditions for the specific mammal primer sets were as follows; initial heat denaturation at 95°C for 5mins, followed by 50 cycles of 95°C for 30sec; with a variable annealing temperature depending on primer set used (Table 4.2.1) for 30sec; 72°C for 45sec followed by a 1°C melt curve and a final extension at 72°C for 10mins.

SI II *DNA extraction from sediments*

All sediments were processed using the Sergey Bulat extraction method optimised for small amounts of material including controls (Haile, 2012). Approximately 2g of sediment was processed in the Bulat buffer and purified over silica columns (see Haile, 2012). Like the bone, sediment DNA extractions were screened using qPCR that employed two generic primer sets for plants; *trn*Lg/h and *rbc*L. The *trn*Lg/h assay amplifies short sections of the *trn*L intron (Taberlet *et al.*, 1991; Taberlet *et al.*,

2007) and the *rbc*L primers targeting a coding segment of the plastid *rbc*L gene (Table 4.2.1) (Gielly *et al.*, 1994; Chiang *et al.*, 1998; Kress *et al.*, 2007).

Concentrations for the sediment digest using the bulat buffer were as follows; 0.02g/mL Sarcosyl, 50mM Tris-HCL (pH 8.0), 20mM NaCl, 50mM DTT, 0.8g/mL Proteinase K. DNA was eluted in 100 µL of 10 mM Tris pH 8.0 (Sigma, Mo, USA) and screened using qPCR at neat, 1/10 and 1/50 dilutions. Each qPCR reaction was made up to a total volume of 25µL, containing 12.5 µl ABI Power SYBR master mix (Applied Biosystems), 0.4 µM of forward and reverse primer, 8.5 µM H_2O and 2 µl sediment aDNA extract. Reactions conditions were as follows: initial heat denaturation at 95°C for 5mins, followed by 50 cycles of 95°C for 30sec; 54°C (for *trn*L) and 53°C (for *rbcL*) for 30sec (annealing step); 72°C for 45sec followed by a 1°C melt curve and final extension at 72°C for 10mins.

SI III *Sequencing of DNA from bone and sediment*

All bone extracts were screened for DNA amplification. Any extracts, which successfully yielded amplicons of the target size were purified. This was done by adding five volumes of PBi buffer (Qiagen, CA, USA) and purified using a Qiagen spin column (Qiagen). Centrifugation ensured the DNA was immobilised on the spin columns and followed with wash buffer PE (Qiagen) and an extra dry spin at 13 000*g*. The DNA was eluted from the silica with 30 µl of 10 mM Tris pH 8.0 (Sigma, Mo, USA). Purified DNA products were packaged for sequencing at a commercial facility, Macrogen (Seoul, South Korea) using BigDye ver. 3.1 (Applied Biosystems) chemistry on a cycling ABI3730XL capillary sequencer (Applied Biosystems). Samples that were found to have a mixed chromatograph reading were prepared for HTS.

DNA extracts chosen for HTS were assigned a unique six base pair (bp) DNA tag (specifically a Multiplex Identifier-tag, MID-tag) and built into primers as fusion tags. All fusion-tagged PCRs were carried out in 25 µl reactions containing; 1X PCR Gold Buffer (Applied Biosystems), 2.5 mM $MgCl_2$ (Applied Biosystems), 0.4 mg/mL BSA (Fisher Biotech, Aus), 0.25 mM of each dNTP (Astral Scientific, Aus), 0.4 µM of forward and reverse primer, 0.25 µL AmpliTaq Gold (Applied

Biosystems), 0.6 µL SYBR Green (1:2,000, Life Sciences gel stain solution) and 2 µL of template. The initial heat denaturation was carried out at 95°C for 5mins, followed by 50 cycles of 95°C for 30sec, annealing temperature varied 54°C - 60°C, then 72°C for 45sec followed by final extension at 72°C for 10mins. Tagged PCR amplicons were generated in triplicates and pooled together depending on the C_T values generated on the qPCR (Applied Biosystems StepONE software version 2.00). This minimised the effects of PCR stochasticity on low template samples. The resultant pooled amplicons were purified using Agencourt AMPure XP PCR Purification Kit (Beckman Coulter Genomics, NSW, Aus), according to manufacturers' instructions and eluted in 30 µL of 10 mM Tris pH 8.0 (Sigma, Mo, USA). Purified amplicons were electrophoresed on 2% agarose gel (Bio-Rad) and pooled in approximately equimolar ratios based on ethidium-stained band intensity. After pooling the tagged amplicons, the library was then quantified with calibrated standards using qPCR (Bunce *et al.*, 2012) to determine appropriate templating ratios for HTS. Emulsion PCR and GS Junior 454 Sequencing were performed as per Roche GS Junior protocols for amplicon sequencing (http://www.454.com).

SI IV *Data analysis and sequence identification*

All sequence reads generated from the two platforms (GS Junior 454 and Ion Torrent PGM) for both bone and sediment extracts were filtered in a similar manner. At first sequences were sorted into sample batches based on the unique MID-tags using Geneious v5.6.5 (Drummond *et al.*, 2011). Tags and primers were trimmed from the sequences allowing for no mismatch in length or base composition. All sequence results that were seemingly less than the expected amplicon size depending on the primer set used, were removed from the analysis. Sequences were searched without a low complexity filter, with a gap penalties existence of five and extension of two, expected alignment value less than 1e-10 and a word count of seven. This was automated in the internet-based bioinformatics workflow environment, YABI (Hunter *et al.*, 2012). The BLASTn results obtained using YABI were imported into MEtaGenome Analyzer v4 (MEGAN) (Huson *et al.*, 2007), where they were taxonomically assigned using LCA-assignment algorithm (min. bit score = 65, top percentage = 10%. Min. support = 1). After sequences were processed, species identified were investigated to determine whether or not they currently or previously

occur in the study region. A combination of FloraBase [http://florabase.dec.wa.gov.au/], eFloraSA [http://www.flora.sa.gov.au] and Atlas of living Australia [http://www.ala.org.au/] were used to facilitate the analyses (Table 4.2.2).

4.3 CHAPTER SUMMARY

This chapter builds upon the methodological advances described in Chapter Two and Three, and further shows the utility of advanced sequencing platforms such as Roche GS-Junior and Ion Torrent (PGM) in ancient DNA research. In this chapter ancient DNA preservation at KHC proved to be exceptional – and could be successfully amplified from all 15 layers of the site, including mammal DNA from bones and plant DNA from sediments. The aims of this study were achieved (from those described in section 4.1) and the results show that there is potential for future work to be carried out on bones from still deeper layers. The innovative findings of this chapter are summarised below:

1) A thorough assessment of DNA preservation from late Pleistocene-Holocene fossils and sediments was successfully conducted.
2) The use of aDNA to successfully identify taxa complementing the morphological studies previously carried out in KHC.
3) The first genetic attempt to use aDNA to study the fauna and flora of KI, facilitated through the use of high-throughput sequencing platforms.
4) Ancient DNA has shown to be a useful tool for providing insights into past biodiversity, and should therefore be instrumental in informing conservation policies and habitat restorations.

The results from this chapter overlayed upon the palaeogeographic history of KI, provides a model context for studying mainland-island interactions (see more discussions in Chapter Six). No other account to date has used molecular approaches to study the fauna and flora on Kangaroo Island, therefore, making the results from this study of importance for island conservation management. The potential for future work to be carried out on this site is highly attractive given the excellent DNA

preservation in deep layers with ages exceeding >20 ka. Isolating aDNA from the deeper layers in the site may however require more complex DNA isolation methods. Such techniques may include enrichment of target DNA using single and/or double stranded DNA library preparations. These methods will be looked at in more detail in the following research (Chapter Five), in which a similar approach is taken (i.e. the use of HTS technology), although this time assessing the DNA preservation of multiple palaeontological cave sites across Australia.

Chapter Five

The exploration of extinct megafaunal kangaroos in Australia using capture and shotgun sequencing of aDNA libraries

Figure 5.1.1 Skull of the short-faced kangaroo used in the (Llamas *et al.*, 2015) study; found inside the Calcite Column Chasm within Mount Cripps in Tasmania's north-west region.

5.1 PREFACE

This chapter consists of a manuscript in preparation for submission to the Journal of Quaternary International, titled ***'Is it viable to obtain ancient DNA from mainland Australian macropods?'*** *Minor modifications have been made to this manuscript in order to incorporate in-thesis referencing.*

During the tenure of this thesis research, Llamas *et al.* (2015) published mtDNA sequences from the fossil remains (e.g. Figure 5.1.1 – skull of ancient kangaroo) of two extinct megafaunal macropods collected from high altitude caves in Tasmania. The researchers set out to show whether ancient mtDNA sequences could help understand macropod evolution in constructing a total-evidence phylogeny that clarifies the phylogenetic placement of the extinct megafaunal kangaroos (Figure 5.1.2). Prior to Llamas *et al.* (2015), no study had generated molecular data from ancient kangaroos. The aims of their study, overlapped with some of the long-term goals of this thesis research and it is therefore appropriate to review the Llamas *et al.* (2015) publication in the context of the data presented in this chapter. Figure 5.1.2 below (reproduced from Llamas *et al.*, 2015), depicts the two species in which ancient mtDNA was recovered from (879 bp for *S. occidentalis* and 2,383 bp *P. anak*). A meta-analysis combining all previously known work from the extinct group, i.e. all morphological and modern DNA was included to produce the total evidence tree.

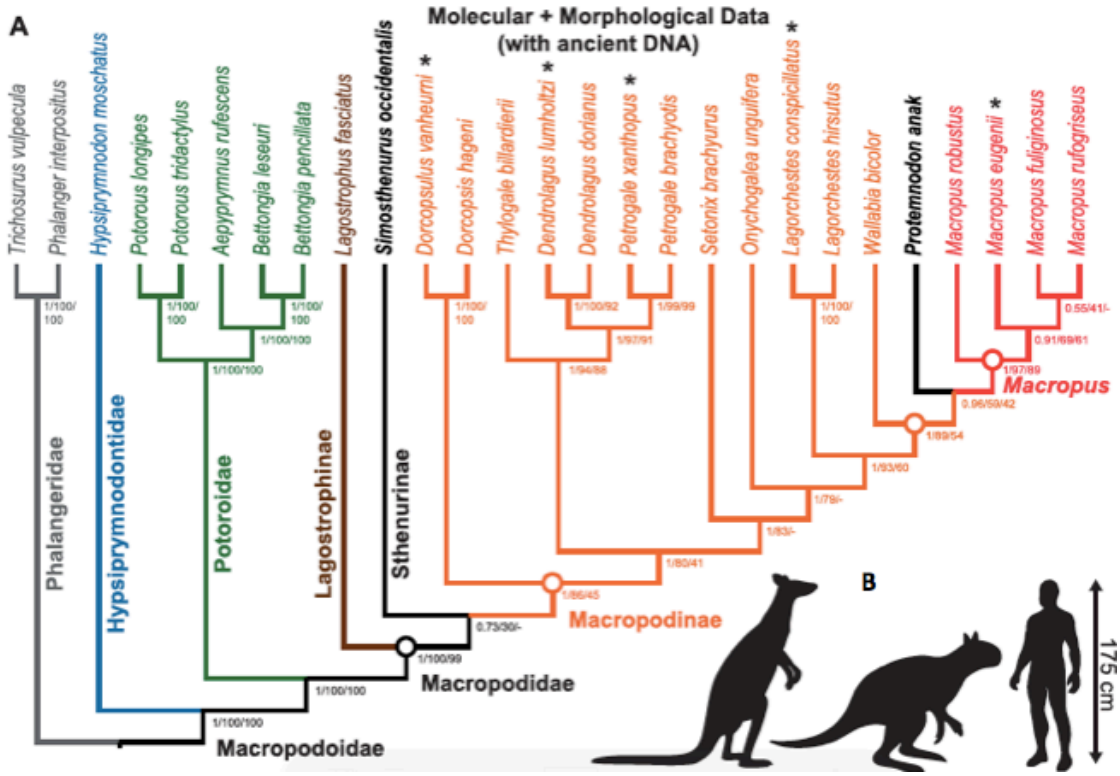


Figure 5.1.2 (A) Phylogenetic relationships of Australian macropodoids (including extinct megafauna) using molecular (aDNA) and morphological data. (B) Adult size of *Protemnodon anak* (left silhouette) and *Simosthenurus occidentalis* (middle silhouette) relative to a 175-cm-tall human. Figure is reproduced from (Llamas et al. 2015), with minor modifications.

Although Llamas *et al.* 2015 has provided some insight into the evolutionary history of Australia's extinct megafaunal kangaroos, the phylogenetic positions of both species were only supported in the total-evidence analysis, which used all morphological and molecular (including aDNA) data (Figure 5.1.2). The combined dataset placed the two species within the family Macropodidae with strong posterior support values, although the branching order between Lagostrophinae, Macropodinae, and Sthenurinae could not be robustly resolved, producing weak posterior support values. In order to fully understand the evolutionary history of these taxa, the development of a comprehensive study, including DNA information of extinct kangaroo species from mainland Australia is required. Based on the information presented in Table 5.1.1, a compilation of extinct kangaroo species taken from Johnson (2006), shows that almost 90% of Pleistocene kangaroo species, had distributions on mainland Australia, with the exception of three individuals highlighted by grey shading Table 5.1.1

Table 5.1.1 Pleistocene macropod species of Australia that did not survive into the Holocene. Species information and known distributions were modified from Johnson (2006). Sections in grey refer to species that are found in Tasmania and mainland Australia, whilst no shading denote species distributed only on the Australian mainland.

FAMILY	SPECIES	DESCRIPTION	KNOWN DISTRIBUTION
Macropodidae			
Macropodinae	*Congruus congruus*	Kangaroo	Southeastern SA
	Macropus ferragus	Kangaroo	Inland southeastern Australia
	M. pearsoni	Kangaroo	Southeastern QLD
	M. piltonensis	Kangaroo	Southeastern QLD
	M. thor	Kangaroo	Southeastern QLD
	Protemnodon anak	Kangaroo	Easter Australia, including Tas.
	P. brehus	Kangaroo	Widespread
	P. roechus	Kangaroo	Southeastern QLD
	Troposodon minor	Kangaroo	Eastern Australia
	Wallabia kitcheneri	Kangaroo	Southwestern WA
Sthenurinae	*Metasthenurus newtonae*	Short-faced kangaroo	Southern and eastern Australia, including Tas.
	Procoptodon Browneorum	Short-faced kangaroo	Southern and eastern Australia
	p. gilli	Short-faced kangaroo	Southern Vic and SA, southeastern NSW
	P. goliah	Short-faced kangaroo	Southeastern Australia, widespread
	P. oreas	Short-faced kangaroo	Southeastern QLD, eastern NSW
	P. pusio	Short-faced kangaroo	Southeastern QLD, eastern NSW
	P. rapha	Short-faced kangaroo	Southeastern Australia, widespread
	P. williamsi	Short-faced kangaroo	Inland southeastern Australia
	Simosthenurus baileyi	Short-faced kangaroo	SA
	Si. brachyselenis	Short-faced kangaroo	Central NSW
	Si. euryskaphus	Short-faced kangaroo	Northeastern NSW
	Si. maddocki	Short-faced kangaroo	Southern and Eastern Australia
	S. occidentalis	Short-faced kangaroo	Southern and eastern Australia including Tas.
	Si. pales	Short-faced kangaroo	Southern Australia widespread
	Sthenurus andersoni	Short-faced kangaroo	Southern Australia widespread
	S. atlas	Short-faced kangaroo	Inland southeastern Australia
	S. murrayi	Short-faced kangaroo	Western NSW
	S. stirlingi	Short-faced kangaroo	Inland SA
	S. tindalei	Short-faced kangaroo	Inland southern Australia

Consistent with containing a wider variety of ecological niches, the Australian mainland is richer in species diversity (Table 5.1.1). If aDNA can be recovered from mainland fossils it will be instrumental in building robust phylogenies, allowing for the extinct macropod lineages to be dated and resolved. The major aim of the study presented in this chapter, was to explore the viability of recovering macropod aDNA from a set of southern Australian mainland sites. This was carried out by: (1) identifying promising paleontological sites, suitable for aDNA analysis; (2) selecting a subset of bones, Late Pleistocene-Holocene in age; (3) screening bone samples for macropod DNA by using newly designed primer sets targeting ultra-short DNA fragments; (4) identifying samples that contain macropod DNA using qPCR and amplicon sequencing; subjecting promising samples to DNA hybridisation capture and enrichment methods; (5) shotgun sequencing the captured and enriched products using High Throughput Sequencing technology; (6) model aDNA preservation and DNA decay over time.

5.1.1 Statement of contribution

As lead author of the following manuscript in preparation, I conducted all laboratory work. This included multiple attempts at using different DNA extraction techniques and sampling an array of megafauna and macropod bones. I performed all PCR, library preparations, capture and enrichment methods, High Throughput DNA Sequencing and bioinformatic analyses. Alicia Grealy, provided assistance with the DNA decay analysis. I wrote the first and final drafts, designed all figures and tables and was provided with comments, edits and input from Prof. Michael Bunce and Dr James Haile.

Dalal Haouchar

5.2 IS IT VIABLE TO OBTAIN ANCIENT DNA FROM AUSTRALIAN MAINLAND MACROPODS?

Dalal Haouchar[1, 2] et al. (*authorship list to be determined*)

[1]Ancient DNA Laboratory, School of Veterinary and Life Sciences, Murdoch University, Perth, Western Australia, 6150, Australia

[2]Trace and Environmental DNA laboratory, Department of Environment and Agriculture, Curtin University, Bentley, WA 6845, Australia

5.2.1 ABSTRACT

Understanding the evolutionary history of Australia's extinct megafauna has been hindered by a relatively incomplete fossil record and convergent morphology. In this study, macropods from four cave sites across the Australian mainland were assessed for aDNA preservation. DNA extraction methods targeting ultra-short fragments were tested on 25 late Pleistocene to Holocene bones. Amplicon sequencing and shotgun analyses were carried out on six samples that yielded traces of aDNA, whilst DNA hybridisation capture methods combined with high-throughput DNA sequencing, was used to target endogenous marsupial DNA in the three samples suitable for mtDNA capture. The results show that pre-capture, endogenous DNA was proportionally less than the proportion of endogenous DNA post-capture, demonstrating that the mtDNA enrichment method worked on most historical samples. One fossil sample identified as a bettong species from the Nullarbor (Nailtail Cave, ~2–3 ka), captured at least 89.6% of the mtDNA genome with ~5.4X coverage. Overall, the rate of DNA decay was rapid, suggesting to little potential of megafaunal DNA survival. Australia's severe climatic conditions and Pleistocene-aged megafaunal fossils, proved to be a disadvantageous combination for ancient DNA analyses. The use of a historical control sample was vital in confirming the application of the ancient DNA techniques used, and warranted the need for more studies, employing larger sampling size and other Australian cave locations. High-quality ancient genomes of Late-Pleistocene aged macropods can provide information about demographic trajectories, which is required to increase our understanding of how species responded to climate changes, which is a key component in conservation strategies.

5.2.2 INTRODUCTION

Giant terrestrial vertebrates, dominated many of the world's ecosystems but by the late Pleistocene (c.125–12 ka), many of these species were extinct (Rule *et al.*, 2012). Australian land animals weighing more than 45 kg (with the exception of a few that were less), some weighing well over 100 kg, constituted the "megafauna". The timing and causes of their downfall have been debated for more than a century (Roberts *et al.*, 2001), with megafaunal extirpation being attributed to the events which occurred around 50 ka (Gillespie *et al.*, 2012), as this is approximately the time in which megafauna rapidly disappeared (Field *et al.*, 2001). Some hypothesise that the primary cause of extinction was successive glacial cycles preceding human arrival (e.g., Horton, 1984; Wroe *et al.*, 2006), although the most widespread view in the late nineteenth and early twentieth century was the drying of the climate, leading to increasing aridity at the end of the Pleistocene (Johnson, 2006; Field *et al.*, 2012). Others have argued for a predominantly human cause (humans arrived on the Australian continent ~50 ka), resulting in over-hunting (e.g., Johnson, 2005) and landscape burning (Miller *et al.*, 2005; Ayliffe *et al.*, 2008). Whilst it is also conceivable that different factors, or a combinations of factors, operated on different species or in different places lead to the disappearance of the megafauna (Johnson, 2006).

The superfamily Macropodoidea commonly referred to as macropods, contain at least 73 extant species in Australia (Van Dyck *et al.*, 2008). More than half the late Pleistocene extinctions were of kangaroos (family Macropodidae). The family Macropodidae is usually divided into four subfamilies: Potoroinae (rat-kangaroo), Lagostrophinae (containing the extant banded hair-wallaby, *Lagostrophus fasciatus*), Macropodinae (containing all other extant macropodids), and Sthenurinae. The Sthenurinae, was extirpated and the surviving subfamily, Macropodinae, was much reduced (Johnson, 2006). The Sthenurinae, a subfamily with six genera and 26 species (Prideaux, 2004), are thought to be a distinct Miocene radiation with recent morphological studies placing the banded hare-wallaby in an isolated lineage that is either sister to all modern macropodines (with the exclusion of sthenurines), or sister to the sthenurine macropodine clade (Prideaux *et al.*, 2012).

The problems encountered in determining the phylogenetic relationships of the sthenurine kangaroos are typical of Australia's extinct megafauna. This is mainly due to the high levels of morphological divergence, which is commonly seen in Australia's marsupial megafauna. Also, the pre-Pleistocene Australian fossil record is rather poor compared with other continents (Archer *et al.*, 1999). Recently extinct megafauna from other continents are usually conducive to DNA preservation. In Australia however, the taphonomy of Australian megafaunal fossil deposits have usually precluded the retrieval of ancient DNA (aDNA) coupled with the warm climates, a common situation on mainland Australia, making the retrieval of DNA extremely challenging. The ability to recover aDNA from extinct specimens such as megafauna, can provide important information and answer questions regarding taxonomy (e.g. Bunce *et al.*, 2003), phylogeny (e.g. Phillips *et al.*, 2013), and population dynamics (e.g. Allentoft *et al.*, 2014).

Ancient DNA can also be used to estimate the timing and impact of important events in the history of a species, such as population expansions and migrations and population-level extinction (Shapiro *et al.*, 2004). The difficulty however still remains in the retrieval of DNA from low altitudes/latitudes across the continent due to high temperatures elevating the rate of DNA decay (Allentoft *et al.*, 2012). It is not known whether the rate of fragmentation can be regarded as constant through time and also, to what extent it varies between specimens from similar depositional environments. A constant rate, however, would imply that DNA decay follows first-order kinetics (i.e. DNA has a half-life; see Figure 5.2.2), and hence that the age of a specimen can potentially work as a proxy for its DNA preservation and *vice versa* (Allentoft *et al.*, 2012). The process of DNA degradation begins soon after cell death when nucleases start to cleave the DNA into fragments (Darzynkiewicz *et al.*, 1997). During decomposition, the DNA is further degraded and digested by micro-organisms (Lindahl, 1993). Constant low temperatures play a central role in the longevity of aDNA molecules because they inhibit nuclease activity (Willerslev *et al.*, 2005). As a result, studies utilising aDNA from a range of substrates in cooler environments have been more successful than those attempting to retrieve aDNA from hot, dry environments. For example, 65 kyr-old bison mtDNA (Shapiro *et al.*, 2004), 300-400 kyr-old plant chloroplast DNA (cpDNA), and 400-600 kyr-old bacterial sequences (Willerslev *et al.*, 2003) are all samples that have been isolated from permafrost settings. In addition, permafrost-preserved bones have permitted the

amplification of DNA between 900-1000 bp in length (Barnes *et al.*, 2002), which is vastly greater then the 100-300bp size range currently being produced (e.g. Haouchar *et al.*, 2013; Murray *et al.*, 2013) from the dry and warm conditions (Hofreiter *et al.*, 2001) exposed to on mainland Australia.

In order for aDNA from Australian megafauna, especially the Sthenurines, to be explored more fully, new and improved DNA extraction, library preparation, and enrichment methods for targeted loci, need to be employed (Shapiro and Hofreiter 2014); however, success is still heavily reliant on the conditions and preservation of the source material. In this study, we screen a range of Late Pleistocene – Holocene ancient megafauna bones (ancient macropods) from four cave sites from the southern parts of Australia; this data is compared to recently published aDNA datasets. The aim of this study was to test and apply a range of new molecular techniques, that is, the capture and enrichment of highly degraded, ultra-short aDNA fragments, from historical samples. These methods were applied to bones previously collected from four Australian mainland (primarily limestone) cave sites that were selected based on their known ability to preserve aDNA and presence of a large selection of megafauna bones. This paper aims to resolve: a) whether new DNA extraction techniques and technology can be successfully used to recover megafaunal DNA in Australia? b) if the targeted capture and enrichment strategies increase the proportion of endogenous DNA recovered from the samples? and c) if it is viable using these techniques to obtain ancient megafaunal DNA from mainland Australian sites? The answer to these questions will either allow even older samples with less favorable preservation to be examined or save specimens from unnecessary destructive analysis until newer techniques are developed.

5.2.3 MATERIALS AND METHODS

5.2.3.1 SAMPLE COLLECTION

A total of 25 faunal fossil remains from four locations around Australia were selected for this study (Table 5.2.1). The selection of material consisted mainly of megafaunal bones, but also included fossil material that could not be morphologically identified to species and genus level (Table 5.2.1). A selection of macropod bones that had previously been known to yield aDNA (e.g. *Bettongia* and

Onychogalea, see Chapter Four and summarised in Table 5.2.1) were also included to perform as a 'younger' positive control in our attempts to isolate DNA from Pleistocene-aged material. Bones were also selected from a cave site that had successfully yielded aDNA prior to this study (Chapter Four). Furthermore, a selection of ancient macropod bones were chosen, that had formerly been morphologically identified to species, genus or family.

Table 5.2.1 Sample identification (ID) of the 25 fossils used in this study, location (also refer to Figure 5.2.1) and other sample information provided in this table. Site locations; Tight Entrance Cave, TEC (1); Kudjal Yolgah Cave, KYC (2); Nailtail Cave, NC (3); Kelly Hill Cave, KHC (4).

SAMPLE ID	SPECIES (based on morphological identification)	LOCATION	SAMPLE INFORMATION (e.g. units, layers, bone/teeth sampled, age of sample if available)
AD 9	*Simosthenurus brownei*	TEC (1)	Bone
AD 101	*Sthenurine*	TEC (1)	Cave walk, femur bone
AD 125	*Sthenurine*	TEC (1)	UH, 27-32 ka
AD 134	*Macropus*	TEC (1)	Layer H, grid D, 27-32 ka
AD 296	*Macropus*	KYC (2)	U2, grid A, upper charcoal <1 ka
AD 315	*Simosthenurus occidentalis*	KYC (2)	U7, south D, bone
AD 319	*Procoptodon browneorum*	KYC (2)	U7, south A, bone
AD 323	*Macropus fuliginosus*	KYC (2)	U7, south B, jaw with molars
AD 326	*Procoptodon browneorum*	KYC (2)	U7, south D, bone
AD 556	*Bettongia penicillata*	NC (3)	Jaw with molars/incisors 2-3 ka
AD 558	*Bettongia penicillata*	NC (3)	Eucla Basin Jaw with molars/incisors
AD 476	*Procoptodon*	KHC (4)	KI, Right femur bone
AD 478	*Procoptodon gilli*	KHC (4)	KI, bone
AD 479	*Procoptodon*	KHC (4)	KI, part right tibia
AD 480	*Macropus*	KHC (4)	KI, tibia
AD 952	*Macropus fuliginosus*	KHC (4)	KI, K1 C U7 S1 left astragalus, 6.8-10 ka
AD 953	*Macropodinae* sp. Indet.	KHC (4)	KI, K1 D U6 S1 caudal vertebra, 6.8-10 ka
AD 960	*Macropodinae* sp. Indet.	KHC (4)	KI, K1 U15 S1 part humerus, >20 ka
AD 961	*Macropodinae* sp. Indet.	KHC (4)	KI, K1 U15 S1 part left long bone >20 ka
AD 966	*Macropodinae* sp. Indet.	KHC (4)	KI, K1 U14 S1 part right humerus >20 ka
AD 1066	*Simosthenurus gilli*	KHC (4)	KI, K1 UA 12S1
AD 1067	*Procoptodon gilli*	KHC (4)	KI, K1 D U12S1
AD 1068	*Sthenurus gilli*	KHC (4)	KI, K1 DU 15S1 (frag. dentary) in-situ
AD 1069	*Sthenurus gilli*	KHC (4)	KI, K1 B U13S-5 jaw
AD 1730	*Procoptodon gilli*	KHC (4)	KI, K1 layer 15 left dentary in-situ >20 ka

5.2.3.2 STUDY SITES

Location one, Tight Entrance Cave (TEC; 34°00'48"S, 115°04'58"E), situated in southwestern Western Australia, was formed within the Tamala Limestone formation (Figure 5.2.1; Prideaux *et al.*, 2010). From this location, four distinct vertebrate fossils were randomly chosen from a vast collection of previously excavated bones (see Table 5.2.1 for more sample information).

Location two, Kudjal Yolgah Cave (KYC; 34°05'55"S, 115°02'45"E), is also situated in southwestern Western Australia in the Leeuwin Naturaliste National Park (Figure 5.2.1). From this location, five vertebrate fossils were selected, from a collection of previously excavated material (see Table 5.2.1 for more sample information).

Location three, Nailtail Cave (NC; 31°12'81"S, 125°24'98"E), is situated in Western Australia in the Nullarbor Plain (Figure 5.2.1). From this location, two vertebrate fossils were selected (see Table 5.2.1 for sample information).

Location four, Kelly Hill Cave (KHC; 35°83'13"S, 137°33'75"E), is situated on Kangaroo Island, South Australia. The KHC K1 site (see Table 5.2.1) was specifically chosen, as previous studies have revealed it has excellent long-term DNA preservation, back to at least 20 ka (Haouchar *et al.*, 2013). A total of 14 macropod fossils from six layers were chosen.

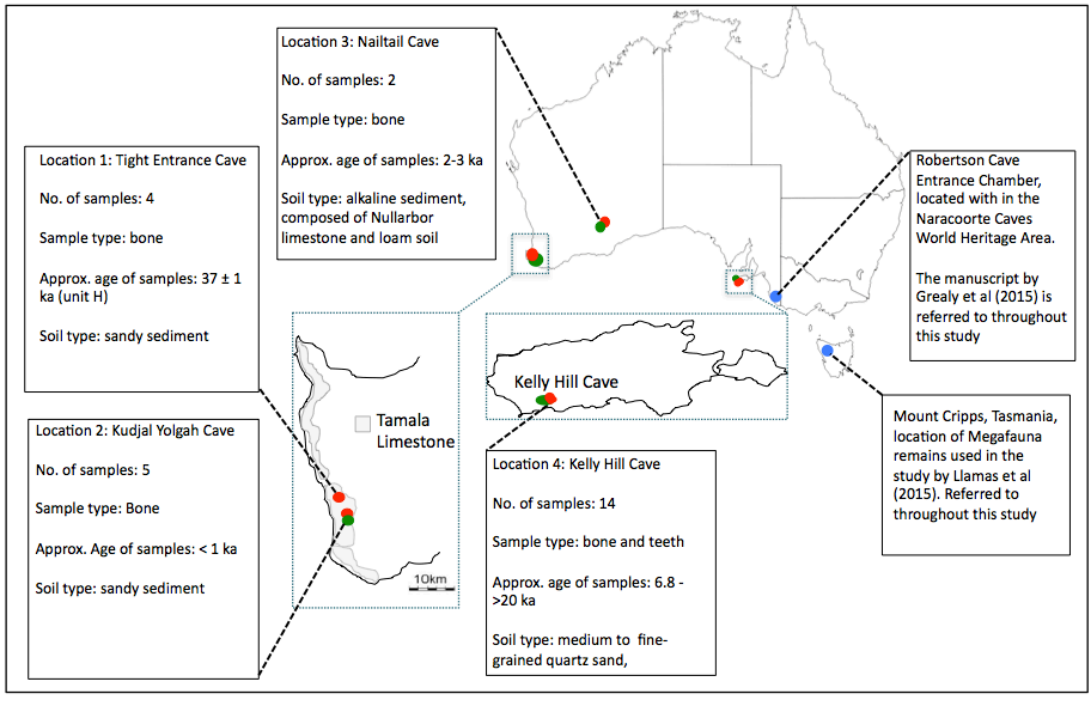


Figure 5.2.1 Map of Australia showing locations where attempts were made to isolate and capture macropod aDNA. Information about each site and the number of samples tested in each location (also refer to Table 5.2.1 for more sample information). Red dots indicate no successful mtDNA captures from the site, while green dots indicate some successful mtDNA captures. Blue dots refer to the locations of other aDNA study sites described in the text.

5.2.3.3 SAMPLE PREPARATION

All 25 samples listed in Table 5.2.1, were prepared for aDNA analysis in a dedicated ultra-clean aDNA facility (TRACE) at Curtin University, Western Australia. Following stringent procedures to guard against contamination from exogenous DNA and PCR amplicons (Willerslev *et al.*, 2005), fossil bones were firstly cleaned of surface material by shaving off the exterior surface of the sampling site. Approximately 50–100 mg of bone powder was obtained using a Dremel drill (part no. 114: Germany) and transferred to a sterile 15 ml tube.

5.2.3.4 DNA EXTRACTION

To extract highly fragmented DNA the method in Dabney *et al.* (2013) was used with minor changes as described below. DNA was eluted in 15 µl EB (QIAGEN) warmed to 37°C, followed by a second elution of 15 µl EB and the final 30 µl eluent was passed back through the column after leaving to incubate for five minutes at room temperature. Eluent was stored in a 1.5 ml Lo-Bind Eppendorf tube containing 1.5 µl of 1% TE-Tween-20. A blank extraction control was included. All extractions and downstream qPCR reactions were prepared in a physically isolated, pre-PCR ultra-clean environment following standard aDNA practice (Willerslev *et al.*, 2005; Shapiro *et al.*, 2012).

5.2.3.5 AMPLICON SEQUENCING

An initial PCR assay was carried out, whereby short specific primers (12sMarsMini, an 85 bp amplicon – see Table 5.2.2) were designed for marsupials in order to demonstrate whether ultra-short aDNA fragments could be successfully amplified. Another primer set was also used concurrently targeting a larger fragment of marsupial DNA (16sMam1, an 150 bp amplicon – Table 5.2.2).

Table 5.2.2 Primers used for the initial qPCR test pre-amplicon sequencing, and further used for amplicon sequencing as the gene specific primers. The architecture of the fusion tag indexed primers are as follows: 5' - MiSeq P5 – 5' Sequencing Adapter – Index – Forward Gene Specific Primer– 3' (forward), and 5' – MiSeq P7 – Index – Reverse Gene Specific Primer -3'(reverse).

Name	Sequence (5'-3')	Target taxa	Gene	Amplicon size (bp)	Annealing Temp (°C)	Reference
16sMam1F	CGGTTGGGGTGACCTCGGA	Mammalian (universal)	16S	150	54	(Taylor, 1996)
16sMam1R	GCTGTTATCCCTAGGGTAACT	Mammalian (universal)	16S		54	(Taylor, 1996)
12sMarsMini (Forward)	TAGTTAGACCTACACATGCAAGTT	Marsupial (universal)	12S	85	55	This study
12sMarsMini (Reverse)	CCTGATACCCGCTCCTRTTR	Marsupial (universal)	12S		55	This study

Combinations of unique indexes were used to eliminate the possibility of sequence contamination arising from previously amplified DNA. Amplicons were purified using an Agencourt AMPure XP PCR purification kit (Beckman-Coulter) at a ratio of 1.2 µl of AMPure XP SPRI beads to 1 µl of PCR product and eluted in 40 µl EB (QIAGEN). Of this eluant, 5 µl of genomic DNA from each sample was combined with 0.5 µl of 6X loading dye and run on a 3% agarose gel. The gel was visualised and photographed using GelRed and a BioRad transilluminator. The amplicons were then pooled in approximate equimolar amounts based on the intensity of the bands on the gel. The total concentration of the sequencing library was quantified via qPCR (SI I), in order to quantify the number of template molecules in the final sequencing reaction. C_T-values were recorded and compared to the standards in order to calculate the number of copies in each dilution, and determine the volume of library to input into the sequencing reaction (SI I). Unidirectional sequencing was performed on Illumina's MiSeq platform by following the manufacturer's instructions for the MiSeq 300 V2 Nano kit using a custom sequencing primer, and minor modifications (SI II).

5.2.3.6 SHOTGUN SEQUENCING AND LIBRARY PREPARATION

Shotgun DNA libraries were built on samples that showed the presence of macropod DNA. To ensure that the library build protocol outlined by Gansauge and Meyer (2013) was optimally working throughout the entire process, a positive macropod

sample (AD556; Table 5.2.1 – Holocene sample) was included (see SI III for minor modifications made to library build). After amplifying libraries with uniquely indexed fusion primers (SI IV), each library was purified using an Agencourt AMPure XP PCR purification kit (Beckman-Coulter) following manufacturer's instructions. The total concentration of each shotgun library was quantified via qPCR using a synthetic standard of known molarity and the libraries were then pooled in equimolar amounts to create a final sequencing library. The absolute concentration of the sequencing library was quantified via qPCR as above to prepare for HTS (SI I). Unidirectional sequencing was performed on Illumina's MiSeq platform following the manufacturer's instructions for the MiSeq 150 V3 kit, with minor modifications (SI V). After sequencing, the data was analysed and subjected to bioinformatic procedures (see section 5.2.3.8). Samples that showed traces of marsupial DNA were selected for mtDNA capture.

5.2.3.7 MITOCHONDRIAL DNA CAPTURE AND INDEXING

Hybridisation baits from MYselect (MYcroarray - Agilent Technologies) were designed such that 50% of the marsupial baits (probes) captured mitochondrial DNA regions and the remaining 50% capturing select nuclear DNA genes. However, the aim of this study was to only target mitochondrial genomic DNA via the hybridisation capture probes (see Table 5.2.3 for species list). Baits were produced in 100 mer lengths with 50 bp tiling density and approximately 1:1 ratio of mitochondrial:nuclear DNA. Modifications were made to the blocking components of the master mix, for compatibility with Illumina sequencing platforms. The capture protocol was carried out using a combination of methods; the MYselect Mycroarray kit combined with the protocol outlined in Li *et al.* (2013) with some modifications (see SI VI for intergraded methods).

Table 5.2.3 Species list and marsupial mitochondrial genomes used in the design of the mtDNA capture probe set. Control region sequences were removed from the mtDNA capture design due to its hypervariablity.

NCBI reference	Species	Common name	mtDNA length; minus control region
AB241054	*Dactylopsila trivirgata*	Striped possum	15462
NC_007630	*Dasyurus hallucatus*	Northern quoll	15495
NC_0022746	*Isoodon macrourus*	Northern brown bandicoot	15447
NC_008136	*Lagorchestes hirsutus*	Rufous hare wallaby	15456
AM262148	*Lagostrophus fasciatus*	Banded hare wallaby	15435
Y10524	*Macropus robustus*	Wallaroo	15468
KJ868119	*Macropus eugenii*	Tammar wallaby	15437
NC_006522	*Notoryctes typhlops*	Marsupial mole	15353
AJ639873	*Potorous tridactylus*	Long-nosed potoroo	15459
AJ303116	*Tachyglossus aculeatus*	Australian echidna	15451
FJ515780	*Thylacinus cynocephalus*	Tasmanian tiger	15389
AJ304826	*Vombatus ursinus*	Common wombat	15458

Once the hybridised libraries (the biotinylated target-bait complex) were bound to the streptavidin-coated beads, and washed accordingly to remove unbound DNA and bound non-target DNA could be removed (Li *et al.*, 2013), a post-hybridisation indexing PCR (off-beads amplification) was carried out (Li *et al.*, 2013). This off-beads amplification ensures that the captured target DNA is amplified off the target-bait-bead complex during the indexing PCR, therefore the beads are still present during the amplification step (Li *et al.*, 2013).

The indexing PCR was carried out in duplicates for 18 cycles (Li *et al.*, 2013 - also see SI VI). The replicates of each sample were combined and the PCR products were cleaned using the AMPure XP kit. A dilution series was carried out and then quantified with set standards (as described in SI I). The qPCR C_T-values were recorded and used to calculate the number copies needed for HTS.

5.2.3.8 BIOINFORMATICS AND RAW SEQUENCING ANALYSIS

The raw data, from the FastQ file generated from the Illumina MiSeq platform were downloaded and imported into Geneious version 7.1.8 for index separation and trimming (Murray *et al.*, 2013). As a form of quality control, sequences that did not perfectly match the expected index combinations were discarded from further analyses. Adapter sequences, forward and reverse primers were also removed. Abundance filtering was not performed on shotgun data, as PCR amplification was stopped during the linear phase resulting in few clonal copies. Sequences shorter than 30 bases were discarded and only unique reads were retained. Taxonomic identification was achieved by BLASTing the filtered datasets against the NCBI's GenBank nucleotide reference database via BLASTn (-F No, -e 0.01, -m Pairwise, -v 20, -b 20) (Altschul *et al.*, 1990), with searches executed in YABI (ccg.murdoch.edu.au/yabi; Hunter *et al.*, 2012). BLAST results were imported into MEGAN version 4.70.4 (ab.inf.uni-tuebingen.de/data/software/megan4; Huson *et al.*, 2007) so that they could be visualised. The LCA parameters used in MEGAN were: min support of 1, min score of 35, top per cent of 10, min complexity 0.44, win score 0. Identifications were based on the percent sequence similarity of the query to the reference across > 80% of the query, in the BLAST hits of the terminal nodes of the MEGAN tree. All sequences were categorised into the following, metatherian, human and bacterial DNA, other and no hits. All metatherian DNA sequences were extracted from the MEGAN analysis and mapped to a reference genome.

5.2.3.9 MAPPING TO A REFERENCE GENOME

Samples that successfully captured marsupial DNA were extracted from MEGAN and imported back into GENEIOUS. Because there was relevant information about the morphological identity of the bones for the successful captures, appropriate mitochondrial genomes could be chosen as reference, and aligned to the mtDNA captures. Alignments were performed in GENEIOUS, using the 'map to reference' function, and increasing iterations to up to 10 times to ensure appropriate filtering of the data. The final iteration was used for obtaining statistical data.

5.2.3.10 MODELING OF aDNA PRESERVATION AND DNA DECAY OVER TIME

According to a model of random fragmentation, the amount of amplifiable template should decline with increasing fragment size (Deagle *et al.*, 2006 - Figure 5.2.2a). Log-transformed copy numbers therefore, have a linear relationship with amplicon length, with the slope of the decline (λ) describing the probability of a bond in the DNA backbone being cleaved (Deagle *et al.*, 2006). Lambda can be converted to a damage rate, when the age of the sample is known (Allentoft *et al.*, 2012 - see Figure 5.2.2b for a theoretical example).

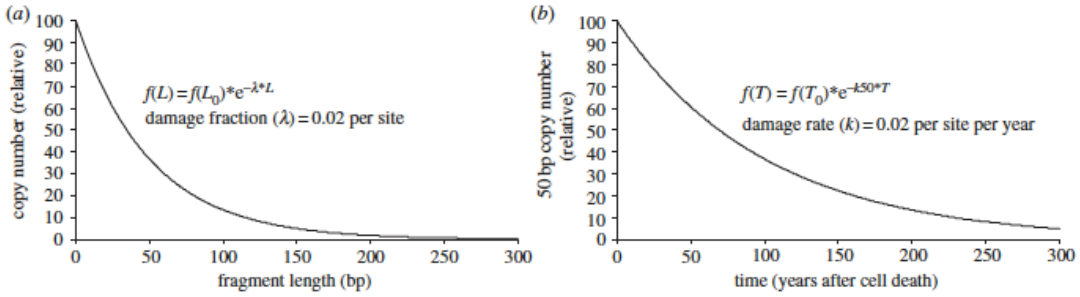


Figure 5.2.2 DNA fragmentation theory taken from Allentoft *et al.* (2012) showing the relationship caused by random fragmentation of DNA. *Post-mortem,* the template fragment length (*L*) distribution follows an exponential decline determined by the proportion of damaged sites (λ). (*b*) A hypothetical signal of temporal DNA decay. The model assumes that the observed damage fraction (λ) can be converted to a rate of decay (*k*) when the age (*T*) of a sample is known. It implies that the number of DNA copies of a given length (L) will decline exponentially with time–hence the notion that DNA has a half-life. Here, the theoretical decay kinetics of a 50bp DNA fragment, assuming a *k* of 2% per site per year. *K* is converted to a 50 bp decay rate (k_{50}), according to a Poisson distribution as: $k_{50} = 1\ (e^{-0.02*50})$.

To fit the decay model, a fragment length distribution using only shotgun metatheria sequences from sample AD556 was firstly constructed. All other samples were excluded from the analysis, as they did not contain a sufficient number of endogenous sequences to generate meaningful estimates (i.e. as the DNA becomes more fragmented, the sequences become too short and less representative of the endogenous sequences). An exponential relationship was modeled using the declining part of the distribution, excluding biases in the distribution tail (Allentoft *et al.*, 2012). This data was used to calculate the exponential co-efficient (λ) and fit (R^2) of the relationship, using Microsoft EXCEL. Once the slope of the decline (λ) and fit (R^2) were determined (Figure 5.2.5), these values were then used to establish the DNA decay rate and DNA survival over a period of time (Allentoft *et al.*, 2012).

The per nucleotide fragmentation rate per year (*k,* per site per year) was calculated for the average, minimum and maximum calibrated ages, by dividing λ by age (ybp). Other relevant information were also extracted using the estimations of *k*: the average fragment length of the DNA in the extract (bp); the number of years until the DNA is completely degraded (i.e. average fragment length = 1 bp); the coefficient of decay for both 30 bp and 100 bp products ($k_l=1-e^{-k*l}$ where l is the fragment length, from Allentoft et al 2012) and the molecular half-life of the smallest informative fragment size of 30 bp (In(2)/k_{30}). The molecular half-life corresponds to the number of years it would take 50% of the 30 bp fragments to be consumed. The proportion of a 30 bp fragment surviving 50,000 years was calculated using $e^{-k30*50,000}$ (see Table 5.2.5 for summary).

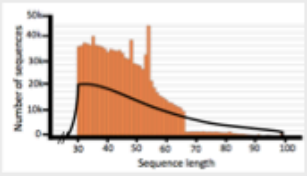
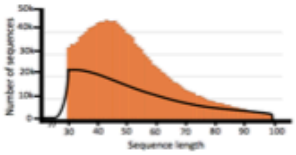
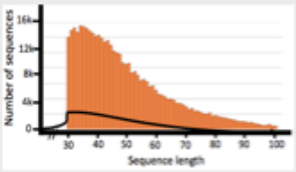
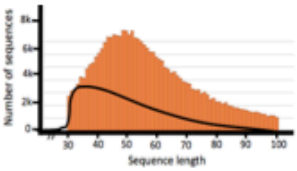
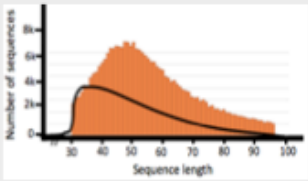
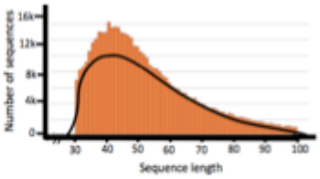
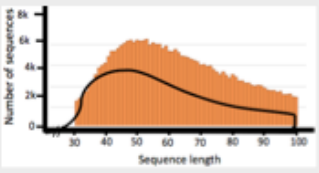
5.2.4 RESULTS

5.2.4.1 SHOTGUN LIBRARIES AND MTDNA CAPTURE

After sifting through the raw data produced from amplicon sequencing, those samples that revealed any traces of endogenous (macropod) DNA, where subject to shotgun sequencing. Building shotgun libraries provides information about the total genomic DNA, the proportion of the DNA that is endogenous (target) versus exogenous (non-target), as well the fragment length distribution of endogenous aDNA (Allentoft *et al.*, 2012). To acquire this type of information, sequencing was carried out prior to, and post mtDNA capture (Table 5.2.4). Six samples were considered for shotgun sequencing, whilst the remaining samples that did not successfully amplify marsupial DNA were no longer examined in this study. The results of shotgun sequencing pre and post capture are summarised in Table 5.2.4.

The mtDNA capture method was only successful in capturing endogenous DNA from three of the six samples, all from different locations in Australia: Nailtail Cave, Koogee Yalder Cave and Kelly Hill Cave (see Figure 5.2.1 for localities). Overall, the average length of the pre-captured sequences ranged between 47 bp to 57 bp, whilst post-captured sequences averaged between 42 bp to 58 bp. Collectively, samples that showed the presence of marsupial DNA (samples AD556, AD296 and AD952), roughly shared similar average sequence lengths (pre-capture; 52 bp, 51 bp and 48bp and post-capture; 45 bp, 43 bp and 42 bp, respectively – Table 5.2.4).

Table 5.2.4 Description of samples subject to shotgun sequencing and mtDNA capture, the number of unique reads and average sequence length. Endogenous DNA was only found in some samples after mtDNA capture as indicated in the table (also see Figure 5.2.3)

Description of source sample	Number of unique sequences pre and post-capture and average length of sequence	Sequence length distributions of shotgun and mtDNA capture Shotgun mtDNA capture	Endogenous DNA recovered YES/NO?
Sample: Bone (AD556) Morphologically identified as: *B. penicillata* Location: Nail-tail cave Nullarbor	Pre-capture: 633,352 Ave. sequence length: 52bp Post-capture: 1,063,409 Ave. sequence length: 45bp		Yes – see Fig. 5.2.3
Sample: Bone (AD296) Morphologically identified as: Macropodinae Location: Koogee Yalder Cave	Pre-capture: 785, 287 Ave. sequence length: 51bp Post-capture: 1,246,692 Ave. sequence length: 43bp		Yes – see Fig. 5.2.3
Sample: Bone (AD952) Morphologically identified as: *Macropus fuliginosus* Location: Kelly Hill Cave, Kangaroo Island, SA	Pre-capture: 94,786 Ave. sequence length: 48bp Post-capture: 37,511 Ave. sequence length: 42bp		Yes – see Fig. 5.2.3
Sample: Bone (AD319) Morphologically identified as: Procoptodon browneorem Location: Koogee Yalder Cave	Pre-capture: 1,071,831 Ave. sequence length: 52bp Post-capture: 275,500 Ave. sequence length: 58bp		No
Sample: Bone (AD315) Morphologically identified as: Sthenurine occidentalis Location: Koogee Yalder Cave	Pre-capture: 2,308,502 Ave. sequence length: 53bp Post-capture: 291,426 Ave. sequence length: 56bp		No
Sample: incisor (1730) Morphologically Identified as: Procoptodon gilli Location: Kelly Hill Cave, Kangaroo Island	Pre-capture: 1,832,270 Ave. sequence length: 47bp Post-capture: 434,241 Ave. sequence length: 51bp		No
Sample: Bone (1730) Morphologically identified as: Procoptodon gilli Location: Kelly Hill Cave, Kangaroo Island	Pre-capture: 1,737,835 Ave. sequence length: 57bp Post-capture: 401,277 Ave. sequence length: 61bp		No

5.2.4.2 EXPLORING TOTAL GENOMIC DNA IN aDNA POSITIVE SAMPLES

The three samples that contained positive traces of marsupial DNA post-capture were further explored. The filtered sequence data for the pre and post capture of these individuals (AD556, AD296, AD952), were assessed in GenBank's nr nucleotide database using BLASTn. This data was assembled into five categories and presented in Figure 5.2.3.

The shotgun results show that for both pre- and post-capture, sample AD952 contains the fewest endogenous sequences, however the proportion of endogenous DNA (pre-capture 0.029% and post-capture 1.186%, respectively), is slightly greater than sample AD296 (pre-capture 0.0007% and post-capture 0.05%, respectively), and may not have enriched further, at least as determined by shot-gun sequencing (Figure 5.2.3). Sample AD556 contains the most endogenous DNA sequences and the highest proportion of DNA (pre-capture 1.122% and post-capture 0.903%). As a result sample AD556 contains over 1500 times more endogenous DNA (pre-capture) then sample AD296, and around 38 times more endogenous DNA then sample AD952 (Figure 5.2.3). Also, the proportion of bacterial DNA between pre-capture and post-capture has shown to decrease between all samples (sample number AD556 contains 3 times less bacterial DNA, AD296 contains 1.5 times less bacterial DNA and AD952 contains 2 times less bacterial DNA). Taken together this indicates that the mtDNA MyBaits enrich endogenous DNA content.

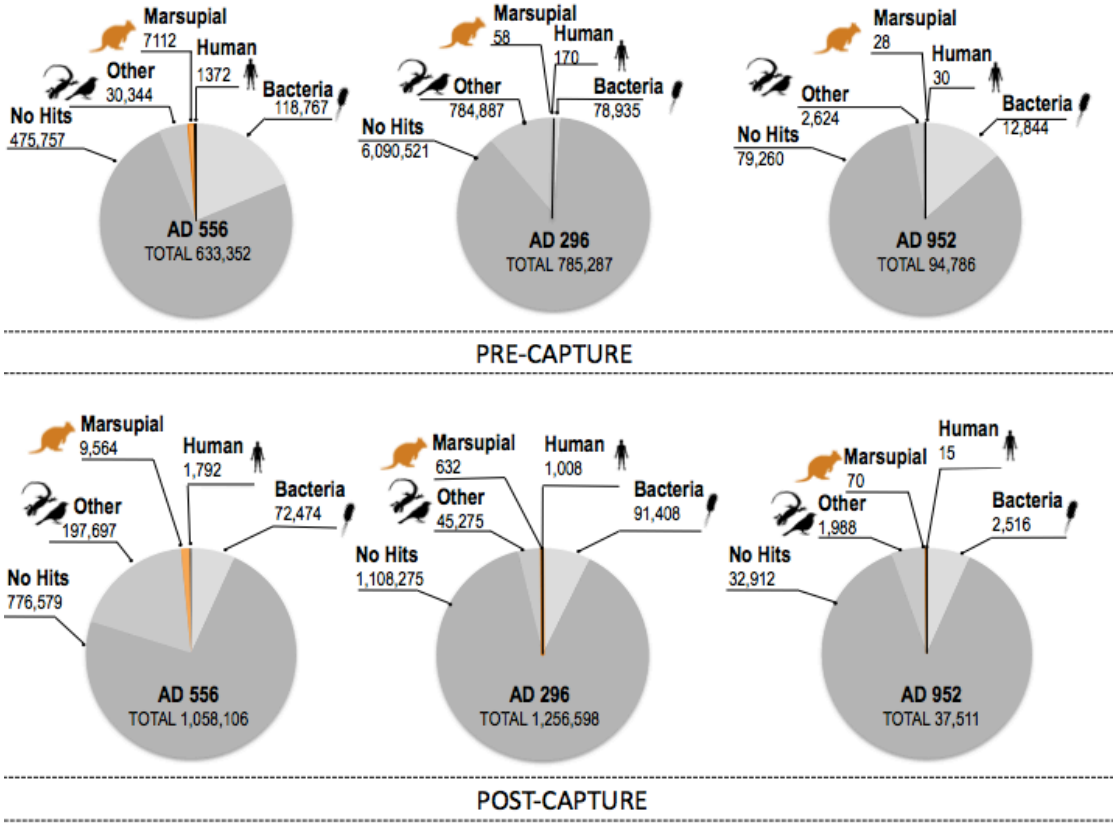


Figure 5.2.3 Pie charts of HTS reads (pre and post-capture) from three macropod fossils. Percentages of reads are shown mapping to bacterial, human, marsupial, other (other organisms), and no hits (not assigned by BLASTn to GenBank's nr nucleotide database).

5.2.4.3 MAPPING SHOTGUN DATA TO REFERENCE MTDNA GENOMES

Generally, the proportion of endogenous mtDNA in post-captured libraries increased for all samples, however, the number of endogenous sequences varied significantly between samples (see later discussion). The total number of unique sequences that were assigned to metatheria from the BLASTn database, as visualised in MEGAN, were extracted and mapped to several reference mtDNA genomes.

Sample AD556 (a bettong species) produced a total of 1,058,106 sequences (post-capture), of these 9,564 unique sequences (a combination of both mtDNA and nuDNA) were assigned to metatheria (Figure 5.2.3), and only 1,920 sequences (i.e. 20% of unique mtDNA reads) mapped to a reference (Figure 5.2.4). Each sample was mapped against at least three complete mtDNA reference genomes chosen based on the morphological identifications given to the samples prior to DNA extraction.

Sample AD556 was morphologically identified as *Bettongia penicillata*; therefore the complete mtDNA genomes (*Bettongia penicillata* – KJ868102, *Bettongia lesueur* – KJ868101, *Aepyprymnus rufescens* – KJ868095) were used for mapping. Unsurprisingly, the unique reads revealed the closest match when aligned to the *Bettongia penicillata* mtDNA and resulted in a ~5.4X coverage (Figure 5.2.4) of 89.6% of the mtDNA genome (11,631 identical sites), with a pairwise identity of 96.2%.

Sample AD296 (a *Macropus* species) produced a total of 1,256,598 sequences (post-capture), of these only 632 sequences were assigned to metatheria. A large portion of the sequences remained unassigned in the BLASTn database (Figure 5.2.3). Only 187 sequences from those that could be assigned were successfully mapped to a reference. As the bone had been morphologically identified as a macropod, three macropod mtDNA genomes were used for mapping (*Macropus fuliginosus* – KJ868120, *Macropus robustus* – NC_001794, *Macropus eugenii* – KJ868119). The 3,826 identical sites mapped to a reference mtDNA genome, resulted in a ~0.8X coverage.

Sample AD952 (a *Onychogalea* species) produced a total of 37,511 sequences (post-capture), of these only 70 sequences were assigned to metatheria (Figure 5.2.3), again a large portion were unassigned via BLASTn searches. Of these 70 sequences, only 23 could be successfully mapped on to a reference mtDNA genome. This sample was previously used in another study and was at first morphologically identified as *Macropus fuliginosus,* although was later genetically shown to be the highly endangered *Onycholgalea fraenata* (Haouchar *et al.*, 2013). As a result, the following mtDNA genomes were used for mapping (*Onychogalea unguifera* – KJ868133, *Macropus fuliginosus – KJ868120, Petrogale xanthopus* – KJ868141). The 771 identical sites mapped to a reference mtDNA genome, resulted in a ~0.3X coverage.

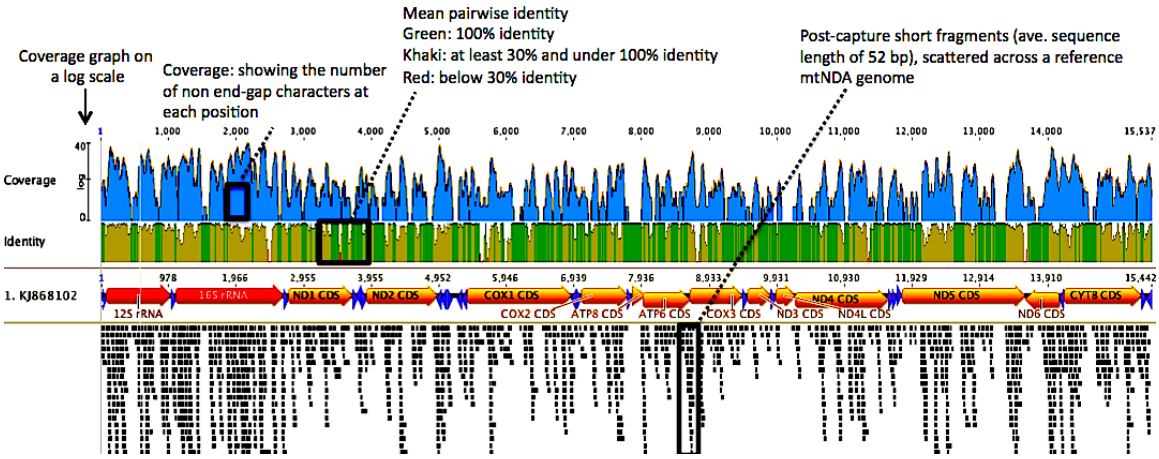


Figure 5.2.4 MtDNA map showing 5.4X coverage of unique endogenous sequences (1,920) of sample AD556 (a bettong species), mapped to a reference genome (*Bettongia penicillata* – KJ868102).

5.2.4.4 DNA DECAY AND PRESERVATION FROM MAINLAND AUSTRALIAN SITES

The distribution of fragment lengths of informative (>30 bp) endogenous (marsupial) DNA (sample AD556) shows that the mean fragment length was around 49 bp (Figure 5.2.5). Using the fragment length distribution observed in Figure 5.2.5, the per nucleotide fragmentation rate, k_{av}, was estimated to be 3.12×10^{-5} (Table 5.2.5). Due to truncation (and removal) of small fragments at the DNA extraction and data filtering stages the mean theoretical length (i.e. the 'real mean') of aDNA fragments in the extract would be estimated to be approximately 12 bp (Table 5.2.5).

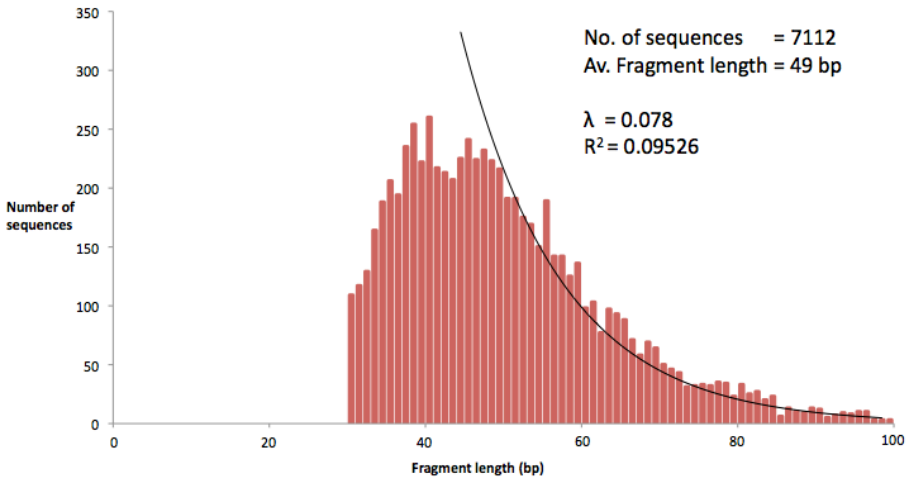


Figure 5.2.5 Sample AD556 (a bettong species) fragment length distribution of unique shotgun reads (7112 sequences represented by orange bars), that mapped to marsupial sequences by BLASTn, from a total of 633,352 sequences. The slope (λ) and fit (R^2) of the exponential relationship (represented by black slope), and average fragment length (bp) is also shown.

The decay rate for sample AD556 was used to estimate the molecular half-life of a 30 bp and 100 bp DNA fragment (Table 5.2.5) and the likelihood of these DNA sequence lengths surviving through time. Using the DNA decay rate model, and factoring in the suggested timeframe of 50,000 years (around the time of the megafaunal extinctions), the average proportion of unique marsupial DNA (30bp in length) surviving after 50,000 years is expected to be around $5.86x10^{-18}$ (Table 5.2.5). Results also show that the molecular half-life (years) of a 30 bp fragment, (30 bp being the smallest informative fragment size), corresponding to the number of years it would take 50% of 30 bp fragments to be consumed, which is likely to occur around 740 years (Table 5.2.5).

Table 5.2.5 Empirical decay (k_{av} per site per year) for sample AD556 modeled using λ as determined from the exponential section of the fragment length distribution (Fig. 5.2.5) and the average, maximum and minimum estimates of bone age, where k = λ/age(Allentoft *et al.*, 2012). The decay constant was not calculated for any other sample, because the sample size (i.e. number of fragments in the distribution) was deemed too small to accurately estimate k.

PARAMETER	MODEL	SAMPLE AD556
Av. empirical decay rate, k_{av} (per site per year) (Min – Max)	λ/ (min – max) age λ = 0.078 (see Fig. 4 for exponential coefficient plot)	$3.12\text{x}10^{-5}$ ($2.60\text{x}10^{-5}$ – $3.90\text{x}10^{-5}$)
Av. fragment length of extract (bp)	1/lambda (λ)	Approx. 12 bp (based on the removal of sequences < 30 bp)
Av. k_{30} (per site per year) (Min – max)	$k_{30}=1-e^{-k*30}$	$9.35\text{x}10^{-4}$ ($1.17\text{x}10^{-3}$ – $7.79\text{x}10^{-4}$)
Av. k_{100} (per site per year) (Min – Max)	$k_{30}=1-e^{-k*100}$	$3.24\text{x}10^{-3}$ ($3.89\text{x}10^{-3}$ – $2.59\text{x}10^{-3}$)
Av. half life (year), 30 bp (Min – Max)	$(\text{In}(2)/k_{30})$	740 (592 – 888)
Av. number of years until the av. fragment length is 1 bp (Min – Max)	$1/k$	32,051 (25,641 – 38,462)
Av. proportion of 30 bp fragments left after 50,000 years (Min – Max)	$e^{-k30*50,000}$	$5.86\text{x}10^{-18}$ ($4.06\text{x}10^{-26}$ – $1.17\text{x}10^{-17}$)

5.2.5 DISCUSSION

5.2.5.1 ANCIENT DNA PRESERVATION FROM FOUR AUSTRALIAN MAINLAND SITES

It is clear from the limited number of aDNA studies carried out on Australian fossils (e.g. Haouchar *et al.*, 2012 - Chapter Two; Murray *et al.*, 2012; Haouchar *et al.*, 2013 - Chapter Four; Murray *et al.*, 2013 - Appendix E; Grealy *et al.*, 2015; Llamas *et al.*, 2015) successful recovery of aDNA is achievable, however, is exceedingly difficult. Like many aDNA studies in temperate regions the fragmented nucleic acids are difficult to obtain requiring many methodological steps, and stringent procedures to avoid contamination (Hofreiter *et al.*, 2001). Numerous studies from all over the world have documented that warm climates are one of the major reasons accelerating DNA damage, resulting in poor DNA preservation (e.g. Hoss *et al.*, 1996; Hofreiter *et al.*, 2001; Gilbert *et al.*, 2005; Mitchell *et al.*, 2005; Rizzi *et al.*, 2012). In Australia, cave systems are common (Field *et al.*, 2013) and represent an environment more conducive to long-term DNA persistence (Stone, 2000) as they offer a cool, often dry, environment that buffers temperature and possibly pH (Elsner *et al.*, 2014). In this study, four Australian cave sites were investigated and the potential of capturing ancient megafauna DNA was explored. All four sites have recently been documented to contain megafaunal remains (Ayliffe *et al.*, 2008; Faith *et al.*, 2011; Field *et al.*, 2013; McDowell, 2013) making them suitable locations for this study. No studies to date have reported the successful capture of ancient megafaunal DNA from mainland Australia (although see Llamas *et al.*, 2015, Tasmanian example). As a result, capturing aDNA from Late Pleistocene Australian megafaunal fossils clearly represent a challenge for aDNA methodologies.

5.2.5.2 IS IT POSSIBLE TO SUCCESSFULLY CAPTURE PLEISTOCENE-AGED MEGAFAUNA DNA ON MAINLAND AUSTRALIA?

Since the first aDNA study (Higuchi *et al.*, 1984) the development of DNA sequencing technology has transformed the field of molecular genetics (e.g. Higuchi *et al.*, 1984; Herrmann *et al.*, 1994; Orlando *et al.*, 2003; Oskam *et al.*, 2010; Rizzi *et al.*, 2012). It is only with these targeted aDNA approaches (for instance, use of species-specific primers and hybridisation capture and enrichment), that researchers

can increasingly harness the information that aDNA has to offer. HTS methods, capable of generating billions of sequence reads per run have greatly advanced paleogenomic research including the generation of entire genomes from samples thousands of years old (e.g. Miller *et al.*, 2008; Taberlet *et al.*, 2012). The conversion of DNA fragments into either single or double stranded DNA libraries is widely used as the method of choice for aDNA studies employing HTS (Gansauge *et al.*, 2013; Murray *et al.*, 2015). These techniques are very effective when used concurrently, and are ideally suited to degraded samples, a characteristic common to most Pleistocene-aged megafaunal fossils (Dabney *et al.*, 2013). Shotgun sequencing also provides important information about the quality of the extract. For example, it allows analysis of the composition of a DNA extract and the determination of the relative abundance of endogenous DNA as shown in Figure 5.2.3 (Knapp *et al.*, 2010), and is a crucial aspect when using highly degraded, historical samples as they become vulnerable to exogenous contamination.

When comparing this study with other aDNA studies in Australia, we see a similar pattern of the proportion of yielded DNA in the sample. In this study, sample AD952 was the oldest fossil (6.8–10 ka) to yield aDNA and contained the fewest unique endogenous DNA sequences (Figure 5.2.3). Sample AD556 was relatively young (2–3 ka old) and contained the most unique endogenous DNA sequences. Similarly, the study by Grealy *et al.* (2015; see Figure 5.2.1 for study site), which also utilised similar aDNA methodologies (i.e. amplicon and shotgun sequencing of fossilised bulk bone from Naracoorte Cave, South Australia), found that samples older in age contained a lower proportion of endogenous DNA, and higher proportion of bacterial DNA (Grealy *et al.*, 2015). Whilst these results demonstrate a general trend, sample AD296, which is younger (<1 ka) than AD556 contained fewer unique endogenous sequences then sample AD556. The amount of endogenous DNA recovered from one sample is, to a degree, stochastic, and can be due to a number of factors, such as burial environment and skeletal element sampled.

Estimating DNA decay rates from shotgun sequencing data can help to better predict the recovery of DNA from ancient macropods. Similar to this study, DNA decay rates from certain layers in Naracoorte cave were used to estimate the molecular half-life of a 30 bp DNA fragment and the likelihood of such a minimally informative DNA sequence surviving through time (Grealy *et al.,* 2015). The results

show that the probability of obtaining informative aDNA sequences from extinct megafaunal bones at Naracoorte caves, is theoretically very low. This is also confirmed by the decay rates model, which shows that the DNA decay rate is rapid (ranges from $1.9x10^{-3}$ to $3.0x10^{-4}$ per site per year). In this study, the rate of DNA decay ($9.35x10^{-4}$; Table 5.2.5) appears to be slower. Although we were able to successfully target marsupial endogenous DNA, the rate of DNA decay is roughly comparable to sites around Australia (although biased to those samples that actually worked), and, like the Naracoorte caves, Nailtail Cave theoretically shows little to no potential of megafaunal DNA survival. The average number of years until the average fragment length is 1 bp is *c.* 32 ka for Nailtail Cave (Table 5.2.5), *c.* 100 ka and *c.* 15 ka for Naracoorte caves Layers 3 and Layers 1, respectively (Grealy *et al.,* 2015). These dates do not fall within the megafaunal extinction window (being *c.* 39-52 ka)(Roberts *et al.*, 2001), and if it does (hence, Layer 3) the length of sequence will appear too short (<30 bp) and uninformative to be of practical use. It is clear from the studies discussed above; the rate of depurination is influenced by temperature, among other factors (Lindahl, 1993), which explains why the most extreme survival of DNA is commonly documented in regions within cooler confinements (Willerslev *et al.*, 2007). Regions that primarily experience warm climate conditions, are conducive to DNA survival, however, from the results presented in this study and of those studies aforementioned, the rate of DNA decay decreases rapidly over time (e.g. Grealy *et al.*, 2016).

A recent study by Llamas *et al.* (2015) however has recently reported the successful capture of extinct megafaunal DNA, at Mount Cripps, Tasmania (Figure 5.2.1). This is currently the only known study to date, to have captured small segments of mtDNA sequences from two extinct megafaunal macropods (*Simosthenurus occidentalis* and *Protemnodon anak*) in Australia (Llamas *et al*., 2015). In the aforementioned study, several sequencing attempts were made using a number of approaches to maximise the capture of endogenous DNA. However, the limited information contained in the short sequence lengths (average read size of 37.1 ± 8.9 for *Protemnodon* and 40.8 ± 15.8 for *Simosthenurus*), coupled with the absence of a comparison genome and low library complexity, meant that molecular data alone could not fully resolve the taxonomy. The overall average read length of samples within all three locations presented in these studies (Grealy *et al.,* 2015; Layer 1 – 46.9 bp, Layer 2 – 49.1 bp; Llamas *et al.,* 2015; 37 bp – 40 bp and this study; sample

AD556 – 49 bp, see Figure 5.2.1 for locations) confirms the nature of these highly degraded aDNA samples. Another similarity common between these studies; is that the largest proportion of sequences could not be assigned by BLASTn to GenBank's nr nucleotide database (e.g. Figure 5.2.3) suggesting a high level of background from exogenous DNA.

A significant difference that could have possibly influenced the outcome of successfully achieving extinct megafaunal DNA, between the Llamas *et al.* (2015) study, and the Grealy *et al.,* (2015) study and this one, are the sample locations. The samples used in Llamas *et al.* (2015) are from a cave on Mount Cripps within the State Forest, Tasmania (Figure 5.2.1) and is located northwest of the Central Plateau Conservation Area (CPCA – one of the coldest places in Australia). The maximum temperature in Tasmania averages 20 °C to 24 °C, whilst on mainland Australia, towards the southern Nullarbor region, the maximum average temperature is 26 °C, with maximum temperatures in the summer reaching to almost 49 °C (Australian Bureau of Meteorology). As previously discussed, the success rate of DNA amplification tends to decline with increasing average temperatures. As a result, some studies have reported a 2–4%, amplification success rate in regions exposed to arid, hot climates (Pruvost *et al.*, 2007). Whilst in more moderate temperatures, the amplification success rate increased to 23–67%, and in cool, permafrost settings, the amplification success rate was reported to be around 78% (Pruvost *et al.*, 2007). Aside from temperature being a major factor in the longevity of DNA in fossilised samples, it is definitely not the only factor effecting DNA preservation.

Post excavation conditions can also have damaging results, negatively influencing aDNA recovery. Recently excavated fossil bones contain more endogenous DNA than bones that have been stored (Pruvost *et al.*, 2007). Therefore, to ensure for a controlled preservational bias throughout this study, fossils were sourced from different (geographical) sites and locations, and from samples with differing post-excavation histories. For example, some fossils from the assemblage in Table 5.2.1 (AD556 and AD558) were sub-sampled directly from the museum collections. Whilst other samples were collected from the field and DNA was isolated within one - two years post-collection (e.g. AD9, AD101, AD125 – Table 5.2.1). One fossil was sampled in-situ (AD1730 – Table 5.2.1), unwashed and untreated, with no exposure to bare human skin. Whilst others were excavated, stored under standard aDNA storage conditions (i.e. cool and dry environment), and sampled within a month or

two after collection (e.g. AD952, AD953, AD960, AD961, AD966 – Table 1). In this study, it appears that the effect on the recovery of endogenous DNA was seen regardless of when, where or how the fossil was sampled. For example, the fossil that was collected in-situ (the collector on site was fully gloved and also adhering to all other downstream aDNA protocols Pääbo et al., 2004 – sample AD1730 – Table 5.2.1 and Table 5.2.4), and sampled directly after collection was still devoid of aDNA.

5.2.5.3 ANALYSIS OF SHOTGUN AND CAPTURE aDNA DATA

As previously discussed, all samples prepared for shotgun sequencing had previously been identified morphologically. In the case of sample AD556, previously morphologically identified as *Bettongia penicillata,* our shotgun analysis shows that based on 145 bp of a contiguous region of cytochrome *b* (contiguous – as it was constructed using 3 fragmented sections, < 50 bp each, from a protein coding region within the mtDNA genome), the sequence was 99.3% similar to *Bettongia penicillata* when re-blasted in GENEIOUS utilising GenBank's nr nucleotide database and BLASTn. This genetically confirms that the DNA (from the shotgun and captured product) is from a species of *Bettongia.* Due to the positive results retrieved from a Holocene aged sample from this location, it may be practical to re-visit this site targeting a larger subset of macropod fossils (~20–30 samples), in order to provide a better assessment of the limits to aDNA preservation in this site location. The Nullarbor Plain is a well-known arid barrier for mesic species in southern Australia, causing significant phylogenetic structuring in many species (Neaves *et al.*, 2009). Providing a detailed examination of the genetic diversity of a targeted species (such as *Bettongia*) from the Nullarbor region may provide information in helping to resolve taxonomic relationships and gene flow across this 'arid barrier', prior to European arrival (see Chapter 3 for an example). A deeper understanding of these concepts (i.e. taxonomy and gene flow), which was one of the major focuses in Chapter Three, can be facilitated with the inclusion of a larger sampling size allowing for more genetic diversity to be captured.

Working with datasets that contain a large amount of species genetic information are useful, however, in this study samples AD952 and AD296 produced a small quantity of sequence information, which were also short in fragment length. Partial ancient

mtDNA sequences, however, can provide useful information about a species past. For example, sample AD952 was genetically identified as an *Onychogalea* species from Kangaroo Island, a result that was conflicting with the previous morphological identifications (see Table 5.2.1). *O. fraenata* has been morphologically confirmed on the island from previous studies (McDowell *et al.*, 2013), however, GenBank lacks sequence data from all *Onychogalea* species (i.e. there is no *O. lunata* data – a mainland species with overlapping geographic ranges with *O. fraenata*), hampering the possibility of a definite species identification. Based on the genetic data presented in this study it is highly likely that the specimen is *O. fraenata,* as described in other genetic studies (see published manuscript – Chapter Four). Having a clear understanding of which species previously inhabited the island is important for future restorations of the correct species back on the island; as it is now highly endangered, and only persists in managed populations (Kearney *et al.*, 2012; Kingsley *et al.*, 2012).

Similarly, sample AD296 previously morphologically identified as a macropod species was chosen for analysis in this study on the basis that this site (i.e. KYC, see Figure 5.2.1) contained a large assemblage of megafaunal bones (Prideaux *et al.*, 2010). The genetic data as revealed from the very short (< 50 bp) partial regions of mtDNA, unfortunately revealed very little information about the species identity, although the genetic data did show that this sample was in fact a member of the *Macropus* genus, as GenBank assignments showed its closest match with *Macropus eugenii* with 95 – 97% nucleotide similarities. The ability to extract a larger subset of nucleotide information (i.e. complete mitochondrial genomes – using similar molecular approaches taken in this study) would facilitate current morphological studies.

5.2.5.4 THE ONGOING HUNT FOR PLEISTOCENE-AGED ADNA IN AUSTRALIA.

Regardless of all the aspects contributing to the survival of DNA, ancient DNA analysis in Australia continues to be an enormous methodological and conceptual challenge for paleogeneticists. However, the results presented in this study (i.e. the *Bettongia*, *Onychogalea* and Macropod fossils) confirm that it is possible to capture endogenous DNA from old and degraded samples in Australia, although the question remains as to what are the limits to DNA preservation going back into the Late Pleistocene on mainland Australia?

Analysis of the 25 fossils inspected here; the results of Grealy *et al*. (2016 - Naracoorte) and; Llamas *et al*. (2015 - Tasmania), indicate that Pleistocene-aged aDNA in Australia will continue to be a challenge. The current molecular techniques used in this study (isolating ultra-short DNA fragments), and more specifically the DNA capture methods, which is now becoming a widely used technique common in aDNA research (Der Sarkissian *et al.*, 2015), provides an important new approach for the continual exploration of aDNA from Pleistocene sites. Furthermore, the techniques used herein, were found to be comparable with those in other studies assessing DNA preservation (e.g. Grealy *et al.*, 2016) or attempting to retrieve megafaunal DNA in Australia (Llamas *et al.,* 2015). Revisiting Table 5.1.1 presented at the front-end of this chapter, it is clear that the majority (~90%) of the Pleistocene kangaroo species existed on the mainland. The results from the Llamas *et al.,* (2015) study, could not fully resolve the branching order between Lagostrophinae, Macropodinae, and Sthenurinae. Therefore, in order to fully attain a greater understanding of the evolutionary history of these extinct taxa, it is important to direct the molecular approaches used in this research (which are similar to those used in the Llamas *et al.* 2015 study), to assess DNA preservation of megafaunal bones collected from palaeontological assemblages found on mainland Australia.

Continual sourcing of recent/freshly-excavated material coupled with ongoing exploration of variable sites around Australia should be the aim in future studies. Screening individual samples for aDNA preservation (as conducted in this chapter) can be time-consuming. However a current technique that can help to screen and survey, hundreds of bones in an efficient and cost-effective manner is the bulk-bone metabarcoding tool (see Appendix E – Murray *et al*., 2013; for work that I have co-authored). Using this approach not only allows a large number of samples to be screened, but also helps to explore differences in DNA preservation between cave locations and between layers within caves (e.g. Grealy *et al.*, 2016) – collectively this data may facilitate predictive DNA preservation maps to be constructed.

5.2.6 CONCLUSIONS

Using the latest DNA extraction procedures, library building methods and capture protocols, specifically designed to deal with short and damaged aDNA (Dabney *et al.*, 2013; Gansauge *et al.*, 2013), we were able to retrieve genomic information from samples as old as 6.8 – 10 ka, showing that genetic data from Pleistocene specimens

remains a challenge. The molecular techniques used in this study and put to the test on an ancient bettong sample, concurrently analysed alongside the megafauna fossils, we were able to successfully capture 89.6 % of mtDNA from a historical *Bettongia penicillata* specimen with 5.4X coverage. Partial mtDNA genomes were also recovered from a species of *Onychogalea* and *Macropus*. This study demonstrates that successful recovery of megafaunal DNA is highly dependent on the individual sample, and that many environmental factors play a role in the outcome. Further screening of sites for aDNA preservation will maximise the chance of finding mainland Australian sites that are able to preserve Pleistocene-aged macropods.

5.2.7 REFERENCES

Allentoft, M. E., M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, T. M. P. Gilbert, E. Willerslev, G. Zhang, P. Scofield, R. N. Hodaway & M. Bunce (2012). The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of Royal Society B-Biological Sciences* **279**, 4724-4733.

Allentoft, M. E., R. Heller, C. L. Oskam, E. D. Lorenzen, M. L. Hale, T. M. P. Gilbert, C. Jacomb, R. N. Holdaweay & M. Bunce (2014). Extinct New Zealand megafuana were not in decline before human colonization. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4922-4927.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers & D. J. Lipman (1990). Basic local alingment search tool. *Journal of Molecular Biology* **215**, 403-410.

Archer, M., R. Arena, M. Bassarova, K. Black, J. Brammall, B. N. Cooke, K. Crosby, H. Godthelp, M. Gott & S. J. Hand (1999). The evolutionary history and diversity of Australia's mammals. *Australian Mammalogy* **21**, 1-45.

Ayliffe, L. K., G. J. Prideaux, M. I. Bird, R. Grun, R. R. G, G. A. Gully, R. Jones, K. L. Fifield & R. G. Cresswell (2008). Age constraints on Pleistocene megafuana at Tight Entrance Cave in southwestern Australia. *Quaternary Science Reviews* **27**, 1784-1788.

Barnes, I., P. Matheus, B. Shapiro, D. Jensen & A. Cooper (2002). Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science* **295**, 2267.

Bunce, M., T. H. Worthy, T. Ford, W. Hoppit, E. Willerslev, A. J. Drummond & A. Cooper (2003). Extreme reversed sexual size dimporphism in the extinxt New Zealand moa Dinornis. *Nature* **425**, 172-175.

Dabney, J., M. Knapp, I. Glocke, M. T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. Garcia, S. Paabo, J. Arsuaga & M. Meyer (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15758-15763.

Darzynkiewicz, Z., G. Juan, X. Li, W. Gorczyca, T. Murakami & F. Traganos (1997). Cytometry in cell necrobiology: analysis of apoptosis and accidental cell death (necrosis). *Cytometry* **27**, 1-20.

Deagle, B. E., J. P. Eveson & S. N. Jarman (2006). Quantification of damage in DNA recovered from highly degraded samples - a case study on DNA in faeces. *Frontiers in Zoology* **3**, 11.

Der Sarkissian, C., M. E. Allentoft, M. C. Avila-Arcos, R. Barnett, P. F. Campos, E. Cappellini, L. Ermini, R. Fernandez, R. da Fonseca, A. Ginolhac, A. J. Hansen, H. Jonsson, T. Korneliussen, A. Margaryan, M. D. Martin, J. V. Moreno-Mayar, M. Raghavan, M. Rasmussen, M. S. Velasco, H. Schroeder, M. Schubert, A. Seguin-Orlando, N. Wales, M. T. P. Gilbert, E. Willerslev & L. Orlando (2015). Ancient genomics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**, 1-12.

Elsner, J., J. Schibler, M. Hofreiter & A. Schlumbaum (2014). Burial condition is the most important factor for mtDNA PCR amplification success in Palaeolithic equid remians from the Alpine foreland. *Archeological and Anthropological Sciences* **6**, 4.

Faith, T. J. & J. F. O'Connell (2011). Revisting the late Pleistocene mammal extinction record at Tight Entrance Cave, southwestern Australia. *Quaternary Research* **76**, 397-400.

Field, J. & R. Fullagar (2001). Archeaology and Australian Megafuana *Science* **294**, 5540.

Field, J. & S. Wroe (2012). Aridity, faunal adaptations and Australian Late Pleistocene extinctions. *World Archaeology* **44**, 56-74.

Field, J., S. Wroe, C. N. Trueman, J. Garvey & S. Wyatt-Spratt (2013). Looking for the archaeological signature in Australian megafaunal extinctions. *Quaternary International* **285**, 76-88.

Gansauge, M.-T. & M. Meyer (2013). Single-stranded DNA library preparation for the sequencing of ancinet or damaged DNA. *Nature Protocols* **8**, 737-748.

Gilbert, M. T. P., H. J. Bandelt, M. Hofreiter & I. Barnes (2005). Assessing ancient DNA studies. *Trends in Ecology & Evolution* **20**, 541-544.

Gillespie, R., A. B. Camens, T. H. Worthy, N. J. Rawlence, C. Reid, F. Bertuch, V. Levchenko & A. Cooper (2012). Man and megafauna in Tasmania. *Quaternary Science Reviews* **37**, 38-47.

Grealy, A., A. Macken, M. E. Allentoft, N. J. Rawlence, E. Reed & M. Bunce (2016). An assessment of ancient DNA preservation in Holocene–Pleistocene

fossil bone excavated from the world heritage Naracoorte Caves, South Australia. *Journal of Quaternary Science* **31**, 33-45.

Grealy, A. C., M. C. McDowell, P. Scofield, D. i. C. Murray, D. A. Fusco, J. Haile, G. J. Prideaux & M. Bunce (2015). A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages. *Quaternary Science Reviews* **128**, 37-47.

Haouchar, D., J. Haile, M. C. McDowell, D. C. Murray, N. E. White, R. J. N. Allcock, M. J. Phillips, G. J. Prideaux & M. Bunce (2013). Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late Pleistocene-Holocene cave deposit on Kangaroo Island, South Australia. *Quaternary Science Reviews* **84**, 56-64.

Haouchar, D., J. Haile, P. B. S. Spencer & M. Bunce (2012). The identity of the Depuch Island rock-wallaby revealed through ancient DNA. *Australian Mammalogy* **35**, 101-106.

Herrmann, B. & S. Hummel (1994). *Ancient DNA: recovery and analysis of genetic material from paleontological, archaeological, museum, medical, and forensic specimens,* New York, Springer-Verlag.

Higuchi, R., B. Bowman, M. Freiberger, O. A. Ryder & A. C. Wilson (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature* **312**, 282-284.

Hofreiter, M., D. Serre, H. N. Poinar, M. Kuch & S. Pääbo (2001). Ancient DNA. *Nature Reviews Genetics* **2**, 353-359.

Horton, D. R. (1984). *Red Kangaroos: last of the Australian megafuana. In: Martin, P. S., Klein, R. G. (Eds)*.

Hoss, M., P. Jaruga, T. H. Zastawny, M. Dizdaroglu & S. Paabo (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research* **24**, 1304.

Hunter, A. A., A. B. Macgregor, T. O. Szabo, C. A. Wellington & M. I. Bellgard (2012). Yabi: An online research environment for grid, high performance and cloud computing. *Source Code for Biology and Medicine* **7**, e1.

Huson, D. H., A. Auch, J. Qi & S. C. Schuster (2007). MEGAN analysis of metagenome data. *Genome Research* **17**, 377-386.

Johnson, C. (2006). *Australia's Mammal Extinctions: a 50000 year history*, America; Cambridge University Press.

Johnson, C. N. (2005). What can the data on late survival of Australian megafuana tell us about the cause of their extinction? *Quaternary Science Reviews* **24**, 2167-2172.

Kearney, F., R. R. J. McAllister & N. D. MacLeod (2012). Conservation and grazing in Australia's north-east: the bridled nailtail wallaby. *Pastoralism: Research, Policy and Practice* **20**, 227-244.

Kingsley, L., A. Goldizen & D. O. Fisher (2012). Establishment of an Endangered species on a private nature refuge: what can we learn from reintroductions of the bridled nailtail wallaby *Onychogalea fraenata*? *Oryx* **46**, 240-248.

Knapp, M. & M. Hofreiter (2010). Next generation sequencing of ancient DNA: requirements, strategies and perspectives. *Genes* **1**, 227-243.

Li, C., M. Hofreiter, N. Straube, S. Corrigan & G. J. P. Naylor (2013). Capturing protien-coding genes across highly divergent species. *BioTechniques* **54**, 321-326.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715.

Llamas, B., P. Brotherton, K. J. Mitchell, J. E. Templeton, V. A. Thomson, J. L. Metcalf, K. N. Armstrong, M. Kasper, S. M. Richards, A. B. Camens, L. S. Y. Michael & A. Cooper (2015). Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies. *Molecular Biology and Evolution* **32**, 574-584.

McDowell, M. C. (2013). Late Quaternary fuanal responses to environmental change and isolation on a large Australian land-bridge island. *School of Biological Sciences*). Flinders University, South Australia.

McDowell, M. C., E. A. Bestland, F. Bertuch, L. K. Ayliffe, J. C. Hellstrom, G. E. Jacobsen & G. J. Prideaux (2013). Chronology, stratigraphy and palaeoenvironmental interpretation of a Late Pleistocene to mid-Holocene cave accumulation on Kangaroo Island, South Australia. *Boreas* **42**, 974-994.

Miller, G. H., M. L. Fogel, J. W. Magee, M. K. Gagan, S. J. Clarke & B. J. Johnson (2005). Ecosystem collapse in Pleistocene Australia and a human role in megafuanal extinction. *Science* **309**, 287-290.

Miller, W., D. I. Drautz, A. Ratan, B. Pusey, J. Qi, A. M. Lesk, L. P. Tomsho, M. D. Packard, F. Zhao & A. Sher (2008). Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* **456**, 387-390.

Mitchell, D., E. Willerslev & A. Hansen (2005). Damage and repair of ancient DNA. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **571**, 265-276.

Murray, D. C., M. L. Coghlan & M. Bunce (2015). From benchtop to desktop: Important considerations when designing amplicon sequencing workflows. *Plos ONE* **4**, e10.

Murray, D. C., J. Haile, J. Dortch, N. White, D. Haouchar, M. I. Bellgard, R. Allcock, J, G. J. Prideaux & M. Bunce (2013). Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports.* **3**, 3371.

Murray, D. C., S. G. Pearson, F. Richard, B. M. Chase, J. Houston, J. Atchison, N. E. White, M. I. Bellgard, E. Clarke, M. Macphail, T. M. P. Gilbert, J. Haile & M. Bunce (2012). High-throughput sequencing of ancient plant and mammal

DNA preserved in herbivore middens. *Quaternary Science Reviews* **58**, 135-145.

Murray, P. F. (1995). The postcranial skeleton of the Miocene knagaroo *Hadronomas puckridgi* Woodburne (Marsupialia, Macropodidae). *Alcheringa* **19**, 119-170.

Neaves, L. E., K. R. Zenger, R. I. T. Prince, M. D. B. Eldridge & D. W. Cooper (2009). Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, *Macropus fuliginosus*. *Molecular Ecology* **18**, 3363-3378.

Orlando, L., J. A. Leonard, A. Thenot, V. Laudet, C. Guerin & C. Hanni (2003). Ancient DNA analysis reveals woolly rhino evolutionary relationships. *Molecular Phylogenetics and Evolution* **28**, 485-499.

Oskam, C. L., J. Haile, E. McLay, P. Rigby, M. E. Allentoft, M. E. Olsen, C. Bengtsson, G. H. Miller, J. L. Schwenninger & C. Jacomb (2010). Fossil avian eggshell preserves ancient DNA. *Proceedings of the Royal Society B: Biological Sciences* **277**, 1991-2000.

Phillips, M. J., D. Haouchar, R. C. Pratt, G. C. Gibb & M. Bunce (2013). Inferring kangaroo phylogeny from incongruent nuclear and mitochondrial genes. *PloS ONE* **8**, e57745.

Prideaux, G. J. (2004). *Systematics and evolution of the sthenurine kangaroo,* Berkeley and Los Angeles, California, University of California Press.

Prideaux, G. J., G. A. Gully, A. M. C. Couzens, L. K. Ayliffe, N. R. Jankowski, Z. Jacobs, R. G. Roberts, J. C. Hellstrom, M. K. Gagan & L. M. Hatcher (2010). Timing and dynamics of Late Pleistocene mammal extinctions in southwestern Australia. *Proceedings of the National Academy of Sciences* **107**, 22157-22162.

Prideaux, G. J. & R. H. Tedford (2012). *Tjukuru wellsi*, gen. et sp. nov., a lagostrophine kangaroo (Diprotodontia, Macropodidae) from the Pliocene (Tirarian) of northern South Australia. *Journal of Vertebrate and Paleontology* **32**, 717-721.

Pruvost, M. l., R. Schwarz, V. B. Correia, S. Champlot, S. v. Braguier, N. Morel, Y. Fernandez-Jalvo, T. Grange & E.-M. Geigl (2007). Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences* **104**, 739-744.

Rizzi, E., M. Lari, E. Gigli, G. De Bellis & D. Caramelli (2012). Ancient DNA studies: new perspectives on old samples. *Genetics Selection Evolution* **44**, 21-39.

Roberts, R. G., T. F. Flannery, L. K. Ayliffe, H. Yoshida, J. M. Olley, G. J. Prideaux, G. M. Laslett, A. Baynes, M. A. Smith & R. Jones (2001). New ages for the last Australian megafauna: continent-wide extinction about 46,000 years ago. *Science* **292**, 1888-1892.

Rule, S., B. W. Brook, S. G. Harberle, C. S. M. Turney, P. A. Kershaw & C. N. Johnson (2012). The aftermath of megafuana extinction: ecosystem transformation in pleistocene Australia. *Sceince* **335**, 1483-1486.

Shapiro, B., A. J. Drummond, A. Rambaut, M. C. Wilson, P. E. Matheus, A. V. Sher, O. G. Pybus, M. T. P. Gilbert, I. Barnes & J. Binladen (2004). Rise and fall of the Beringian steppe bison. *Science* **306**, 1561-1565.

Shapiro, B. & M. Hofreiter (2012). *Ancient DNA: methods and protocols,* New York, Humana Press.

Shapiro, B. & M. Hofreiter (2014). A paleogenomic perspective on evolution and gene function: New insights from Ancient DNA. *Science* **343**, 1-6.

Stone, A. C. (2000). Ancient DNA from skeletal remains. *Biological anthropology of the human skeleton*. (ed. by M. A. Katzenberg and S. R. Saunders), pp. 351–371. New York Wiley-Liss.

Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann & E. Willerslev (2012). Towards next generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* **21**, 2045-2050.

Taylor, P. G. (1996). Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution* **13**, 283-285.

Van Dyck, S. & R. Strahan (2008). *The mammals of Australia*, Sydney, Reed New Holland.

Willerslev, E., E. Cappellini, W. Boomsma, R. Nielsen, M. B. Hebsgaard, T. B. Brand, M. Hofreiter, M. Bunce, H. N. Poinar, D. Dahl-Jensen, S. Johnsen, J. P. Steffensen, O. Bennike, J.-L. Schwenninger, R. Nathan, S. Armitage, C.-J. de Hoog, V. Alfimov, M. Christl, J. Beer, R. Muscheler, J. Barker, M. Sharp, K. E. H. Penkman, J. Haile, P. Taberlet, M. T. P. Gilbert, A. Casoli, E. Campani & M. J. Collins (2007). Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. *Science* **317**, 111-114.

Willerslev, E. & A. Cooper (2005). Ancient DNA. *Proceedings of the Royal Society of London B: Biological Sciences* **272**, 3-16.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky & A. Cooper (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* **300**, 791-795.

Wroe, S. & J. H. Feild (2006). A review of the evidence for a human role in the extinction of the Australian megafuana and an alternative interpretation. *Quaternary Science Reviews* **25**, 2692-2703.

5.2.8 SUPPLEMENTARY INFORMATION

SI I *PCR reaction and thermocycling conditions for quantifying the sequencing library*

The sequencing library was diluted 1/10, 1/100, 1/1000, 1/5000, 1/25000, 1/125000, 1/625000 in EB buffer (QIAGEN). Each dilution was qPCR amplified in duplicate along side a standard of known concentration (10^8, 10^6, 10^5, 10^4, 10^3, 10^2 molecules) using primers complementary to the sequencing adapters, in order to quantify the number of template molecules in the final sequencing library. PCR 'no-template controls' were included. The PCR reaction contained reagents in final concentrations of: 1X ABI Power SYBR Master Mix, 0.4 µM IDT forward primer P5, 0.4 µM IDT reverse primer P7 (SI Table I), and 2 µl of library in a total reaction of volume of 25 µl (including 8.5 µl GIBCO HPLC-grade water). Thermocycling conditions were: 95°C for 5 min, 40 cycles of 95°C for 30 sec, 60°C for 45 sec, followed by a melt curve of 95°C for 15 sec, 60°C for 1 min, and 95°C for 15 sec. CT values were recorded and compared to the standard in order to calculate the number of copies in each dilution, and determine the volume of library to input into the sequencing reaction.

SI TABLE I Primers used for quantifying the final sequence library by qPCR

Name	Sequence (5'-3')	Target	Annealing Temp (°C)
P5	AATGATACGGCGACCACCGAGATCTACAC	Forward qPCR quant primer	60
P7	CAAGCAGAAGACGGCATACGAGAT	Reverse qPCR quant primer	60

SI II *Modifications to the MiSeq 300 v2 kit*

15 billion input library template molecules were targeted for the sequencing reaction using a MiSeq 300 v2 kit. 0.9 µl of the neat blend (blend included equimolar amounts of all samples) and combined with 8.1 µl of EB buffer (QIAGEN), and 1 µl of 1M molecular biology-grade NaOH and incubated for 5 minutes at 25°C, then placed on ice. 10 µl of this mixture was then added to 990 µl of HT1 buffer, and placed on ice. 550 µl of this mixture was then combined with 50 µl of 20 pM

denatured PhiX, and placed on ice. 600 µl of this mixture was added to the reagent cartridge in slot 17. Slot 12 was pierced with a pipette tip, and the contents were removed using a Pasteur pipette and placed in a 1.5 mL Eppendorf tube. 3 µl of the custom sequencing primer was added to this, and the mixture was placed back into slot 2 using a Pasteur pipette. The options selected when creating the sample sheet were: Other/ FastQ only/ Sample preparation kit = TruSeq LT/ No index reads/ Single end/ 325 cycles/ No custom primer/ No trimming.

SI III *Modifications to the single-stranded library building method by (Gansauge & Meyer 2013)*

Modifications to the adapters used are listed in SI table III. An extraction control, template (water) control, and CL104 positive control were also included in the library building process. At step 1, reactions were performed in 0.2 ml 8-well PCR strip tubes. 12 µl of DNA extract was used, and Afu UDG was replaced by Ultra-pure water. At step 7, ligation products were stored overnight at -20°C. At step 13, tubes were incubated in a rotating hybridisation oven for 2 min at 65°C as opposed to a thermal shaker. At step 13, tubes were transferred to a thermal shaker pre-cooled to 15°C as opposed to a thermocycler. Steps 14, 15, 18, 19, 23, and 25 were performed in a thermal shaker. Step 22 was performed in a rotating hybridisation oven. At step 25, the supernatant was stored in a 1.5 mL Lo-Bind Eppendorf tube at -20°C. After step 28, the PCR products were run on a 2% agarose gel electrophoresis in order to confirm the library building process worked.

SI TABLE III Ligation adapters and sequencing primers used for shot-gun sequencing (modifications from Gansauge *et al.*, 2013) The architecture of the fusion-tag indexed primers is: 5' - MiSeq P5 – RD1–Index—5' SS adapter -3' (forward), and 5' - MiSeq P7—RD2—Index—3' SS adapter - 3' (reverse). Unique indexes are represented by NNNNNNN (any base) below.

Name	Sequence (5'-3')	Function	Annealing Temp (°C)
CL53	ACACGACGCTCTTC-ddC	Double-stranded adapter, strand 1	RT
CL78	[Phosphate]AGATCGGAAG[C9Spacer]3[TEG-biotin]	Single-stranded adapter	60
CL105_106_Std	ACACTCTTTCCCTACACGACGCTCTTCCTCGTCGTTTGGTATGGCTTCTA TCGUATCGATCGATCGACGATCAAGGCGAGTTACATGAAGATCGGAA GAGCACACGTCTGAACTCCAGTCAC	Synthetic qPCR standard	-
P5-RD1-Index-Fwd adapter	AATGATACGGCGACCACCGAGATCTACAC-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-NNNNNNN ACACTCTTTCCCTACACGACGCTCTT	Forward indexing primer	60
P7-RD2-Index-Rev adapter	CAAGCAGAAGACGGCATACGAGAT-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-NNNNNNN-GTGACTGGAGTTCAGACGTGT	Reverse indexing primer	60
RD1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	Standard sequencing primer	-

SI IV *PCR reaction and thermo cycling conditions for the shotgun library fusion tag/indexing PCR*

At step 30 (Gansauge & Meyer 2013), the libraries were amplified in quintuplicate with unique fusion-tag indexing primers (SI table III). The PCR reaction contained reagents in final concentrations of: 1X ABI Power SYBR Master Mix, 0.4 µM forward indexing primer, 0.4 µM reverse indexing primer, 1 µl of neat library, made up to a total of 25 µl final volume with HPLC-grade water. Thermocycling conditions were: 95°C for 2 min, 26 cycles of 95°C for 15 sec, 60°C for 30 sec, 72°C for 1 min. Replicate reactions were combined and purified using the AMPure XP PCR purification kit (Beckman-Coulter; SI VI). Step 33 was not performed. Libraries were quantified via qPCR as per SI VIII. Libraries were pooled in equal amounts and the final shotgun sequencing library was quantified once again via qPCR as per SI I.

SI V *Modifications to the MiSeq 150 v3 kit*

At step 34 (Gansauge & Meyer 2013), a standard sequencing primer was used (SI table III). 20 billion input library template molecules were targeted for the sequencing reaction using a MiSeq 150 v3 kit. 4.0 µl of the 1/100 dilution of the library was combined with 12 µl of EB buffer (QIAGEN), and 2 µl of 1M molecular biology-grade NaOH and incubated for 5 minutes at 25°C, then placed on ice. 10 µl of this mixture was then added to 990 µl of HT1 buffer, and placed on ice. 550 ul of this mixture was then combined with 50 µl of 20 pM denatured PhiX, and placed on ice. 600 µl of this mixture was added to the reagent cartridge in slot 17. The options selected when creating the sample sheet were: Other/ FastQ only/ Sample preparation kit = TruSeq LT/ No index reads/ Single end/ 175 cycles/ No custom primer/ No trimming.

SI VI *Modifications to the mitochondrial capture protocol*

MySelect hybridisation capture was carried out on the MiSeq (Illumina), therefore new custom primers were designed and used instead the Block #3 component provided initially as 454 LibA primers. The new block #3 component for the library master mix, now consisted of P5 Forward (0.08 µl at 1.74 µM) and P7 Reverse Block (0.08 µl at 1.74 uM. Block #2 (salmon sperm) was not used in the master mix; instead only Block #1 (Human Cot-1 DNA – 2.5 µl at a final concentration of 0.27 µg/µl) was used. The library master mix was combined and 2.66 µl of this was aliquot into the PCR strip. Prior to PCR amplification 6.5 µl of library was added to the tubes and an additional 10 µl of mineral oil was added to avoid evaporation from tubes. The hybridisation and capture baits master mix followed the Myselect protocol. Thermocycler conditions for the hybridising PCR were carried out using the Li *et al.,* (2013) methods. Following the PCR, the Li *et al.,* (2013) protocol was also used for the steps involved in hybridising libraries to beads, and wash steps.

Once the beads were thoroughly washed and eluted in 50 µl of nuclease free water, an additional 1 µl of 10% Tween was suspended in the mixture. A post-hybridisation indexing PCR was carried out (off-beads amplification) using the following master mix components; 5 µl of AccuPrime PFx Reaction mix (1X), 0.5 µl of AccuPrime Pfx polymerase (0.025 U/µl), 0.5 µl of P5 – forward (0.1 µM), 0.5 µl of P7 – reverse

(0.1 µM) and the remaining volume was made up with nuclease free water, to a total of 27 µl. 23 µl of library was added to each tube and an indexing PCR was carried out for 18 cycles and the following thermocylcer conditions; 98°C for 15 sec, 65°C for 30 sec, and 72°C for 45 sec, followed by 72°C for 1 min, and hold at 4°C for 10 min.

Replicates of each sample were combined in a 1.5 µl Eppendorf tube, and followed by an Agencourt AMPure XP cleanup. The final step was eluted in 20 µl of nuclease free water. A dilution series was then carried out following the steps mentioned in section SI I, and proceeded with the quantification of products alongside a set of known standards, also see SI I. From the qPCR a calculated amount of molecules were 8.5 billion copies and 1.3 µl of captured product loaded on to the MiSeq (Illumina) platform using the 150 v3 kit.

5.3 CHAPTER SUMMARY

Chapter Five has collectively explored the practicality of capturing highly degraded aDNA from some of Australia's unique megafauna species. The combination of using new and current aDNA techniques, capture and enrichment methods, coupled with HTS technology has proved to be a useful tool in investigating the viability of isolating DNA from these species. Although this study did not successfully generate useful DNA sequence information from the megafaunal fossils (the original aim of this work), the work did however, add to the paucity of studies that are attempting to characterise aDNA from Australia. Having successfully attained almost an entire mtDNA genome (~5.4X coverage) from a fossil bettong proves that these techniques are practical, especially for those bones found in the dry and warm Australian climate, however the question remains - what are the limits to DNA preservation going back into the Late Pleistocene?

This work, coupled with recent publications (i.e. Haouchar *et al.*, 2013; Murray *et al.*, 2013; Llamas *et al.*, 2015; Grealy *et al.*, 2016) indicates that aDNA preserved under Australian conditions are possible but that it is on the limits of preservation and may be restricted to mtDNA. Utilising nuclear DNA (nuDNA) is now recommended and a common procedure to enhancing phylogenetic resolution and

dating in 'modern' phylogenies (see Phillips *et al*. 2013 study on which I am a co-author for outcomes of integrating nuDNA into macropod phylogeny). More fundamentally, moving into high-quality ancient genomes together with, genome-wide single nucleotide polymorphisms (SNPs) can provide a unique snapshot of past demographic trajectories. This type of approach can aid our current understanding of how species responded to major climatic changes in the past, and may shed light on the events leading up to first human arrival. Although this task is pertinent and is considerably the next step in aDNA research, uncertainty still remains as to whether Pleistocene aged material can only be captured from within the cooler regions in Australia (i.e. Tasmania).

Overall, in order to further explore the viability of successfully capturing DNA from bone; Pleistocene in age, from mainland Australia, requires the analysis of a much larger subset of megafaunal bones. One way in taking a practical approach in screening a large sum of bones, is using the bulk-bone method, as mentioned earlier (also see Appendix E – for work I have co-authored in, showing the development of this new method), in order to build predictive maps, focusing our energy on those locations and/or layers within sites, as a targeted approach for future DNA analysis. From here, facilitating DNA decay models, and predicting which sites are likely to contain bones conducive to DNA survival is also pertinent. Once informed decisions are made about the DNA preservation at each, the newly developed extraction methods tailored for ultra-short DNA fragments and target-enrichment capture techniques used in this study should be once again utilised.

Some of the key points described in this chapter summary, will form part of the discussions in the subsequent and final chapter (Chapter Six). Also a review of all the chapters collectively and main findings will be discussed, leading into the future directions, an important component for the continuation and exploration of macropod studies in Australia.

CHAPTER SIX – GENERAL DISCUSSION AND FUTURE DIRECTIONS

6.1 GENERAL DISCUSSION

The work presented in this thesis has collectively explored how ancient DNA isolated from Australian fossils and sediments can advance our knowledge of past biodiversity and ecosystems. Themes within this thesis research cross a range of fields including, phylogenetics, phylogeography, extinction and conservation. The rapid advancement of sequencing platforms over the tenure of this thesis research has provided new opportunities for the applications of aDNA to study some unresolved questions relating to a number of Australian macropods.

The utility of mtDNA sequencing on single-source samples, as well as bulk bone sampling methods, with the aid of current DNA extraction techniques targeting ultra-short DNA fragments, has proved to be an invaluable tool. High-throughput sequencing has been utilised in multiple ways, including amplicon sequencing (e.g. Chapters Three, Four and Five) and shotgun sequencing (e.g. Chapters Three and Five). Combining the latest HTS technology and the most recent DNA extraction techniques, this research further presents a way forward for future attempts to capture aDNA from extinct Australian megafauna.

This discussion chapter summarises how this thesis research has advanced aDNA in Australia and how it has facilitated a greater understanding of Australian macropods, especially bettongs. The chapter will also discuss future directions that ancient DNA in Australia could take.

6.1.1 SIGNIFICANCE OF THESIS RESEARCH

The extraction of aDNA from historical samples (ranging from Late Pleistocene to Holocene in age) is a major theme explored in each manuscript within this thesis (Chapters Two - Five). In Chapter One, an overview of all the (non-human) aDNA research conducted in Australia was collated (Table 1.5.1). The information in the table shows that, since the first aDNA study carried out in 1989 and now more than 20 years later, only a handful of aDNA publications have arisen from the Australian region. It is important to note that the majority of studies listed in Chapter One (Table 1.5.1), have arisen in the last five years, which is largely due to the advancements seen in the DNA sequencing technologies (i.e. HTS) coupled with the improvements in DNA extraction techniques of small and degraded DNA fragments (i.e. enrichment and capture). The following discussion will explore some of the key concepts played out in Australian mammal history, necessitating the need for a greater understanding of extinct and extirpated macropod research in Australia, which has largely been aided through the integration of aDNA data.

When humans first arrived in Australia (*c*. 50 kyr) there was a wave of extinction. Recent studies and modeling, present strong evidence for anthropogenic drivers of this extinction event (Saltre *et al.*, 2016). Since European settlement of the continent in 1788 the extinction rate has intensified further; ~22 native mammal species have become extinct (Burbidge *et al.*, 1998; McKenzie *et al.*, 2007). Since 1788 the majority of the mammal extinctions in Australia were more likely in arid and semi-arid regions, for medium-sized species, smaller macropods (e.g. Potoroidae, Macropodidae) and larger dasyurids (Dasyuridae) (Fitzsimons *et al.*, 2010; Woinarski *et al.*, 2011). In order to develop strategic responses to manage this biodiversity decline, areas of research need to ultimately focus on the most pressing matters, that is, to identify the principal cause. Conservation managers over the years have actively attempted to conserve and restore Australia's biodiversity (Hayward, 2009). However, before European settlement little of the continent's fauna was adequately documented before it became locally extinct – severely hampering the development of appropriate, well-informed conservation strategies (McDowell, 2014). Many of the mammals that became very rare (or extirpated) on mainland Australia performed valuable ecosystem services such as seed and spore dispersal, facilitation of seed germination, soil aeration, incorporation of organic matter and

improvement in moisture infiltration (Fleming *et al.*, 2013 - also see later discussions for an Australian species example). Restoration of ecosystems is a topical issue in conservation biology and a greater understanding of former composition and functional diversity of faunal communities is an important first step. The integration of Late Pleistocene and Holocene data (both morphology and genetics) can assist management and restoration decisions. As a result of the findings presented in this thesis, and in the context of temporally-informed conservation science, the following sections will briefly highlight the importance of the work presented in Chapters Two, Three, Four and Five, summarising some of the key outcomes in an attempt to explore what ancient DNA can add to our knowledge of Australia's past biodiversity.

6.1.2 CHAPTER HIGHLIGHTS AND IMPLICATIONS OF RESEARCH

6.1.2.1 CHAPTER TWO SYNOPSIS

In Chapter Two, an isolated rock-wallaby species (previously inhabiting an off-shore island off WA's Pilbara coast) was completely extirpated from its natural habitat, before any secure species identification could be made. On mainland Australia, rock-wallaby species in the Pilbara region are becoming increasingly fragmented over time, and are suffering from low levels of genetic diversity (Eldridge *et al.*, 2012). Therefore, an attempt to maintain and conserve current populations on the mainland, Depuch Island, provides a suitable site (i.e. fox-free) for the tranlocation of the native rock-wallaby species back onto the island, and is likely to have positive effects on the island's ecosystem. However, before any reintroduction attempts could be considered, a genetic characterisation of the species of rock-wallaby previously on the island needed to be confirmed. This was the major aim in Chapter Two, and is ultimately the first step taken in aiding conservationists with information required in actively managing rock-wallaby populations. The results from this chapter, clearly underline the importance of aDNA information in aiding conservationists on what species is best for future restorations and is a simple example of what aDNA can offer, as the results are becoming more widely interrogated into conservation practices (see Leonard, 2008 for a review). Appendix D is another example of a study (work which I co-authored) that uses aDNA techniques, and builds on how interpreting aDNA data can influence wildlife management decisions.

6.1.2.2 CHAPTER THREE SYNOPSIS

In Chapter Three, the research set out to explore how bettongs evolved and speciated across Australia. Members of the genus *Bettongia*, for example *B. tropica,* feeds mainly by digging up fungi ("truffles"), which in the process improves soil quality and enhances sites for seed germination. The brush-tailed bettong species are important 'ecosystem engineers' although are yet another example of an Australian mammal that has been severely affected (i.e. experienced major range contractions) since European settlement. The ultimate aim of the work presented in Chapter Three, was to increase our understanding of the evolutionary history of the brush-tailed bettongs. The integration of aDNA information into modern bettong datasets, can add a temporal context to current bettong taxonomy, which is fundamental for best-practice conservation management of these endangered taxa. Like Chapter Two, Chapter Three also uses ancient mtDNA for species identification, although is applied on a broad scale (i.e. across the Australian continent), as apposed to a single species from an offshore island. The work has resulted in an appreciation for how brush-tailed bettongs have become completely fragmented, how they were once interconnected, and the extent of biodiversity loss (also see co-authorship Appendix D that focuses on biodiversity loss in woylies). The results produced from a combination of historical and modern datasets, has provided a deeper understanding of the speciation of bettongs across Australia, which may aid the effective recovery of extirpated bettong species, back into their former distributions. A practical example of this restoration is the Mulligans Flat–Goorooyarroo Woodland Experiment, which aims to protect and recover 20 threatened mammals by 2020. *Bettongia gaimardi* and the *Bettongia penicillata ogilbyi* are included in 'The Threatened Species Strategy Action Plan 2015-2016' (Shorthouse *et al.*, 2012).

6.1.2.3 CHAPTER FOUR SYNOPSIS

In Chapter Four, the use of aDNA coupled with HTS technology is a foray into understanding ecosystems composition. Recent molecular advancements in aDNA analyses have made it possible to audit the genetic composition of past biodiversity by directly sampling DNA from Holocene fossils. The late Holoecene fossil record is however, rarely consulted by neo-ecologists, even though it can make substantial contributions to ecosystem management and restoration (McDowell, 2014). By

thoroughly assessing a fossil assemblage in Kelly Hill Cave, KI, the work set out to investigate the impact(s) of past environmental change on biodiversity. The integration of information from the fossil record added deeper-temporal dimensions to the analyses. Furthermore, the discovery of aDNA persistence in sediments is very rarely documented in Australia, and to-date this study is the first record on KI, which has helped in contributing to our knowledge of palaeo-vegetation and climate changes. A new bulk bone method, which was recently developed on caves in southwestern Australia (see Appendix E, for published work I have co-authored), was also put to the test in this chapter. Collectively the data proved to not only be a cost effective way for identifying bone fragments, but also an efficient use of those fossils that would have traditionally been disregarded (or discarded) because they are taxonomically unidentifiable. It is evident from the mtDNA results, the recovery of certain taxa, which were often not picked up in morphological analyses, were also not previously present in the paleontological record. For example, certain species of *Dasyurus* were previously known to have been on the island (e.g. *Dasyurus maculatus*). However, with the use of ancient mtDNA analyses, coupled with the results from radiocarbon ages, *D. viverrinus* was identified as being present on the island, and appeared to have been lost from KI during the mid-Holocene. Taken together the work highlights the benefits afforded by overlaying *sed*aDNA and fossil aDNA onto existing, more traditional, methodologies.

6.1.2.4 CHAPTER FIVE SYNOPSIS

The focus of Chapter Five was to evaluate how deep in time aDNA techniques could be extrapolated on mainland Australia. The initial objective of the research presented in Chapter Five, was to capture aDNA from extinct Pleistocene-aged megafaunal macropods. However, consistent with the 20-year history of ancient DNA, the 'real' result may be that DNA preservation is rare and may not be viable. In chasing elusive Pleistocene-aged megafaunal DNA sequences, a deeper understanding of the challenges of working in Australian palaeontological cave sites became apparent. Until recently (Llamas *et al.*, 2015), megafaunal macropods from the Australian continent had never been genetically characterised. The pre-Pleistocene fossil Australian record is poor compared with other continents (Archer *et al.*, 1999) and many of the extinct Australian marsupial megafauna were morphologically highly divergent, evolving into extreme forms unlike any living species and thereby

hindering phylogenetic inferences. The retrieval of aDNA for resolving the phylogenetic relationships of many living and extinct species has been instrumental in previous studies (e.g. Bunce *et al.*, 2009; Mitchell *et al.*, 2014), although the retrieval of DNA from lower altitudes/latitudes (typical in Australia) remains problematic due to high temperatures and rates of DNA decay. The major objective of Chapter Five was therefore re-focused to assess how viable it is to obtain ancient DNA from mainland Australian macropods. The successful retrieval of megafaunal DNA from the Australian mainland could shed light on the origins, taxonomy and evolutionary history of marsupials. The study presented in Chapter Five, shows that DNA preservation between sites is highly variable and it was evident that the climate conditions (e.g. temperature, humidity) played a major role on the rate of DNA decay. New HTS approaches enable ultra-short DNA fragments to be retrieved from highly degraded fossilised material and the DNA capture and enrichment methods, selectively capturing low-concentration endogenous DNA was the main method of choice. Although these techniques have been successful in previous studies (Llamas *et al.*, 2015), the results showed limited success when attempting these techniques on Late-Pleistocene bones from mainland cave sites. Collectively, Chapter Five is the first to try to understand relative preservation of DNA across Australia, although more attempts and efficient ways of screening fossil sites on a large scale are needed.

6.2 FUTURE DIRECTIONS

The rapid development of molecular biology techniques has been instrumental in the field of aDNA. As advancements are continuously being produced, new opportunities for aDNA applications are opening up new avenues into the exploration of Australia's macropods. Based on these new improvements and during the course of this thesis research, questions remained unanswered. The following four sections below highlight some areas of interest that require additional study – these four points are not meant as a comprehensive list of what should be done, rather it is a discussion of the more pressing areas of research that, time permitting, I would explore further.

6.2.1 CONTINUED INTEGRETION OF aDNA INTO AUSTRALIA'S CONSERVATION STRATEGIES

The number of studies employing ancient DNA techniques for the application of conservation and management issues in Australia are gradually increasing, as techniques for the retrieval of DNA from historic specimens have been refined. Australia's small mammals have declined rapidly since European settlement, leading to high extinction rates, habitat destruction and fragmentation. Because of these factors, the use of modern DNA analyses alone may present an incomplete representation of a species past as sampling material is becoming limited to living taxa. As a result, the retrieval of DNA from historical samples is becoming a useful tool in answering questions about a species evolutionary history. Now with the development of new aDNA techniques and HTS technology, accessing information about a species past has helped to consolidate future conservation plans, highlighting the importance of continuous aDNA research in this area.

Islands have long been pursued as secure sanctuaries, providing a refuge for declining mainland species. Finding suitable islands also remains a challenge, however current sites such as Faure Island and Dirk Hartog Island for example are important places in the Shark Bay World Heritage Area for ecosystem restoration (DEC, 2005). Kangaroo Island (explored in Chapter Four) currently retains large proportions of uncleared native vegetation, and remains free of rabbits and foxes – a vital requirement for island sanctuaries. Other islands, such Salisbury and Barrow Island that are also fox free (mainly due to their isolation), have also served as places of conservation importance (Eldridge *et al.*, 1999). Based on the current trends (e.g. predation by feral cats, the increase in inappropriate fire regimes, spread of disease and habitat change), Australia's biodiversity is diminishing at an alarming rate. As a result, species such as *B. lesueur* (the burrowing bettong), once found widespread across mainland Australia, is now confined to island refugia. Despite a number of conservation success stories (e.g. Faure Island run by the Australian Wildlife Conservancy – www.australianwildlife.org), if we do not have an understanding of what species previously inhabited these areas, may make it difficult to appropriately rehabilitate them or to assess the likelihood of successful translocations. Ancient DNA analyses provide an advantage for solving these types of questions as species (plant and animal) as well as haplotype information can be extracted.

6.2.2 FUTURE DIRECTIONS IN BETTONG RESEARCH

Bettongs were a major focus presented in this thesis research. Chapter Three explored the widespread brush-tailed bettongs across the Australian continent. Some of the data generated in Chapter Four from a bettong species found on Kangaroo Island is also integrated in to the analyses presented in Chapter Three, adding more scope to this study. In Appendix C, an extinct bettong species (*Bettoniga anhydra*), was genetically characterised for the first time, and was also integrated into the phylogenetic analysis presented in Chapter Three. Bettongs are important ecosystem engineers and the impact of their removal from the ecosystem is yet to be fully understood.

Mitochondrial DNA is a common locus in aDNA research and was used in all of the studies aforementioned to answer species-level questions. However, like many of Australia's endangered marsupials, questions regarding population-level genetics can provide crucial information about speciation events and population structure. Utilising nuclear DNA (nuDNA) markers to explore questions at a population-level can help to provide information based on gene flow and connectivity. The study presented in Appendix D (work I have co-authored) is an example of the type of information we can acquire once we start to look at nuclear historical data. To answer some of the pressing questions that have come up during the course of this research, the integration of nuDNA to strengthen current aDNA analyses is critical. Obtaining nuDNA from historical and degraded material still remains a challenge due to the lower relative abundance of nuDNA compared to mtDNA in a cell. However, nuDNA is becoming more accessible in historical material, with the developing molecular techniques and HTS platforms. Incorporating aDNA and nuDNA (shotgun sequencing, RADseq or capture) data into current-day conservation strategies is clearly a need for future bettong research – the recovery of partial mtDNA genomes and identification of nuDNA markers (Chapter Three) point towards the viability of this approach. Ultimately, information on historical nuclear gene flow is needed before meaningful decisions can be made regarding translocation – the end goal being management that is cognisant of what is a species boundary and how best to conserve what is left of the genetic diversity in the brush-tailed bettongs.

6.2.3 SYSTEMATICALLY SCREENING FOSSIL SITES IN AUSTRALIA

From a number of historical bones analysed in Chapter Four, it was clear that some samples were successfully producing relatively large fragments of aDNA (>100 bp) from layers dating >20 ka. This is a positive indication that Kelly Hill Cave, Kangaroo Island, clearly preserves DNA well and highlights the potential for ongoing sampling from deeper layers, in order to execute the megafauna deposits. Another study, successfully extracted aDNA from bone deposits at Devil's Lair, dated between 44,260 – 46,890 ka, which is genuinely the oldest aDNA recovered from Australia to date (see Appendix E for published work I have coauthored). Archeological sites, with paleontological significance such as Devil's Lair and Tunnel Cave in Western Australia, are successfully yielding aDNA sequences from dates that coincide within the megafauna extinction window and should therefore be identified as valid sites for future research (see Figure 6.2.1). Similarly, another study assessing the preservation of aDNA in historical bones from the Naracoorte Caves in South Australia (see Chapter Five for more discussion on this study), are yielding short aDNA fragments > 18 ka (Grealy *et al.*, 2016), however may not offer the potential to explore specimens older in age (i.e. Pleistocene material > 40 kyr). Finally, another recent study by (Llamas *et al.*, 2015), successfully recovered aDNA from two extinct megafauna species from the high altitude caves in Tasmania.

Although only a few aDNA studies have been documented here, aDNA research carried out in Australia is still rare (see Chapter One – Table 1.5.5 for aDNA studies in Australia) and the preservation of aDNA at each site is highly variable. Therefore developing a systematic approach to screening aDNA bones, Australia wide, is needed to concentrate efforts in efficiently sourcing cave sites with bone material conducive to DNA survival. Generating a 'predictive' DNA preservation map taking into account temporal, spatial and physical attributes may assist in future aDNA studies. Importantly for the development for such a map, the newly developed 'bulk-bone' methods (see Chapter Four and co-authorship in Appendix E) provides a mean for efficiently screening hundreds of low value bones from sites within Australia and globally. Figure 6.2.1, is a 'road map', to systematically screening a selection of temporal and spatially diverse fossil sites across Australia.

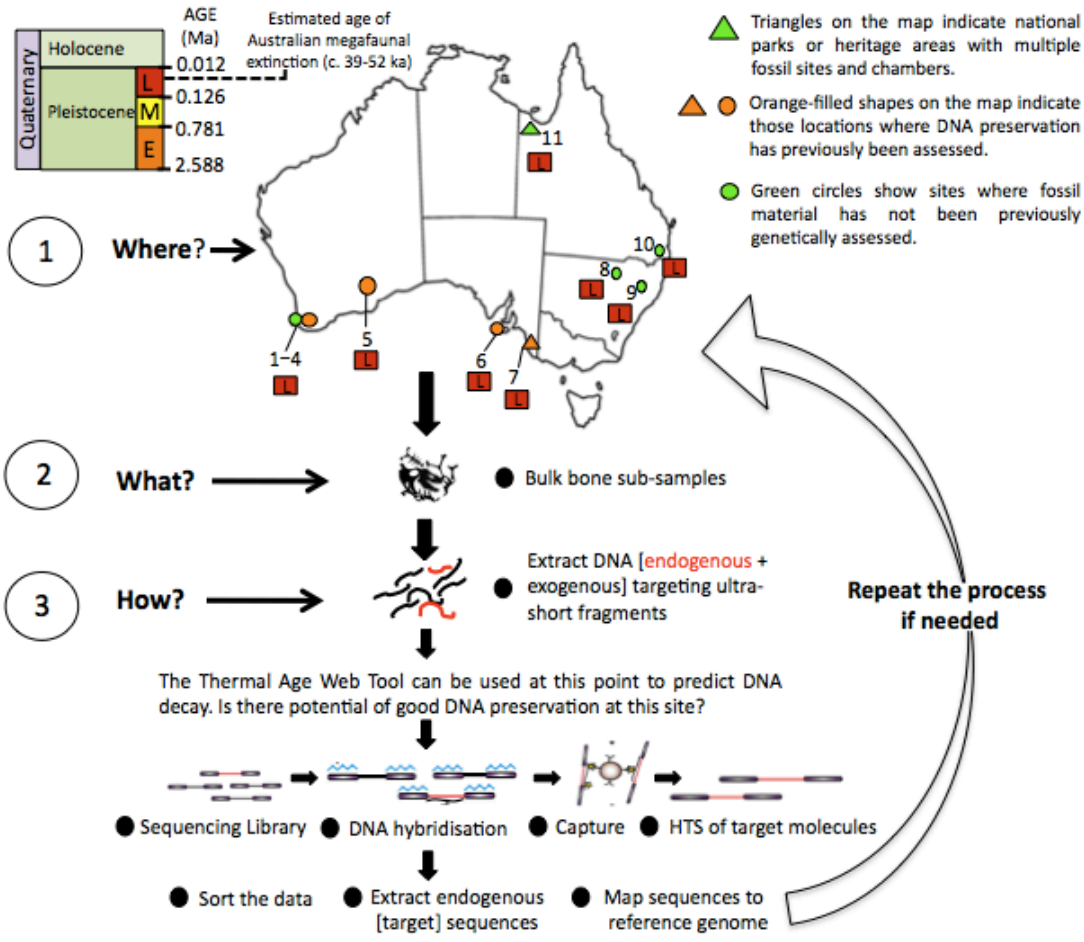


Figure 6.2.1 Predictive map and experimental design strategy for screening Late Pleistocene and Holocene Australian fossil sites. Regions on the map show some key megafaunal assemblage sites found to contain a number of extinct macropods. Site numbers: 1, Devil's Lair; 2, Kudjal Yolgah Cave; 3, Mammoth Cave; 4, Tight Entrance Cave; 5, Nailtail Cave; 6, Kelly Hill Cave; 7, Naracoorte Caves; 8, Cuddie Springs; 9, Mooki River; 10, Ned's Gully; 11, Riversleigh World Heritage Area. Green circles on the map show sites where fossil material have not been previously genetically assessed, but have megafaunal deposits formerly dated to Late Pleistocene, hence red boxes with "L" and should be genetically assessed using the experimental design outlined above. The "H" blue filled boxes, denote to areas which contain Holocene material successfully yielding DNA. Orange-filled shapes on the map indicate those locations where DNA preservation has previously been assessed, although these sites should be revisited using freshly excavated material (i.e. from the locations highlighted in step 1) targeting the megafaunal layers. An experimental design (steps 2 – 3) should then be followed, which includes [i] an initial test for the preservation of DNA at the site [ii] using a web tool predicting the rate of DNA decay and the possibility of DNA surviving into the Late Pleistocene [iii] DNA hybridisation and capture to enrich for endogenous DNA [iv] high-throughput sequencing (HTS) followed mapping/analysis.

Figure 6.2.1, highlights the limited number of sites which hold Pleistocene-aged material (Roberts *et al.*, 2001; Prideaux *et al.*, 2010), which coincide with chronologies of megafaunal extinctions (c. 39-52 ka) (Roberts *et al.*, 2001; Gillespie *et al.*, 2012). Also, some locations indicated on the map refer to areas which contain multiple cave sites, for example Victoria Fossil Cave at Naracoorte, is considered to be Australia's largest fossil site and contains some of the best-preserved fossil deposits in the world (Wells *et al.*, 1984). Riversleigh World Heritage Area is another site and contains over 300 fossil localities (Archer *et al.*, 1997), which has become one of the most important fossil sites for understanding the early evolution of modern Australasian marsupials (Cooke *et al.*, 2015). In order to explore the DNA preservation across multiple sites concurrently, the bulk bone sampling method (as indicated in Figure 6.2.1 and tested in Chapter Four) is mandatory. Furthermore, the integration of The Thermal Age Web Tool (TTAWT) has shown to be an effective measure in predicting the rate of DNA decay and assessing DNA preservation. The use of these two methods (i.e. bulk bone sampling and TTAWT) in tandem is a means by which to screen a large quantity of fossil material in a quick, cost-effective and efficient manner. Once a location reveals the potential of good DNA preservation from fossils of Late Pleistocene age, they should only then be subject to shotgun approaches or more targeted capture (DNA hybridisation - as successfully tested in Chapter Five and used in other studies, Llamas *et al.,* 2015). If a systematic screening of fossil sites in Australia can be carried out (as advocated in Figure 6.2.1) it paves the way for the integration of aDNA data into a variety of projects across a range of disciplines including; climate change, extinction processes, restoration/conservation and evolution.

6.2.4 GENES TO GENOMES

Moving away from using traditional PCR/Sanger methods and into HTS methods using capture and enrichment techniques has very quickly widened the scope of aDNA research. One of the focus areas of this thesis was to explore techniques that enable historical fossils to be genetically analysed, which would not normally be amplifiable (i.e. highly degraded megafaunal material). The development of customised 'baits', which are synthetically designed to target a group of taxa makes it a desirable feature in aDNA studies. In Chapters Three and Five these baits were marsupial specific. The benefit of this method is that, it not only targets endogenous

DNA of interest but also reduces the proportion of exogenous DNA (mostly contamination) amplified in the mixture. Therefore, the continuous uses of DNA capture and enrichment techniques; coupled with HTS technology, opens an avenue into the continuation of future aDNA studies in Australia.

This era of genomics should enable aDNA researchers to move beyond neutral molecular markers and track functional genes through time. Scaling up to genome-wide data can improve traditional conservation genetic inferences, and provide qualitatively novel insights (Shafer *et al.*, 2015). Moving Australian aDNA into the genomics field however, will remain a challenge as reference genomes are currently limiting - *Macropus eugenii* (Renfree *et al.*, 2011 - tammar wallaby) and *Sarcophilus harrisii* (Miller *et al.*, 2011 - Tasmanian devil) are the only examples of Australian species with low quality 'draft' genomes. The ability to acquire historical bones with sufficient nuclear aDNA preservation however is key to unlocking a wealth of information about extinct and extant populations. Recent theoretical and empirical work have demonstrated that genomic data collected from a small number of (or even single) individuals can provide rich insights into population history and the causative connection between alleles and phenotypes (Li *et al.*, 2011; Orlando *et al.*, 2013). Indeed, palaeogenomes are remarkably powerful lenses through which one can investigate a variety of questions from population size and connectivity through to positive selection and biodiversity loss. If the 'right' fossils can be found there is a lot of potential to explore Australia's lost biodiversity using palaeogenomics.

6.3 CONCLUDING REMARKS

Molecular methods for optimised aDNA retrieval from historical samples are continuously being refined. During the course of this thesis research the most relevant and up-to-date techniques were continuously being sourced – this thesis documents a progression from cloning and Sanger sequencing, leading into the latest HTS technology. These developments have had direct implications for providing best-practice conservation and management applications to some of Australia's macropod species. Using a combination of historical, Holocene – Pleistocene fossils and modern sequence data, this research has provided valuable insights into the conservation of rock-wallabies, bettongs and added a different perspective on the palaeo-ecosystems on Kangaroo Island. Taken together this thesis research has

shown the potential for aDNA in Australia despite conditions that are often not conducive to long-term DNA survival.

6.4 REFERENCES

Archer, M., R. Arena, M. Bassarova, K. Black, J. Brammall, B. N. Cooke, K. Crosby, H. Godthelp, M. Gott & S. J. Hand (1999). The evolutionary history and diversity of Australia's mammals. *Australian Mammalogy* **21**, 1-45.

Archer, M., S. J. Hand, H. Godthelp & P. Creaser (1997). Correlation of the Cainozoic sediments of the Riversleigh World Heritage Fossil Property, Australia; pp. 131–152 in J. P. Aguilar, S. Legendre, and J. Michaux (eds.), Actes du Congrès Biochrom97, Montpellier, 14–17 April 1997. Memoires et Travaux de l'Institut de Montpellier de l'École Pratique des Hautes Ètudes 21.

Bunce, M., T. H. Worthy, M. J. Phillips, R. N. Holdaway, E. Willerslev, J. Haile, B. Shapiro, P. Scofield, A. J. Drummond & P. J. Kamp (2009). The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proceedings of National Academy of Sciences of the United States of America* **106**, 20646-20651.

Burbidge, A. A., N. L. McKenzie & A. J. Brennan (1998). Conservation status and biogeography of Australia's terrestrial mammals. *Australian Journal of Zoology* **56**, 411-422.

Cooke, B. N., K. J. Travouillon, M. Archer & S. J. Hand (2015). Ganguroo robustiter, sp. Nov. (Macropodoidea, Marsupialia), a middle to early late Miocene basal macropodid from Riversleigh World Heritage Area, Australia. *Journal of Vertebrate Paleontology* **35**, e956879.

Eldridge, M. D. B., J. M. King, A. K. Loupis, P. B. S. Spencer, A. C. Taylor, L. C. Pope & G. P. Hall (1999). Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby. *Conservation Biology* **13**, 531-541.

Eldridge, M. D. B., M. P. Piggott & S. L. Hazlitt (2012). Population genetic studies of the Macropodoidea: a review. In Macropods: the biology of kangaroos, wallabies and rat-kangaroos. (Eds G. Coulson and M. D. B. Eldridge.) 35-51. CSIRO Publishing, Melbourne.

Fitzsimons, J., S. Legge, B. Traill & J. Woinarski (2010). *Into the oblivion? The disappearing native mammals of northern Australia*). The Nature Conservancy, Melbourne.

Fleming, P. A., H. Anderson, A. S. Prendergast, M. R. Bretz, L. E. Valentine & G. E. Hardy (2013). Is the loss of Australian digging mammals contributing to a deterioration in ecosystme function? *Mammal Review* **doi:10.1111/mam.12014**.

Gillespie, R., A. B. Camens, T. H. Worthy, N. J. Rawlence, C. Reid, F. Bertuch, V. Levchenko & A. Cooper (2012). Man and megafauna in Tasmania. *Quaternary Science Reviews* **37**, 38-47.

Grealy, A., A. Macken, M. E. Allentoft, N. J. Rawlence, E. Reed & M. Bunce (2016). An assessment of ancient DNA preservation in Holocene–Pleistocene fossil bone excavated from the world heritage Naracoorte Caves, South Australia. *Journal of Quaternary Science* **31**, 33-45.

Hayward, M. (2009). Conservation management for the past, present and future. *Biodiversity and Conservation* **18**, 765-775.

Leonard, J. A. (2008). Ancient DNA applications for wildlife conservation. *Molecular Ecology* **17**, 4186-4196.

Li, H. & R. Durbin (2011). Inferences of human population history from individual whole-genome sequences. *Nature* **475**, 493-496.

Llamas, B., P. Brotherton, K. J. Mitchell, J. E. Templeton, V. A. Thomson, J. L. Metcalf, K. N. Armstrong, M. Kasper, S. M. Richards, A. B. Camens, L. S. Y. Michael & A. Cooper (2015). Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies. *Molecular Biology and Evolution* **32**, 574-584.

McDowell, M. C. (2014). Holocene vertebrate fossils aid the managment and restoration of Australian ecosystems. *Ecological Managment and Restoration* **14**, 58-63.

McKenzie, N. L., A. A. Burbidge, A. Baynes, R. N. Brereton, C. R. Dickman, G. Gordon, L. A. Gibson, P. W. Menkhorst, A. C. Robinson, M. R. Williams & J. C. Z. Woinarski (2007). Analysis of factors implicated in the recent decline of Australia's mammal fauna. *Journal of Biogeography* **34**, 597-611.

Miller, W., V. M. Hayes, A. Ratan, D. C. Petersen, N. E. Wittekindt, J. Miller, B. Walenz, J. Knight, J. Qi, F. Zhao, Q. Wang, O. C. Bedoya-Reina, N. Katiyar, L. P. Tomsho, L. M. Kasson, R.-A. Hardie, P. Woodbridge, E. A. Tindall, M. F. Bertelsen, D. Dixon, S. Pyecroft, K. M. Helgen, A. M. Lesk, T. H. Pringle, N. Patterson, Y. Zhang, A. Kreiss, G. M. Woods, M. E. Jones & S. C. Schuster (2011). Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian devil). *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12348-12353.

Mitchell, K., J, B. Llamas, J. Soubrier, N. J. Rawlence, T. H. Worthy, J. Wood, M. S. Y. Lee & A. Cooper (2014). Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science* **344**, 898-900.

Orlando, L., A. l. Ginolhac, G. Zhang, D. Froese, A. Albrechtsen, M. Stiller, M. Schubert, E. Cappellini, B. Petersen & I. Moltke (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499**, 74-78.

Prideaux, G. J. & N. M. Warburton (2010). An osteology-based appraisal of the phylogeny and evolution of kangaroos and wallabies (Macropodidae: Marsupialia). *Zoological Journal of the Linnean Society* **159**, 954-987.

Renfree, M., A. Papenfuss, J. Deakin, J. Lindsay, T. Heider, K. Belov, W. Rens, P. Waters, E. Pharo, G. Shaw, E. Wong, C. Lefevre, K. Nicholas, Y. Kuroki, M. Wakefield, K. Zenger, C. Wang, M. Ferguson-Smith, F. Nicholas, D. Hickford, H. Yu, K. Short, H. Siddle, S. Frankenberg, K. Chew, B. Menzies, J. Stringer, S. Suzuki, T. Hore, M. Delbridge, A. Mohammadi, N. Schneider, Y. Hu, W. O'Hara, S. Al Nadaf, C. Wu, Z.-P. Feng, B. Cocks, J. Wang, P. Flicek, S. Searle, S. Fairley, K. Beal, J. Herrero, D. Carone, Y. Suzuki, S. Sugano, A. Toyoda, Y. Sakaki, S. Kondo, Y. Nishida, S. Tatsumoto, I. Mandiou, A. Hsu, K. McColl, B. Lansdell, G. Weinstock, E. Kuczek, A. McGrath, P. Wilson, A. Men, M. Hazar-Rethinam, A. Hall, J. Davis, D. Wood, S. Williams, Y. Sundaravadanam, D. Muzny, S. Jhangiani, L. Lewis, M. Morgan, G. Okwuonu, S. Ruiz, J. Santibanez, L. Nazareth, A. Cree, G. Fowler, C. Kovar, H. Dinh, V. Joshi, C. Jing, F. Lara, R. Thornton, L. Chen, J. Deng, Y. Liu, J. Shen, X.-Z. Song, J. Edson, C. Troon, D. Thomas, A. Stephens, L. Yapa, T. Levchenko, R. Gibbs, D. Cooper, T. Speed, A. Fujiyama, J. M Graves, R. O'Neill, A. Pask, S. Forrest & K. Worley (2011). Genome sequence of an Australian kangaroo, Macropus eugenii, provides insight into the evolution of mammalian reproduction and development. *Genome Biology* **12**, R81.

Roberts, R. G., T. F. Flannery, L. K. Ayliffe, H. Yoshida, J. M. Olley, G. J. Prideaux, G. M. Laslett, A. Baynes, M. A. Smith & R. Jones (2001). New ages for the last Australian megafauna: continent-wide extinction about 46,000 years ago. *Science* **292**, 1888-1892.

Saltre, F., M. Rodriguez-Rey, B. W. Brook, C. N. Johnson, C. S. M. Turney, J. Alroy, A. Cooper, N. Beeton, M. I. Bird, D. A. Fordham, R. Gillespie, S. Herrando-Perez, Z. Jacobs, G. H. Miller, D. Nogues-Bravo, G. J. Prideaux, R. G. Roberts & C. J. A. Bradshaw (2016). Climate change not to blame for late Quaternary megafauna extinctions in Australia. *Nature Communications* **7**, 1-7.

Shafer, A. B., J. B. Wolf, P. C. Alves, L. Bergstrom, M. W. Bruford, I. Brannstrom, G. Colling, L. Dalen, L. De Meester & R. Ekblom (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution* **30**, 78-87.

Shorthouse, D. J., D. Iglesias, S. Jeffress, S. Lane, P. Mills, G. Woodbridge, S. McIntyre & A. D. Manning (2012). The 'making of' the Mulligans Flat - Goorooyarroo experimental restoration project. *Ecological Managment and Restoration* **13**, 112-125.

Wells, R. T., K. Moriarty & D. L. G. Williams (1984). The fossil vertebrate deposits of Victoria Fossil Cave, Naracoorte: an introduction to the geology and fuana. *Australian Zoologist* **21**, 305-333.

Woinarski, J. C. Z., S. Legge, J. A. Fitzsimons, B. J. Traill, A. A. Burbidge, A. Fisher, R. S. C. Firth, I. J. Gordon, A. D. Griffiths, C. N. Johnson, N. L. McKenzie, C. Palmer, I. Radford, B. Rankmore, E. G. Ritchie, S. Ward & M.

Ziembicki (2011). The disappearing mammal fauna of northern Australia: context, cause, and response. *Conservation Letters* **4**, 192-201.

Appendix A - Published Manuscripts

The following citations are the published manuscripts arising from this thesis research (1 and 2) with the journal versions (PDF) included on the ensuing pages. Journal publications in which I was a co-author (3 – 6) can be found in Appendix: B, C, D and E.

1. **Haouchar, D**., Haile, J., Spencer, P.B.S., Bunce, M., 2012. The identity of the Depuch Island rock-wallaby revealed through ancient DNA. *Australian Mammalogy* **35**, 101-106.

2. **Haouchar, D**., Haile, J., Mcdowell, M.C., Murray, D.C., White, N.E., Allcock, R.J.N., Phillips, M.J., Prideaux, G.J., Bunce, M., 2013. Thorough assessment of DNA preservation from fossil bone and sediments excavated from a Late Pleistocene-Holocene cave deposit on Kangaroo Island, south australia. *Quaternary Science Reviews* **84**, 56-64.

3. [See Appendix B]
   Phillips, M.J., **Haouchar D**., Pratt, R.C., Gibb, G.C., Bunce, M., 2013. Inferring kangaroo phylogeny from incongruent nuclear and mitochondrial genes. *PloS ONE* **8**, e57745.

4. [See Appendix C]
   McDowell, M.C., **Haouchar, D**., Aplin, K.P., Bunce, M., Baynes, A., Prideaux, G.J., 2015. Morphological and molecular evidence supports specific recognition of the recently extinct *Bettongia anhydra* (Marsupialia: Macropodidae). *Journal of Mammalogy* **96**, 287-296.

5. [See Appendix D]
   Pacioni, C., Hunt, H., Allentoft, M.E., Vaughan, T.G., Wayne, A.F., Baynes, A., **Haouchar, D**., Dortch, J., Bunce, M., 2015. Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. *Molecular Ecology* **24**, 5813-5828.

6. [See Appendix E]

Murray, D.C., Haile, J., Dortch, J., White, N., **Haouchar, D**., Bellgard, M.I., Allcock, R.J., Prideaux, G.J., Bunce, M., 2013. Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports* **3**, 3371.

CSIRO PUBLISHING

Australian Mammalogy
http://dx.doi.org/10.1071/AM11044

The identity of the Depuch Island rock-wallaby revealed through ancient DNA

D. Haouchar[A], *J. Haile*[A], *P. B. S. Spencer*[B] *and M. Bunce*[A,C]

[A]Ancient DNA Laboratory, School of Biological Sciences, Murdoch University, Perth, WA 6150, Australia.
[B]Wildlife Identification Laboratory, School of Biological Sciences, Murdoch University, Perth, WA 6150, Australia.
[C]Corresponding author. Email: m.bunce@murdoch.edu.au

Abstract. Ancient DNA is becoming increasingly recognised as a tool in conservation biology to audit past biodiversity. The widespread loss of Australian biodiversity, especially endemic mammal populations, is of critical concern. An extreme example occurred on Depuch Island, situated off the north-west coast of Western Australia, where an unidentified species of rock-wallaby (*Petrogale* sp.) became extinct as a result of predation by red foxes. Two potential candidate species, *Petrogale lateralis* and *P. rothschildi*, both have ranges adjacent to Depuch Island, making identification based on geography difficult. A museum bone (one of the only surviving Depuch Island specimens) was subjected to standard ancient DNA analyses and procedures. Mitochondrial DNA cytochrome *b* and hypervariable control region were targeted for species identification. Ancient DNA was successfully recovered from the bone: 200 base pairs (bp) of control region and 975 bp of the cytochrome *b* gene. Bayesian phylogenetic analyses were employed to model the Depuch Island rock-wallaby DNA sequences together with sequences of other rock-wallaby taxa from GenBank. Evidence suggests that of the two *Petrogale lateralis* subspecies proposed to have inhabited Depuch Island, *Petrogale lateralis lateralis* was identified as the most likely. The identification of the Depuch Island rock-wallaby population may assist in the reintroduction of an insurance population of *Petrogale lateralis lateralis*, which is becoming increasingly threatened on mainland Australia.

Additional keywords: Past biodiversity, Conservation.

Received 21 November 2011, accepted 25 July 2012, published online 26 October 2012

Introduction

The Australian mammal fauna has undergone several extinctions and population declines since European settlement (Burbidge and Manly 2002; Johnson 2006). Among those species most severely affected have been the medium-sized species, more specifically marsupials that are in the 'critical weight range' (between 35 and 5500 g) (Johnson and Isaac 2009). As a result, combinations of environmental pressures such as introduced feral predators, and demographic and genetic stochasticity have contributed to putting isolated populations of critical-weight marsupials, such as those that exist on islands (Frankham 1997; Eldridge *et al*. 1999), under extreme pressure.

Located off the north-west coast of Western Australia (Fig. 1*a*), Depuch Island (DI) was previously separated from the mainland by ~3 km (2 miles) of shallow water and mud flats (Ride 1964). Whilst neighbouring islands in the Forestier's Archipelago are low and sandy, DI has undulating outcrops of dolerite and granite (Ride 1964; McCarthy 1961), which provide ideal refuge sites and habitat for rock-wallabies. DI, known as Womalantha by Aboriginal people, was named in 1801 by Nicolas Baudin on board the *Le Géographe* and revisited 40 years later by Captain J. C. Wickham, commander of the *H.M.S. Beagle*, who conducted the first biological survey of the island (Ride 1964). DI is well known to anthropologists as an important site of Aboriginal rock art and engravings (Ride 1964) and was declared a sanctuary in 1958 for the protection of engravings and fauna (McCarthy 1961). The island was previously inhabited by several bird, reptile and mammal species including a rock-wallaby (recorded as a small kangaroo by Péron 1807, and as *Petrogale lateralis* by Stokes 1846), which became extinct as a consequence of predation by the introduced red fox (*Vulpes vulpes*) (Kinnear *et al*. 1988, 1998). Rock-wallaby colonies were reported on DI at the mouth of the Balla Balla River and the animals were known to be present in 1964. Judging by the vast quantities of their droppings around the island, the rock-wallabies seemed plentiful at the time (Ride 1964). However, 20 years later they were considered locally extinct (Hall and Kinnear 1991) before any secure species identification could be made. Prior to any reintroduction attempts, a sound species and genetic characterisation of the extinct wallaby should be undertaken.

At present, the taxonomic identity of the Depuch Island rock-wallaby is unresolved, with two candidate species inhabiting north-west Australia: the black-footed (black-flanked) rock-wallaby (*Petrogale lateralis*) and Rothschild's rock-wallaby (*Petrogale rothschildi*). The location of DI within the range of both species makes it difficult to predict which taxon might have formerly occupied this island (see Fig. 1*a*). Ride (1964) set out to confirm Stokes' (Stokes 1846) observation of an unknown

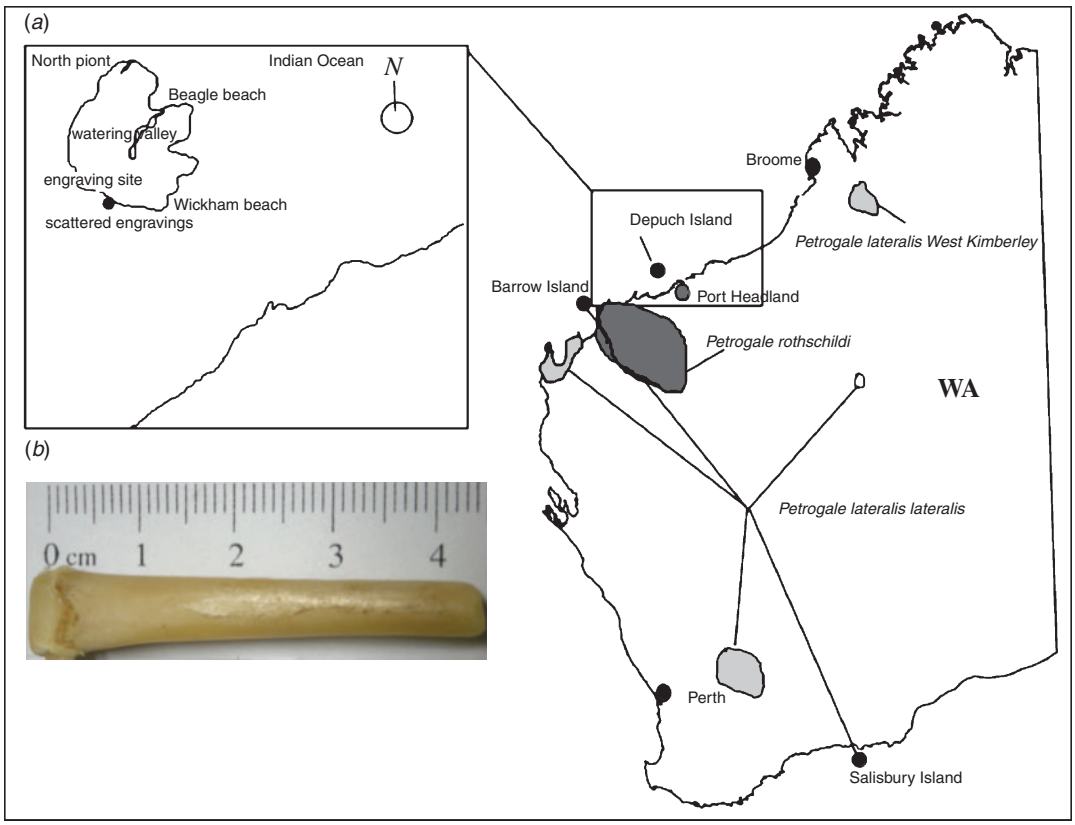


Fig. 1. (*a*) Current species distribution of *Petrogale rothschildi* (dark grey), *Petrogale lateralis lateralis* (light grey) and *Petrogale lateralis* West Kimberley race (medium grey) in relation to Depuch Island. Inset shows an enlargement of Depuch Island, showing sites mentioned in the text. (*b*) Western Australian Museum fossil bone of fourth metatarsal (proximal end missing) (M5233; Depuch No. 48) from Depuch Island (collected on 15 June 1962) used in this study. Map (*a*) based on Mason *et al*. (2011) and (Eldridge *et al*. 1994).

species of rock-wallaby on DI, and stressed the importance of identifying the species for zoogeographical and taxonomic reasons. Ride (1964) noted that two species occurred in the vicinity of the island, *P. rothschildi* inhabiting islands of the Dampier Archipelago and *P. lateralis* being dispersed throughout the Western Australian mainland and also Barrow Island (Ride 1964).

However, *Petrogale lateralis* is polytypic, consisting of five distinct subspecies/races – *P. l. lateralis*, *P. l. hacketti*, *P. l. pearsoni*, *P. l.* 'MacDonnell Ranges' and *P. l.* 'West Kimberley' (Sharman *et al*. 1990; Eldridge *et al*. 1994, 2001; Eldridge and Close 1997; Pearson and Kinnear 1997; Potter *et al*. 2012) – some of which are readily distinguished by traditional morphology (e.g. skull/body morphology/size, coat colour and markings). However, such features are highly variable within rock-wallaby species and are therefore an unreliable means of classification (Eldridge and Close 1992, 1997; Eldridge and Spencer 1997). Prior to DNA studies one of the most convincing genetic species determiners had been chromosome and allozyme similarities (Sharman *et al*. 1989; Eldridge *et al*. 1991*a*, 1991*b*, 1994). However, the use of chromosomes to provide definitive species identification is limited by the technique's requirement for living tissue from which rapidly dividing cells can be cultured for analyses (Eldridge and Spencer 1997; Loupis and Eldridge 2001). Now with DNA studies becoming a diagnostic tool for species identification, it is mandatory that this research be used in identifying which subspecies/races of *P. lateralis* likely occurred on DI. Given the location of *P. l. lateralis* and *P. l.* 'West Kimberley' near the island, either of these could be potential candidates. Therefore, the use of mitochondrial DNA (mtDNA) targeting cytochrome *b* and control region can be useful in telling these taxa apart.

Previously, chromosome analyses categorised *Petrogale* species into three distinct groups: the *xanthopus*, *brachyotis* and *lateralis–penicillata* groups (Eldridge and Close 1992, 1997). However, recent molecular analyses shows that the *xanthopus* group is not monophyletic and comprises independent lineages for *P. xanthopus*, *P. persephone* and *P. rothschildi*, despite their shared ancestral karyotypes (Potter *et al*. 2012). Further analyses show that four distinct lineages were identified on the basis of mitochondrial and nuclear DNA: (1) the *brachyotis* group, (2) *Petrogale persephone*, (3) *Petrogale xanthopus* and (4) the *lateralis–penicillata* group (Potter *et al*. 2012).

The aim of this study was to isolate ancient DNA from a museum specimen for species identification, namely targeting mtDNA. A marginally degraded subfossil museum bone (Fig. 1*b*) was subjected to mitochondrial DNA (mtDNA) analysis of the cytochrome *b* gene and hypervariable control region. Until now

no attempts had been made to clearly identify the endemic species of rock-wallaby that previously inhabited the island, which is important as it will allow informed decisions to be made in any biodiversity restoration projects onto the now fox-free DI. Since this species is completely extinct on the island and there are only a few trace rock-wallaby museum specimens left, ancient DNA analysis is the most suitable technique and is highly recommended in order to handle and retrieve relevant data for species identification useful in this study.

Materials and methods

Sample collection and DNA extraction

Due to sampling restrictions, a single DI rock-wallaby subfossil bone (fourth metatarsal) was accessed from the Western Australian Museum (M5233; collected on 15 June 1962) and sampled for DNA (Fig. 1*b*). DNA extraction procedures were carried out in a dedicated ancient DNA laboratory, minimising contamination from PCR amplicons and modern DNA. A sample of ~100 mg of bone powder was obtained using a Dremel tool (part no. 114; Germany) at high rotational speeds. Bone powder was collected in a 1.5-mL Eppendorf tube, weighed and stored for later digestion.

The digestion step included an overnight incubation at 55°C with rotation in a 1.5-mL digestion buffer containing 20 mM Tris pH 8.0 (Sigma, MO, USA), 10 mM dithiothreitol (Thermo Fisher Scientific, MA, USA), 1 mg mL^{-1} proteinase K powder (Amresco, OH, USA), 0.48 M EDTA (EDTA) (Invitrogen) and 1% Triton X-100 (Invitrogen). After digestion, centrifugation at 13 000*g* was initiated for 1 min to pellet undigested material. The supernatant containing the DNA was concentrated to ~100 µL in a Vivaspin 500 column (MWCO 30000; Sartorius Stedim Biotech, Germany) at 13 000*g*, and then combined with five volumes of PBi buffer (Qiagen, CA, USA). DNA was immobilised on silica spin columns (Qiagen) and washed with 700 µL of AW1 and AW2 wash buffers (Qiagen). Finally, the DNA was eluted from the silica in 50 µL of 10 mM Tris pH 8.0 (Sigma, MO, USA).

Primer design

Primers were designed in order to amplify small segments of DNA spanning the targeted mtDNA regions, ranging in size from 150 to 420 base pairs (bp) (see Table 1 for primer combinations). PCR amplification of 975 bp of cytochrome *b* gene and 200 bp of the hypervariable control region was targeted by designing specific primers for rock-wallabies using GENEIOUS 4.8 (Biomatters, New Zealand).

PCR of cytochrome b and control region

Amplification of DNA product was carried out using real-time PCR to assess the amount of DNA preserved in the sample, a protocol advocated in many ancient DNA procedures (Cooper and Poinar 2000; Pruvost and Geigl 2004). The StepOne real-time PCR system (Applied Biosystems) was used with a final reaction concentration of: 1 × High Fidelity PCR Buffer (Invitrogen), 50 mM $MgSO_4$ (Invitrogen), 0.25 mM of each dNTPs (Austral Scientific), 8 µM of each primer, 0.25 U HIFI Taq polymerase (Invitrogen), 1 µL (10 mg mL^{-1}) bovine serum albumin (Fisher Biotech), 0.6 µL SybrGreen (Invitrogen cat no S7563, 1 : 2000 dilution), ultrapure H_2O and 2 µL of DNA extract in a 25-µL reaction. PCR thermal cycling was initiated with a 5-min denaturation step at 94°C, followed by 50 cycles of 94°C for 45 s, with a variable annealing temperature either 53°C or 57°C according to primer set used (see Table 1), followed by 68°C for 45 s and a final extension at 72°C for 10 min. Quantitative PCR data were analysed using Applied Biosystems StepONE software version 2.00 software. The relative yield of DNA was assessed between each dilution sample according to the C_T values. Samples were visualised on a 2% w/v DNA-grade agarose gel electrophoresis (Bio-Rad) stained with ethidium bromide.

Cloning mtDNA

All DNA products were amplified at least twice, and control region PCR products that contained amplified fragments of the appropriate size were purified using a Qiaquick Purification kit (Qiagen) and cloned using pGEM-T vector system (Promega). Successfully cloned products were chosen after being screened for blue (which do not contain the pGEM-T vector) and white (which contain the pGEM-T vector) colonies on LB/ampicillin/IPTG/X-Gal plates. A random selection of ~10 white colonies and at least one blue colony was selected for size comparison and screened on a 2% agarose w/v gel. Amplicons were checked

Table 1. Mitochondrial control region and Cytochrome b primer sequences and amplification conditions used to genetically characterise the DI rock wallaby

Primer name	Primer sequence 5′ → 3′	Annealing temperature (°C)	Amplicon size (bp)
DL1F	CCACAACACATCAACTYATTTG	53	150
DL1R	ATTCATTTTATGTATTACTAGAATTATGTA		
DL3F	TGTATTAAGACAGATATGTATAAAGT	53	250
DL2R	ATTCATTTTATGTATTACTAGAATTATGTA		
DL3F	TGTATTAAGACAGATATGTATAAAGT	53	280
DL3R	AGTCAGAGATTTGTTAGGTACG		
Cytb_WallabyF1	GACACCCTAACAGCCTTCTCATCAG	57	260
Cytb_WallabyR1	CGGTAGCTCCTCAGAATGATATTT		
Cytb_WallabyF2	AAATATCATTCTGAGGAGCTACCG	57	340
Cytb_WallabyR2	GAGAAGTTGTCTGGGTCGCC		
Cytb_WallabyF3	GGCGACCCAGACAACTTCTC	57	240
Cytb_WallabyR3	GGCTGTAAGGATTCAGAATAGGAT		

to ensure that products were of predicted size, purified and finally prepared for sequencing. This cloning step ensured the production of clean chromatographs facilitating unambiguous base calls, as well as allowing an assessment of DNA damage and polymerase error.

Sequencing of Cytochrome b and control region

Sanger sequencing was carried out at the commercial facility Macrogen (Seoul, South Korea) using BigDye ver. 3.1 (Applied Biosystems) chemistry on a cycling ABI3730XL capillary sequencer (Applied Biosystems). Sequences were analysed using GENEIOUS 5.4.3 (Biomatters, New Zealand) and deposited on GenBank under accession numbers JN898804 and JN898805.

Phylogenetic analysis of Cytochrome b and control region sequence

Alignments of nucleotide sequences were carried out in GENEIOUS 5.4.3, with any ambiguities resolved by eye. Reference species used in this analysis were derived from published sequence data (Potter *et al*. 2012). GenBank accession numbers for cytochrome *b* – *P. lateralis lateralis* (JQ042127-S972, Nangeen Hill, south-west Western Australia), *P. rothschildi* (JQ042134-S204, Rosemary Island, Dampier Archipelago) and *P. lateralis* West Kimberley race (JQ042130) – were used for comparison. Control region of *P. lateralis lateralis* (AF348675-S207, Ningaloo) was also used, along with *P. rothschildi* (not published) and *P. lateralis* West Kimberley race (AF348688) for comparison. To ensure validity of DNA sequences and to overcome ancient DNA damage, multiple-sequence datasets were created and an overall consensus was drawn for use in the final phylogenies. Topology analysis was conducted using the Bayesian phylogenetic program MrBayes (Huelsenbeck and Ronquist 2001), a plug-in application provided through GENEIOUS 5.4.3. The MrBayes analysis extended over 1 million iterations, and genealogies and model parameters were sampled every 1000 iterations with 10% burn-in. An HKY85 substitution model was imposed and an invariance gamma model for among-site variation was chosen according to the simulations run in MrjModel Test (Nylander 2004). Results were visualised and examined in FigTree ver. 1.2.2.

Results

MtDNA analysis of Cytochrome b and control region

The cytochrome *b* sequence alignment of 975 bp of the DI rock-wallaby specimen to the two reference species, *P. l. lateralis* and *P. rothschildi*, consisted of 99% and 93% of identical sites, respectively. Among-site variation between the DI rock-wallaby and *P. l. lateralis* revealed 10 single-nucleotide polymorphisms (SNPs), nine transitions (7 C → T and 2 G → A) and one transversion (A → C). Between the DI rock-wallaby and *P. rothschildi* 69 SNPs were identified, 58 being transitions (48 C → T and 10 G → A) and 11 transversions (7 A → C and 4 A → T).

The control region (Domain 1) alignment of 200 bp of the DI rock-wallaby specimen and *P. l. lateralis* and *P. rothschildi* consisted of 92.3% and 63.2% identical sites, respectively. Among-site variation revealed 15 SNPs between the DI rock-wallaby and *P. l. lateralis*, eight transitions (2 G → A and 6 C → T) and seven transversions (2 A → C, 4 A → T and 1 G → C). Among-site variation between the DI rock-wallaby and *P. rothschildi* consisted of 41 SNPs, 16 transitions (13 C → T and 3 G → A) and 25 transversions (3 A → C, 13 A → T, 3 G → C and 6 T → G).

Phylogenetic analysis

Mitochondrial DNA sequences were positively identified on the basis of congruent phylogenies for both the cytochrome *b* gene and control region. This is a strong indication that sequences were not artefacts of nuclear copies. Modelling the concatenated consensus sequence data for both cytochrome *b* and control region was carried out using Bayesian phylogenetic methods. The *Thylogale* (pademelon) was chosen as the outgroup for this study. Phylogenetic analysis shows clear support that the unidentified rock-wallaby specimen collected from DI is most similar to *P. l. lateralis* at the mtDNA level (Fig. 2). *P. l. lateralis* consistently forms a sister clade with the DI specimen (100% posterior probabilities) in both phylogenies to the exclusion of *P. rothschildi* and other closely neighbouring rock wallabies such as *P. lateralis* West Kimberley (Fig. 2).

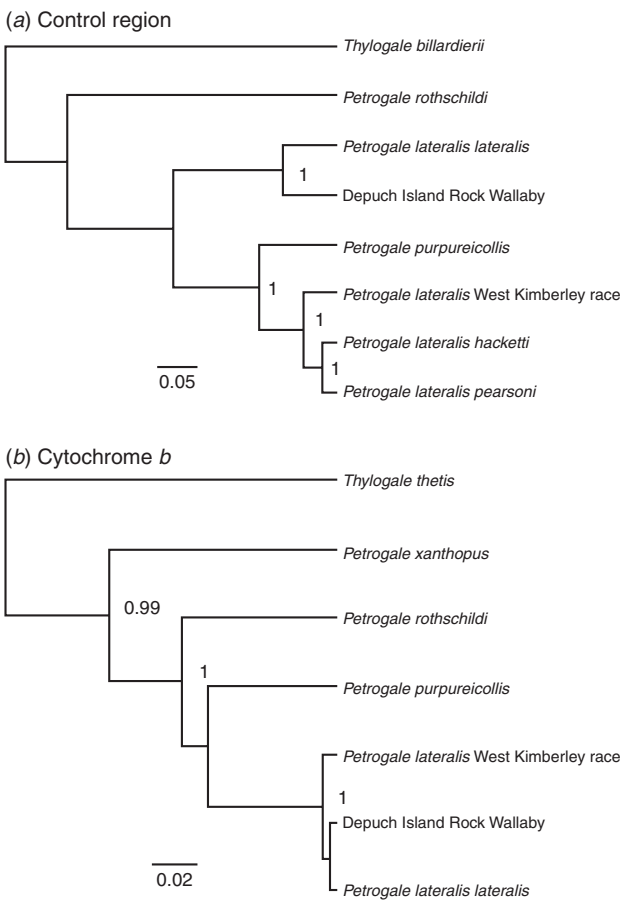


Fig. 2. Bayesian phylogenetic tree showing the relationship of the Depuch Island rock-wallaby sequence with other available rock-wallaby sequence data: (*a*) based on 200 bp of control region and (*b*) based on 975 bp of cytochrome *b* sequence. Posterior probabilities greater than 90 are shown on nodes. The tree was built using a HKY85 model and invariant gamma was assumed and imposed with a relaxed molecular clock.

Discussion

Mainland Australia has recorded the world's highest rate of recent mammal extinctions (Short and Smith 1994) and mainland *Petrogale* populations have not been immune to these changes. Having undergone severe population retractions, the remaining animals are highly vulnerable to introduced predators (Kinnear *et al*. 1988, 1998), inbreeding depression and a multitude of interacting anthropogenic factors that further threaten population viability. Offshore islands have provided a refuge for many species, and currently harbor 14 species of Australian mammals (Eldridge *et al*. 1999) that are now extinct on the mainland. Such isolated islands can offer valuable insights into speciation and adaptive radiation and provide a refuge and conservation site for terrestrial mammal species (Maxwell *et al*. 1996; Abbott 2000). To date, species such as *Bettongia lesueur*, *Lagorchestus fasciatus*, *Perameles bougainville*, *Pseudomys fieldi*, amongst others, would now be extinct had they not been conserved on Western Australian islands (Burbidge *et al*. 1997).

DI is currently isolated, with the land bridge connecting it to the mainland being permanently under water, creating a barrier to mammal dispersal to the island. Prior to this separation it was connected by ~3 km of mud flats, exposed at low tide, which enabled access to the island and therefore gene flow and migration of several species, including introduced predators such as the red fox. Rock-wallaby (*Petrogale* spp.) populations were noted to exist on the island at one time and on other offshore islands of Western Australia, such as Salisbury and Barrow Island (Hall and Kinnear 1991), which have remained fox free, mainly due to their isolation (Eldridge *et al*. 2001). The species of rock-wallaby on DI, however, was never identified before it became locally extinct.

Our ancient DNA study has established that *P. l. lateralis* is most closely related to the DI rock-wallaby population and therefore the most probable rock-wallaby taxon that once inhabited DI. Having determined this makes future reintroduction attempts possible, with the exclusion of other potential rock-wallaby taxa like the *P. lateralis* 'West Kimberley' and *P. rothschildi* as potential candidates.

Previous attempts to identify unknown taxa using traditional morphological criteria have sometimes been challenging when dealing with morphologically cryptic species such as *Petrogale* (Loupis and Eldridge 2001). Other researchers have shown that although karyotypes can be useful in differentiating rock-wallaby taxa (Eldridge *et al*. 1991*a*), a major limitation is the necessity of living tissue that can then be cultured (Eldridge and Spencer 1997). The advancement of molecular biology techniques and ancient DNA techniques means that even samples that have been subjected to severe degradation can be used as a valuable source of information to investigate past populations.

Petrogale l. lateralis is classified as Vulnerable by the *Environment Protection and Biodiversity Conservation Act 1999* and the *Western Australian Wildlife Conservation Act*. To ensure survival of this species in the face of the continuing decline of mainland populations (Hall and Kinnear 1991; Mason *et al*. 2011), we would advocate the relocation of a viable population of *P. l. lateralis* onto the now fox-free DI, which would provide an insurance population for the species, at least for the short term. Recent studies by Mason *et al*. (2011) and others (Frankham 1997; Frankham *et al*. 2002) highlight the importance of preserving fragmented mainland populations of rock-wallabies rather than translocating populations to islands, because of the increasingly low levels of genetic diversity. However, given that *P. l. lateralis* was the original inhabitant of DI, repopulating it will be a significant step in restoring the island's past biodiversity and one that is likely to have positive effects on the island's ecosystem. Additionally, a sound understanding of the genetic diversity of mainland populations is a necessary prerequisite to enhance the reintroduced population's survival in the new island environment (Mason *et al*. 2011). The ultimate aim must be that, following successful reintroductions, such island sanctuaries can in the future be used as sources for returning species back onto the mainland, and, in doing so, restore at least some of the lost biodiversity of the recent past.

Acknowledgements

Thanks to Mark Eldridge for providing several critical reference sequences for this study to be carried out and to the Western Australian Museum (Dr Ric How) for providing the subfossil samples used in this study. This research was funded by the Australian Research Council grant 'An inventory of past biodiversity in Western Australia using ancient DNA' (FT0991741) awarded to MB.

References

Abbott, I. (2000). Improving the conservation of threatened and rare mammal species through translocation to islands: case study Western Australia. *Biological Conservation* **93**, 195–201. doi:10.1016/S0006-3207(99)00144-5

Burbidge, A. A., and Manly, B. F. J. (2002). Mammal extinctions on Australian islands: causes and conservation implications. *Journal of Biogeography* **29**, 465–473. doi:10.1046/j.1365-2699.2002.00699.x

Burbidge, A. A., Williams, M. R., and Abbott, I. (1997). Mammals of Australian islands: factors influencing species richness. *Journal of Biogeography* **24**, 703–715. doi:10.1046/j.1365-2699.1997.00145.x

Cooper, A., and Poinar, H. N. (2000). Ancient DNA: do it right or not at all. *Science* **289**, 1139.

Eldridge, M. D. B., and Close, R. L. (1992). Taxonomy of rock wallabies, *Petrogale* (Marsupialia, Macropodidae). I. A revision of the eastern *Petrogale* with the description of three new species. *Australian Journal of Zoology* **40**, 605–625. doi:10.1071/ZO9920605

Eldridge, M. D. B., and Close, R. L. (1997). Chromosomes and evolution in rock-wallabies, *Petrogale* (Marsupialia: Macropodidae). *Australian Mammalogy* **19**, 123–135.

Eldridge, M. D. B., and Spencer, P. B. S. (1997). Sources and uses of genetic material in the study of *Petrogale* (rock wallabies) and other mammals. *Australian Mammalogy* **19**, 265–278.

Eldridge, M. D. B., Close, R. L., and Johnston, P. G. (1991*a*). Chromosomal rearrangements in rock-wallabies, *Petrogale* (Marsupialia, Macropodidae). IV. G-Banding analysis of the *Petrogale lateralis* complex. *Australian Journal of Zoology* **39**, 621–627. doi:10.1071/ZO9910621

Eldridge, M. D. B., Johnston, P. G., and Close, R. L. (1991*b*). Chromosomal rearrangements in rock-wallabies, *Petrogale* (Marsupialia, Macropodidae). V. Chromosomal phylogeny of the *lateralis–penicillata* group. *Australian Journal of Zoology* **39**, 629–641. doi:10.1071/ZO9910629

Eldridge, M. D. B., Kinnear, J. E., and Close, R. L. (1994). Identification of rock wallabies in the Calvert ranges, Little Sandy Desert, Western Australia, as *Petrogale lateralis lateralis*. *Australian Mammalogy* **17**, 129–131.

Eldridge, M. D. B., King, J. M., Loupis, A. K., Spencer, P. B. S., Taylor, A. C., Pope, L. C., and Hall, G. P. (1999). Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby. *Conservation Biology* **13**, 531–541. doi:10.1046/j.1523-1739.1999.98115.x

Eldridge, M. D. B., Wilson, A. C. C., Metcalfe, C. J., Dollin, A. E., Bell, J. N., Johnson, P. M., Johnston, P. G., and Close, R. L. (2001). Taxonomy of rock-wallabies, *Petrogale* (Marsupialia: Macropodidae). III. Molecular data confirms the species status of the purple-necked rock-wallaby (*Petrogale purpureicollis* Le Souef). *Australian Journal of Zoology* **49**, 323–344. doi:10.1071/ZO00082

Frankham, R. (1997). Do island poulations have less genetic variation than mainland populations? *Heredity* **78**, 311–327. doi:10.1038/hdy.1997.46

Frankham, R., Ballou, J., and Briscoe, D. A. (2002). 'Introduction to Conservation Genetics.' (Cambridge University Press: England.)

Hall, G. P., and Kinnear, J. E. (1991). Recovery plan for the black-flanked rock-wallaby *Petrogale lateralis lateralis* (Gould) (Australian National Parks and Wildlife Service Endangered Species Program, Project 149). Department of Conservation and Land Management, Western Australia.

Huelsenbeck, J. P., and Ronquist, F. (2001). Mr Bayes: Baysian inference of phylogenetic trees. *Bioinformatics (Oxford, England)* **17**, 754–755. doi:10.1093/bioinformatics/17.8.754

Johnson, C. (2006). 'Australia's Mammal Extinctions: a 50 000 Year History.' (Cambridge University Press: America.)

Johnson, C. N., and Isaac, J. L. (2009). Body mass and extinction risk in Australian marsupials: the 'Critical Weight Range' revisited. *Austral Ecology* **34**, 35–40. doi:10.1111/j.1442-9993.2008.01878.x

Kinnear, J. E., Onus, M. L., and Bromilow, R. N. (1988). Fox control and rock-wallaby population dynamics. *Australian Wildlife Research* **15**, 435–450. doi:10.1071/WR9880435

Kinnear, J. E., Onus, M. L., and Sumner, N. R. (1998). Fox control and rock-wallaby population dynamics – II. An update. *Wildlife Research* **25**, 81–88. doi:10.1071/WR96072

Loupis, A. K., and Eldridge, M. D. B. (2001). Restriction fragment length polymorphism (RFLP) analysis of three nuclear genes in rock-wallabies (*Petrogale*: Marsupialia: Macropodidae): a search for genic markers to identify taxa within the *Petrogale lateralis–penicillata* group. *Australian Journal of Zoology* **49**, 27–36. doi:10.1071/ZO00078

Mason, R., Browning, T., and Eldridge, M. (2011). Reduced MHC class II diversity in island compared to mainland populations of the black-footed rock-wallaby (*Petrogale lateralis lateralis*). *Conservation Genetics* **12**, 91–103. doi:10.1007/s10592-009-9993-y

Maxwell, S., Burbidge, A. A., and Morris, K. (1996). 'Action plan for the Australian Marsupials and Monotremes' for the Australasian Marsupial and Monotreme Specialist Group, IUCN Species Survival Commission. Wildlife Australia, Canberra.

McCarthy, F. D. (1961). The rock-engravings of Depuch Island, north-west Australia. *Records of the Australian Museum* **25**, 121–148. doi:10.3853/j.0067-1975.25.1961.660

Nylander, J. A. A. (2004). MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala Univeristy.

Pearson, D. J., and Kinnear, J. E. (1997). A review of the distribution status and conservation of rock-wallabies in Western Australia. *Australian Mammalogy* **19**, 137–152.

Péron, F. (1807). 'A Voyage of Discovery to the Southern Hemisphere Performed by the Order of the Emperor Napolean during the Years 1801, 1802,1803, and 1804. (London.) pp. 129–132.

Potter, S., Cooper, S. J. B., Metcalfe, C. J., Taggart, D. A., and Eldridge, M. D. B. (2012). Phylogenetic relationships of rock-wallabies, *Petrogale* (Marsupialia: Macropodidae) and their biogeographic history within Australia. *Molecular Phylogenetics and Evolution* **62**, 640–652. doi:10.1016/j.ympev.2011.11.005

Pruvost, M., and Geigl, E. M. (2004). Real-time quantitative PCR to assess the authenticity of ancient DNA amplification. *Journal of Archaeological Science* **31**, 1191–1197.

Ride, W. D. L. (1964). Depuch Island. *The Western Australian Museum* **2**, 1–89.

Sharman, G., Close, R., and Maynes, G. (1989). Chromosome evolution, phylogeny and speciation of rock wallabies (*Petrogale*, Macropodidae). *Australian Journal of Zoology* **37**, 351–363. doi:10.1071/ZO9890351

Short, J., and Smith, A. (1994). Mammal decline and recovery in Australia. *Journal of Mammalogy* **75**, 288–297. doi:10.2307/1382547

Stokes, J. L. (1846). 'Discoveries in Australia during the voyage of *H.M.S. Beagle* in the years 1837–38–39–40–41–42–43.' (T. and W. Boone: London.)

Quaternary Science Reviews 84 (2014) 56–64

Contents lists available at ScienceDirect

Quaternary Science Reviews

journal homepage: www.elsevier.com/locate/quascirev

Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late Pleistocene–Holocene cave deposit on Kangaroo Island, South Australia

Dalal Haouchar [a], James Haile [a], Matthew C. McDowell [b], Dáithí C. Murray [a], Nicole E. White [a], Richard J.N. Allcock [c], Matthew J. Phillips [d], Gavin J. Prideaux [b], Michael Bunce [a,e,*]

[a] Australian Wildlife Forensic Services and Ancient DNA Laboratory, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia
[b] School of Biological Sciences, Flinders University, Bedford Park, SA 5042, Australia
[c] LotteryWest State Biomedical Facility: Genomics, School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, WA 6009, Australia
[d] School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia
[e] Trace and Environmental DNA Laboratory, Department of Environment and Agriculture, Curtin University, Perth, WA 6845, Australia

ARTICLE INFO

Article history:
Received 9 August 2013
Received in revised form
6 November 2013
Accepted 10 November 2013
Available online

Keywords:
Pleistocene–Holocene
Quaternary
Fossils
Ancient DNA
Biodiversity

ABSTRACT

Fossils and sediments preserved in caves are an excellent source of information for investigating impacts of past environmental changes on biodiversity. Until recently studies have relied on morphology-based palaeontological approaches, but recent advances in molecular analytical methods offer excellent potential for extracting a greater array of biological information from these sites. This study presents a thorough assessment of DNA preservation from late Pleistocene–Holocene vertebrate fossils and sediments from Kelly Hill Cave Kangaroo Island, South Australia. Using a combination of extraction techniques and sequencing technologies, ancient DNA was characterised from over 70 bones and 20 sediment samples from 15 stratigraphic layers ranging in age from >20 ka to ~6.8 ka. A combination of primers targeting marsupial and placental mammals, reptiles and two universal plant primers were used to reveal genetic biodiversity for comparison with the mainland and with the morphological fossil record for Kelly Hill Cave. We demonstrate that Kelly Hill Cave has excellent long-term DNA preservation, back to at least 20 ka. This contrasts with the majority of Australian cave sites thus far explored for ancient DNA preservation, and highlights the great promise Kangaroo Island caves hold for yielding the hitherto-elusive DNA of extinct Australian Pleistocene species.

1. Introduction

Islands have long provided a natural laboratory for the study of evolutionary processes because evolutionary changes on them are often magnified, simplified and therefore more readily interpretable (e.g., Darwin and Wallace, 1858; MacArthur and Wilson, 1967; Losos and Ricklefs, 2010). The study of genetic variation on islands also has a long history (Lomolino et al., 1989; Van der Geer et al., 2010). However, ancient DNA (aDNA) analyses applied to stratified, dated faunal successions can add a temporal context, allowing the ebb and flow of genes, species and communities to be assessed, particularly in combination with more traditional analyses of vertebrate and plant macrofossils and pollen. A necessary prerequisite for aDNA research is adequate biomolecular preservation. Cave systems represent an ideal environment for palaeontological investigations as they often contain relatively complete and undisturbed stratigraphic deposits that harbour several environmental proxies (White, 2007; Butzer, 2008) that have been subjected to minimal temperature and humidity fluctuations; conditions that favour DNA persistence (Stone, 2000). Such caves represent archives of well-preserved Quaternary vertebrate assemblages (Prideaux et al., 2007, 2010), with the ability to preserve invaluable repositories of past biodiversity. All samples (bones and sediments) analysed in this study were obtained directly from Kelly Hill Cave (KHC), Kangaroo Island (KI) with the aim of conducting a

* Corresponding author. Trace and Environmental DNA Laboratory, Department of Environment and Agriculture, Curtin University, Perth, WA 6845, Australia.
E-mail address: michael.bunce@curtin.edu.au (M. Bunce).

0277-3791/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.quascirev.2013.11.007

thorough assessment of DNA preservation in KHC to determine whether genetic data could enhance temporal information about faunal change on KI. Moreover, as part of this study the preservation of plant DNA obtained directly from sediments was assessed (Haile et al., 2007; Willerslev et al., 2007) with the aim to provide palaeovegetation data to complement fossil data.

In this study a combination of techniques such as a relatively new bulk-bone method (Murray et al., 2013) and High-Throughput Sequencing (HTS) technology was used in order to capture aDNA from a variety of samples collected from KHC. Also we show how the addition of a palaeontological molecular perspective may complement existing morphology based studies allowing identification of osteologically absent and cryptic species, and the investigation of genetic change over time. These results are overlayed upon the palaeogeographic history of KI, which provides a model context for studying mainland–island interactions. This is pertinent to KHC, because it contains an excellent Late Quaternary vertebrate fossil assemblage (Pledge, 1979) that records the response of Australian native fauna to both the Last Glacial Maximum (LGM) and isolation of KI caused by rising sea levels at 8.9 ka (McDowell, 2013).

2. Kangaroo Island

Kangaroo Island lies at the entrance to Gulf St Vincent in South Australia (Fig. 1), and is the third-largest land-bridge island in Australia (4405 km^2), with a length of 145 km and width of 55 km at its widest point (Abbott, 1974; Hope et al., 1977; Lampert, 1981; Twidale and Bourne, 2002) (Fig. 1). It is geologically continuous with the adjacent Fleurieu Peninsula, but was isolated by glacial erosion during the Late Carboniferous and Early Permian (Belperio and Flint, 1999). Today it is separated from Yorke Peninsula by Investigator Strait, a 50-km stretch of 30–35 m deep water (Fig. 1).

During the late Pleistocene, sea levels were 120 m lower than at present (Chappell and Shackleton, 1986; Yokoyama et al., 2001) and KI was connected to the mainland. Global warming during the early Holocene caused a rapid rise in sea-level, isolating KI from the mainland at 8.9 ka (Belperio and Flint, 1999; Bradley, 1999; Cutler et al., 2003). Prior to its isolation, gene flow was presumably continuous between the mainland KI which supported a species-rich fauna (Abbott, 1974; McDowell, 2013). Once isolated, the newly marooned organisms would have become more susceptible to genetic drift; island floras and faunas tend to be diverse after initial isolation but subsequently suffer elevated selection pressures, loss of genetic diversity and elevated rates of extinction (Diamond, 1972; Foufopoulos and Ives, 1999; Stiller et al., 2010).

Despite the loss of genetic diversity caused by island living, many species that have become extinct or endangered on the mainland find refuge on islands due to relaxed competition and reduced predation pressures (Lister, 2004). KI retains the largest proportion of uncleared native vegetation of any southern Australian agricultural district. In addition KI remains free of rabbits and foxes that have had such a catastrophic impact on mainland biota (Robinson and Armstrong, 1999) enhancing its conservation importance. Biodiversity management can be complimented by aDNA analyses and assessments of the fossil record, providing parameters such as population sizes, levels of gene flow and population relatedness (Ramakrishnan et al., 2005; Leonard, 2008; De Bruyn et al., 2011).

2.1. Study site

The KHC complex on KI (35.83° S, 137.33° E) is ideally suited to explore biomolecule preservation as it has an already well-studied and well-dated palaeontological record that spans the terminal late Pleistocene to the middle Holocene (McDowell, 2013). In addition numerous surveys of the island's modern flora and fauna have also been made (Robinson and Armstrong, 1999).

KHC is the focus of an ongoing palaeontological project that investigates how climate change and isolation due to sea level rise has affected the fauna of KI (McDowell et al., 2013). To date, a fauna rich in mammals, birds, reptiles, frogs and land-snails has emerged

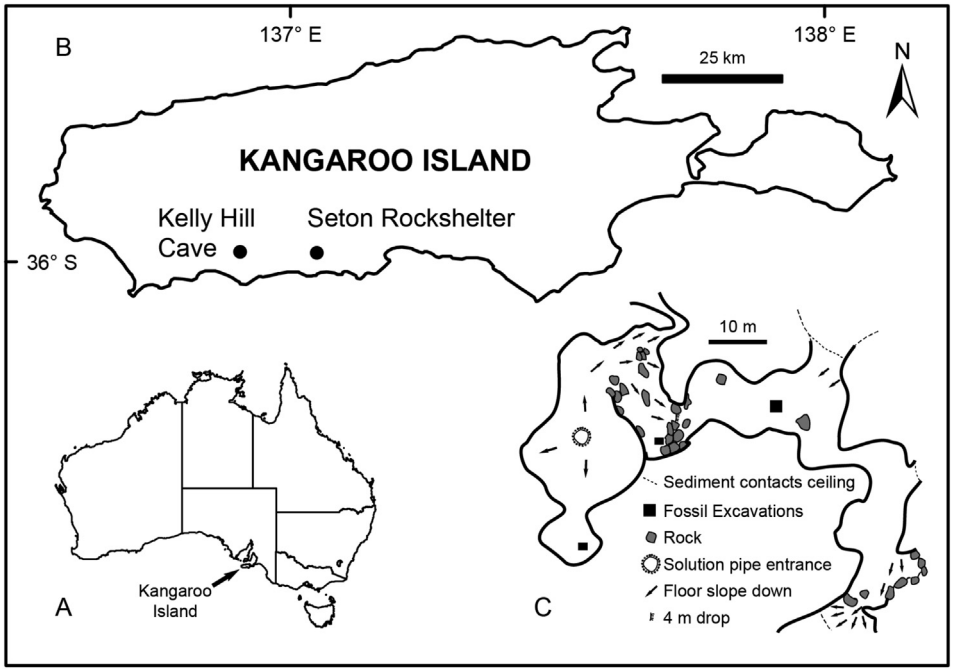


Fig. 1. A. Location of Kangaroo Island relative to the Australian mainland. B. Location of Kelly Hill Cave and Seton Rock Shelter, Kangaroo Island. C. Relevant map section of Kelly Hill Cave showing the location of modern solution pipe entrances, fossil excavation and the blocked palaeo-entrance through which excavated sediments and bones entered the cave (McDowell, 2013).

from an excavation 4 m^2 × 1.5 m deep, and includes the remains of several species not previously recorded on KI. This site was selected based on the presence of fossil bones on the cave floor surface, depth of sediment, likelihood of stratigraphic integrity and likelihood of encountering speleothems that can be U/Th dated (McDowell, 2013; McDowell et al., 2013).

3. Materials & methods

We used 2nd generation DNA sequencing technologies (Roche GS-Junior and Ion Torrent, Personal Genome Platforms (PGM)) to target chloroplast (plastid) DNA (cpDNA) and mitochondrial DNA (mtDNA) from sediment samples and assessed aDNA preservation of bones collected from several sedimentary layers ranging in age from >20 to 6 ka (McDowell et al., 2013). This study uses a combination of techniques to recover aDNA from multiple samples including a novel bone-grind technique, hitherto applied only to two cave sites in Western Australia (Murray et al., 2013) To test the veracity or to refine morphology-based identifications, aDNA was extracted from 70 complete macropodid and reptile postcranial bones. Some specimens had been identified morphologically to species level, while others were only identified to family level (McDowell, 2013).

The layers span the terminal Late Pleistocene to the middle Holocene (>20 ka–6 ka). Bones from layers 3–11 contained adequate collagen to be AMS radiocarbon dated (McDowell et al., 2013). Skeletal remains from layers 12–15 lacked viable collagen and what little remained was too degraded to be radiocarbon dated (McDowell et al., 2013). Standard aDNA protocols specifies that pre-PCR procedures are conducted in a dedicated aDNA clean room, with subsequent post-PCR work carried out in a separate laboratory in order to minimise contamination (Pääbo, 1989; Cooper and Poinar, 2000; Fulton, 2012). DNA extractions and amplifications were conducted at Murdoch University, whilst Sanger sequencing was performed at a commercial facility (South Korea) and HTS was carried out at Murdoch University (Roche, 454 GS-Junior) and the Lotterywest State Biomedical Facility Genomics Node at Royal Perth Hospital (Ion Torrent, PGM). A more detailed version of the molecular methods (Sections 3.2–3.4) can be found in the Supplementary data.

3.1. Background to sediments and sample collection

Numerous sediment core samples were collected from all exposed layers of the KHC excavation using sterile equipment and protective clothing to preserve their genetic integrity. Prior to sediment collection approximately 5 cm of surface soil was removed from the wall prior to coring to minimise possible contamination. A 50 mL falcon tube was pushed into the pit wall to core out the sediment of each distinctive layer. Each tube was sealed, labelled appropriately and stored at 4 °C to be used in subsequent aDNA work. Previous sedimentary analyses of all 15 layers has been conducted to assess grain composition, size, colour, petrography and geochemistry (McDowell et al., 2013). Sub-samples from a total of 20 sediment cores were then subjected to aDNA analysis, targeting plant cpDNA and vertebrate mtDNA.

3.2. DNA extraction methods

All samples were extracted using methodologies designed for optimal aDNA recovery. Therefore, different extraction protocols were applied to bones and sediments and are described independently.

3.2.1. DNA extraction of bone

Fossil remains were ground to powder using a Dremel tool (part no. 114: Germany) set to high rotational speeds using DNA extraction method described by Haouchar et al. (2013). Briefly, ~0.2 g of bone powder was collected in a 1.5 mL Eppendorf tube, weighed and stored at 4 °C for later digestion. A number of bones were individually extracted (i.e. one bone per extraction digest), although where bones were morphologically uninformative and unidentified, a bulk-bone method was applied. This entailed grinding up to six discrete bones together and extracting DNA from the resulting powder for molecular analysis Murray et al. (2013).

Silica-based DNA extractions were performed (Haouchar et al., 2013) and all extracts quantitatively screened using primer sets targeting marsupials 12S and 16S rRNA markers (see Table 1 for primer combinations). Quantitative PCR assays were carried out using SYBR-green qPCR (Bunce et al., 2012) as described in the Supplementary data.

Table 1
Mitochondrial 12s rRNA/16s rRNA and Cytochrome *b* primer sequences and conditions used in this study to genetically characterise marsupial/mammal fossils and plants from sediments from KHC.

Primer name	Primer sequence 5′→3′	Annealing temperature (°C)	Amplicon size (bp)	Primer info
Primers for mammal/marsupial fossils				
12s_Macro_40F	GAYCTACACATGCAAGTTTCCGC	53	175	This study
12s_Macro_240R	CGGTGGCTGGCACGAGATTTAC			
12s_Macro_725F	GGAAAGYAATGGGCTACATTTTCTAA	60	115	This study
12s_Macro_843R	GCCTATTTCAATTAAGCTCTCTATTCT			
12s_Mars_520F	GGTCATAGCATTAACCCAAATTAACAG	55	170	This study
12s_Mars_690R	CTAATCCCAGTTTGTCTCTTAGCT			
16s_Mam1_F	CGGTTGGGGTGACCTCGGA	54	150	(Taylor, 1996)
16s_Mam1_R	GCTGTTATCCCTAGGGTAACT			
Cytb_Macropod_250F	CACGCTAACGGAGCATCCATATTC	56	160	This study
Cytb_Macropod_450R	GCCGATGTAGGGGATAGCGG			
Cytb_Macropod_400F	TACCGTGAGGACAAATATCATTCTGA	56	160	This study
Cytb_Macropod_600R	GAGCCTGTTTCGTGTAGGAATAG			
Primers for plants from sediment				
*trn*Lg	GGGCAATCCTGAGCCAA	54	90	(Taberlet et al., 2007)
*trn*Lh	TTGAGTCTCTGCACCTATC			
*rbc*L_F	GGCAGCATTCCGAGTAACTCCTC	53	100	(Chiang et al., 1998)
*rbc*L_*R*	CGTCCTTTGTAACGATCAAG			
Primers for snake fossils				
Cytb_Snake_55F1	CTCCACCTGATGAAACTTCGG	54	145	This study
Cytb_Snake_220R1	ATATGGATGCGCCGATTGCG			

3.2.2. DNA extraction from sediment

All sediments were processed using the Sergey Bulat extraction method optimised for small amounts of material including controls (Haile, 2012). Briefly, ~2 g of sediment were processed in the Bulat buffer and purified over silica columns (see Haile, 2012). Like the bone, DNA extractions were screened using qPCR that employed two generic primer sets for plants; *trn*Lg/h and *rbc*L. The *trn*Lg/h assay amplifies short sections of the *trn*L intron (Taberlet et al., 1991, 2007) and the *rbc*L primers target a coding segment of the plastid *rbc*L gene (Table 1) (Gielly and Taberlet, 1994; Chiang et al., 1998; Kress and Erickson, 2007). Concentrations for the DNA digest, qPCR set up and cycling conditions are further described in the Supplementary data.

3.3. Sequencing of DNA from bone and sediments

All bone extracts were screened for DNA amplification. Any extracts, which successfully yielded amplicons of the target size were purified (see Supplementary data for procedure) and sequenced. For single-source bone samples Sanger Sequencing was employed (using ABI, BigDye chemistry at Macrogen), for mixed samples and sediments a next generation HTS approach was used.

DNA extracts chosen for HTS were assigned a unique six base pair (bp) DNA tag (specifically a Multiplex Identifier-tag, MID-tag) and built into primers as fusion tags. All fusion-tagged PCRs were carried out in 25 μl reactions (see Supplementary data for master mix) and imaged by qPCR (Bunce et al., 2012). The general cycling conditions and amplicon purification can be found in Supplementary data. After pooling the amplicons, the library was then quantified with calibrated standards using qPCR (Bunce et al., 2012) to determine appropriate templating ratios for HTS. Emulsion PCR and GS Junior 454 Sequencing were performed as per Roche GS Junior protocols for amplicon sequencing (http://www.454.com).

Sediment DNA was prepared for GS junior 454 and Ion Torrent PGM sequencing using both *trn*Lg/h and *rbc*L primer sets (Table 1). Extracts were quantified in the same manner as the 454. All qPCRs were generated in triplicate and pooled accordingly. Amplicon pools were cleaned using Agencourt AMPure XP PCR Purification Kit (Beckman Coulter Genomics, NSW, Aus) and separated by gel electrophoresis. GS Junior 454 Sequencing and emulsion PCR were performed as per Roche GS Junior protocols. Ion Torrent emulsion PCR was performed on One Touch 2 system and sequencing was performed on an Ion Torrent (PGM).

3.4. Data analysis and sequence identification

All sequence reads generated from the two platforms (454 GS Junior and Ion Torrent PGM) for both bone and sediment extracts were filtered in a similar manner. First, sequences were sorted into sample batches based on the unique MID-tags using the program

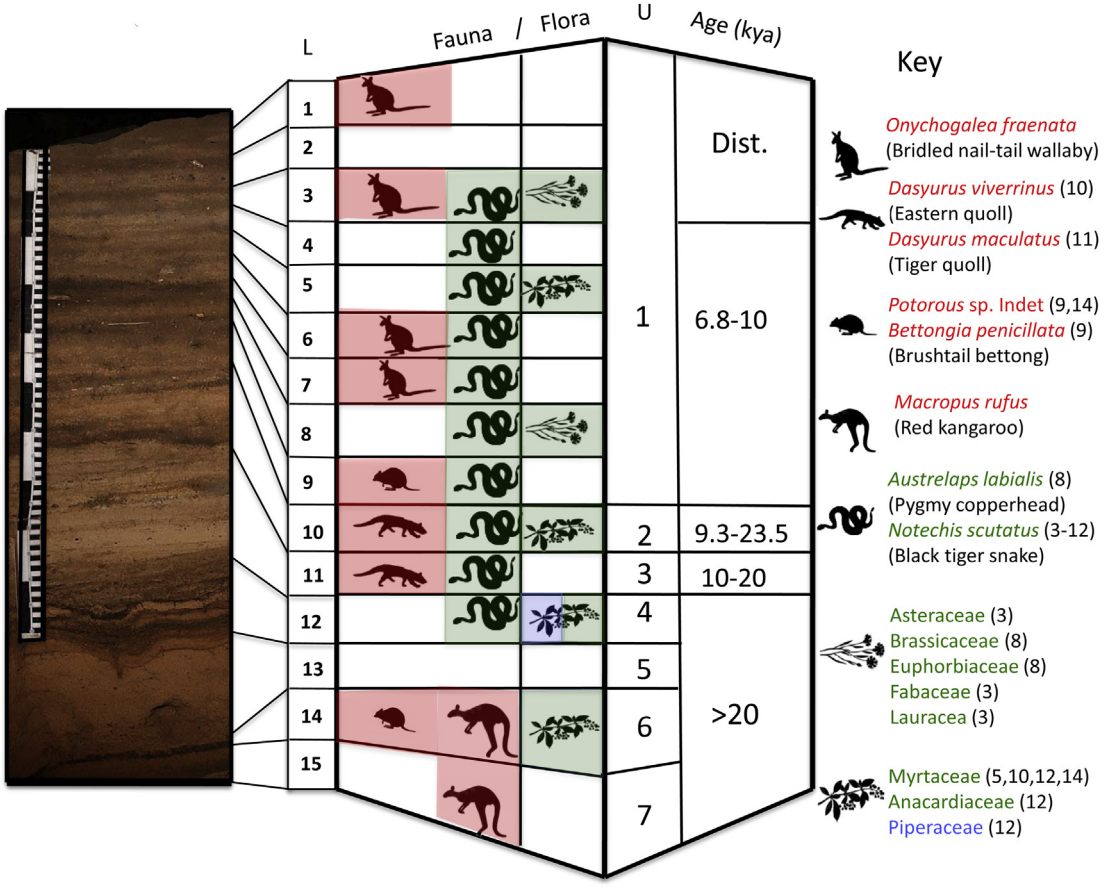


Fig. 2. A snapshot of the fauna and flora recovered throughout the layers of the excavation site in KHC. Units are indicated (U: 1–7) as well as Layers (L: 1–15). Age (ka) indicates the time over which each accumulated based on both U–Th and radiocarbon ages. Dating results suggest that a depositional hiatus occurred between accumulation of Units 3 and 2 and that Unit 2 has been reworked. The first three layers represent a disturbance (Dist.) followed by twelve undisturbed layers (4–15). Scale bar is 1.0 m in length (McDowell et al., 2013). Red shading indicates fauna and flora extirpated from the island; green shading indicates species still present on KI today; blue shading shows species not previously found on KI. Numbers in the key represent the layers each species was found, when multiple species are found in one layer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Geneious v5.6.5 (Drummond et al., 2011). Tags and primers were trimmed from the sequences allowing for no mismatch in length or base composition. All sequence results that were seemingly less than the expected amplicon size depending on the primer set used, were removed from the analysis. Once sequences were searched to contain all relevant information, i.e., they all retained the unique MID-tag, forward and reverse primer and adaptor primer, they were trimmed and searched against the NCBI GenBank nucleotide database (Benson et al., 2006) using BLASTn version 2.2.23 (Altschul et al., 1990) to identify reads. BLASTn datafiles were analysed in the program MEtaGenome Analyzer v4 (MEGAN) (Huson et al., 2007) (see Supplementary data for a detailed description of the analyses).

4. Results and discussion

4.1. Overview of sequence data of bones from KHC

Approximately 350,000 sequence reads were generated from multiple 454 (GS Junior) and Ion Torrent (PGM) runs. DNA from a range of taxa including *Macropus*, *Onychogalea*, *Potorous*, *Bettongia*, *Dasyurus*, *Rattus* and *Notechis* was amplified using an array of primers suitable to target the gene and organisms of interest (Table 1). The following results sections summarise the DNA analyses from KHC bones (Fig. 2). Initially, a number of bones that had previously been identified morphologically at family level were screened to investigate DNA preservation at the site, and confirm the taxonomic identity. Approximately 70% of the 19 bones that were randomly selected (dispersed evenly throughout the 15 layers) from the site yielded DNA, including those from the deeper layers 14 and 15 enabling amplification of between 100 and 250 bp sequences at a time. After this initial screening, another 45 bone fragments were subject to aDNA methods including an inventory of fragments from macropodid and murid limbs and vertebrae, and snake vertebrae.

4.2. Red kangaroo

Two postcranial bones identified as 'Macropodinae genus and species indeterminate' (sp. indet.) from layers 14 and 15 which are >20 ka (Fig. 2), yielded two sequences, one being 98% and the other 99% similar to red kangaroo (*Macropus rufus*), respectively. Morphologically diagnostic specimens of this species were not detected by McDowell (2013), but the species has previously been recorded on KI on the basis of three tooth fragments retrieved from Seton Rockshelter, nearby archaeological assemblage (Hope et al., 1977, Fig. 1). The presence of *M. rufus* in layers 14 and 15 likely represents a time when KI was connected to the mainland and probably preceded the LGM. Red kangaroos are currently widely distributed through the drier regions of mainland Australia but typically occur in arid and semi-arid regions (Van Dyck and Strahan, 2008; Jackson and Vernes, 2010). This finding is significant as it demonstrates that KHC is capable of long-term DNA preservation and provides additional impetus for further excavations to be

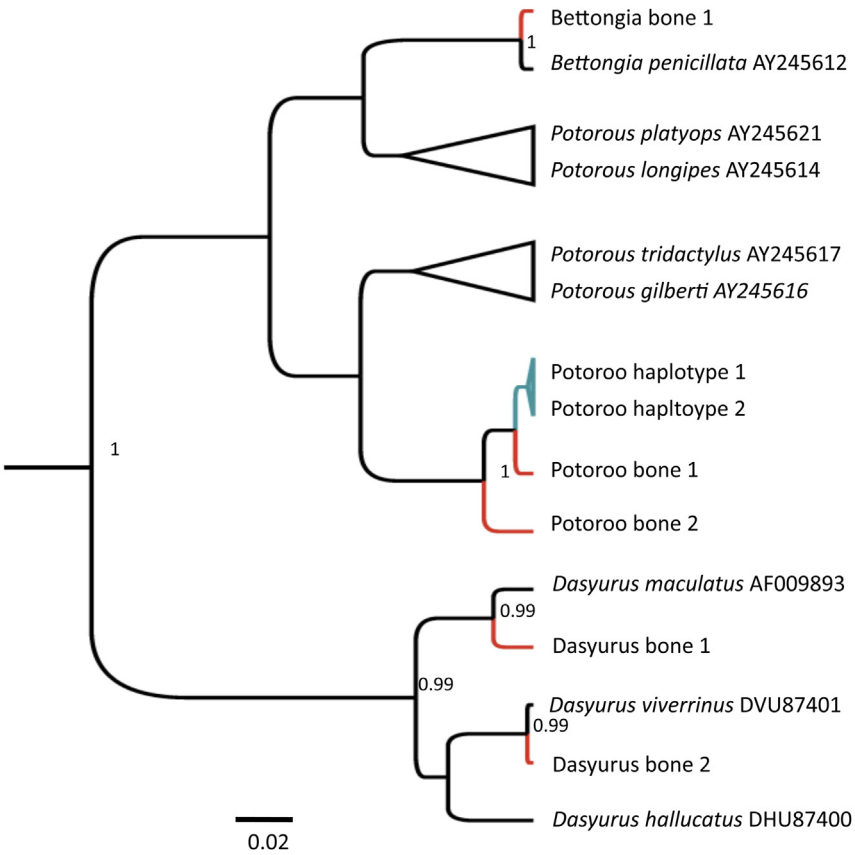


Fig. 3. Bayesian 12s rRNA phylogeny from 180 bp alignment showing the closest genetic match of the fossils from KHC to the mainland reference sequences retrieved from GenBank (accession numbers shown). Sequence data illustrates the relationships between *Bettongia*, *Potoroo* and *Dasyurus*. Blue coloured nodes are the result from bulk bone sampling; whilst red coloured nodes are single bones, single extraction samples. The tree was built using a HKY model and Yule tree prior with invariant gamma sites and imposed with a relaxed molecular clock. Values on node show >90% posterior probabilities. The scale represents the number of nucleotide substitutions per site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Summary of plants achieved using High-Throughput DNA sequencing and number of sequences from sediment cores taken from six layers in KHC, KI. Eight families were detected using a two-locus approach targeting the chloroplast *trn*L and *rbc*L genes, resulting in over 100,000 sequences. Families indicated in bold and genus is italicized. Approximate layer ages are indicated in Fig. 1. Plant locations were sourced from FloraBase [http://florabase.dec.wa.gov.au/], eFloraSA [http://flora.sa.gov.au] and Atlas of living Australia [http://www.ala.org.au/].

Plant taxa identification	Sedimentary layer and sampling depth from surface (KHC, KI)					
	Layer 3 (−10 cm)	Layer 5 (−18 cm)	Layer 8 (−59 cm)	Layer 10 (−80 cm)	Layer 12 (−98 cm)	Layer 14 (−130 cm)
Anacardiaceae					$n = 100$	
Anacardium sp.					$n = 36$	
Asteraceae	$n = 200^{\#\S}$					
Brassicaceae			$n = 11400^{\#\S}$			
Brassica sp.			$n = 10^{\S}$			
Euphorbiaceae			$n = 260^{\#\S}$			
Hevea sp.			$n = 160$			
Fabaceae	$n = 4000^{\#\S}$					
Lauraceae	$n = 350^{\#\S}$					
Cinnamomum sp.	$n = 7^{\S}$					
Myrtaceae		$n = 26500^{\#\S}$		$n = 10500^{\#\S}$	$n = 20700^{\#\S}$	$n = 14200^{\#\S}$
Myrcia sp.		$n = 7$			$n = 24$	$n = 5$
Eucalyptus sp				$n = 13^{\#\S}$	$n = 22^{\#\S}$	$n = 6^{\#\S}$
Leptospremum sp.		$n = 765^{\#\S}$		$n = 250^{\#\S}$	$n = 875^{\#\S}$	$n = 375^{\#\S}$
Syzygium sp.						$n = 8^{\S}$
Piperaceae					$n = 1000$	
Piper sp.					$n = 263$	

Key: # − Presence documented on Kangaroo Island; § − presence documented on South Australia mainland.
n = number of sequences in each family/genera detected using HTS.

carried out at this site. *M. rufus* probably became extirpated on KI as more wooded vegetation returned to KI.

4.3. Bridled nail-tail wallaby

Bridled nail-tail wallaby (*Onychogalea fraenata*) sequences were obtained from a combination of single-bone analysis and from bulk-bone sampling methods. A total of four haplotypes were observed with genetic differences distinct from the mainland individuals (71% to *O. unguifera* and 98% to *O. fraenata*). GenBank lacks reference sequences of the Crescent Nail-tail Wallaby (*O. lunata*) and although the geographic ranges of *O. fraenata* and *O. lunata* overlapped on the mainland until the late 19th century, only *O. fraenata* has been identified from KI based on morphological evidence (McDowell, 2013). *O. fraenata* is now highly endangered (Gordon and Lawrie, 1980; McKnight, 2008) and persists only in managed populations (Van Dyck and Strahan, 2008; Kingsley et al., 2012). *O. lunata* is extinct and *O. unguifera* (Northern Nail-tail wallaby) is widespread across northern Australia. Prior to European settlement *O. fraenata* and *O. lunata* were common throughout southern and eastern Australia (Van Dyck and Strahan, 2008). The critical status of *O. fraenata,* and its susceptibility to foxes, stock grazing and habitat destruction has resulted in a massive decline over the last 100 years. Pending further research, KI might be considered as a potential reintroduction site for *O. fraenataplease*.

4.4. Potoroo

DNA sequences with a potoroo affinity were obtained using both a single-bone extraction and bulk-bone methods. Overall, three haplotypes were observed which varied considerably from the reference mainland potoroo sequences acquired from GenBank. Two bones identified as Macropodinae sp. Indet. were individually extracted and analysed. A number of bones (5 in particular) identified as *Potorous platyops* from various layers collected (including surface scatter, layers 5 and 7) were subjected to the bulk bone method. An unresolved polytomy (Fig. 3) is revealed from the three samples (two single bone extractions and one bulk sample) with low posterior support values for the four haplotypes observed on KI, making species identification problematic. The cluster does however show that the species belongs in *Potorous*, since the percentage similarities between the mainland and island sequences show little discrepancy. Potoroo bone 1 (Fig. 3) and potoroo bone 2 share a similar match of 92.8% and 91.9% to *Potorous tridactylus* whilst haplotype 1 and 2 have a 98.3% and 98.5% similar match to *Potorous gilbertii*. As the species of *Potorous* is not clearly identified using this dataset, more bone fragments from KHC should be sequenced to identify whether the variability in this sequence cluster represents elevated genetic drift due to island isolation or potential for a new *Potorous* species, endemic to KI.

Prior to European settlement, potoroos were widely distributed across the continent. However, the combined pressure of habitat loss and introduced predators and competitors (Frankham et al., 2012) resulted in dramatic range reductions for most potoroos. *P. platyops* is extinct, *P. tridactylus* is listed as vulnerable, *P. longipes* is endangered and *P. gilbertii* is critically endangered. Since the time of European settlement *P. platyops* was only recorded as living in the south-west of Western Australia and was likely already extinct on KI at that time (Robinson and Armstrong, 1999). Fossil remains of *P. tridactylus* have been recovered from KHC (McDowell, 2013), but it appears to have been extirpated well before European settlement.

Another bone recovered from KHC analysed using the single bone extraction method, showed a genetic similarity of 99.3% to sequence from mainland specimens of the Brushtail Bettong or Woylie (*Bettongia penicillata*) (Fig. 3; Bettongia bone 1). Bettongs have also suffered dramatic range reductions due to anthropogenic effects and introduced predators throughout Australia (Claridge et al., 2007). The now extinct subspecies *B. penicillata penicillata* was once plentiful on the South Australian mainland, whilst the Burrowing Bettong or Boodie (*B. lesueur*) and *B. penicillata* have been identified from KI fossil bones (Hope et al., 1977; Robinson and Armstrong, 1999; McDowell, 2013). Both have been now extirpated from the island. Attempts to reintroduce the burrowing bettong (*B. lesueur*) back onto KI have met with limited success (Short et al., 1992; Robinson and Armstrong, 1999).

4.5. Dasyuridae

Sequences of quoll species (*Dasyurus*) were obtained using a single-bone extraction method. The results show that one yielded a DNA sequence 99% similar to Tiger Quoll, *Dasyurus maculatus* (Fig. 3; *Dasyurus* bone 1). Whilst DNA extracted from a small right femur provisionally identified by one of us (MCM) as cf. *Phascogale*, was found to be 99.5% similar genetically to mainland Eastern Quoll, *D. viverrinus* (Fig. 3; *Dasyurus* bone 2). These findings support physical fossil evidence that both species were present on KI. *D. viverrinus* appears to have been lost from KI during the mid-Holocene (McDowell, 2013) but *D. maculatus* persisted until it was extirpated by Europeans. During the 19th century, quolls were reported as being extant on KI until ca 1886 (Robinson and Armstrong, 1999), although no specimens are known to have been lodged in a museum. The presence of *D. maculatus* in a 200-year-old European fur-trapper accumulation at Bales Bay (Walshe, 2013) strongly points to this as the persistent species. Rehabilitation efforts and applications of genetic management to conserve quolls on the mainland (Jones et al., 2003) have taken place to reintroduce certain species back to their former ranges (Firestone et al., 1999, 2000). Direct evidence of fossils (McDowell, 2013), and now *Dasyurus* mtDNA haplotypes from KI suggests this top predator may be a candidate for reintroduction.

4.6. Muridae

A number of rodents have been collected from KI e.g., *Pseudomys shortridgei* (Heath rat), *Rattus lutreolus* (Swamp rat) and *Rattus fuscipes* (Bush rat). In this study crania morphologically identified as *R. fuscipes*, were genetically assessed using the bulk-bone method to confirm their species identity. The results show that all three groups were 98–99% similar to *R. fuscipes,* confirming the morphological identifications. Genetic variations within single nucleotide polymorphisms (SNPs) occur between samples and layers and probably between the different individuals sampled. However to prove this, further analyses of all bones sampled should be reanalysed individually using a single-bone extraction method in order to compare the SNPs throughout the layers and further compare this to the genetic population on the mainland.

R. fuscipes is common on KI and the mainland as well as 13 other continental islands off the coast of South Australia (Hinten et al., 2003). As a result of elevated sea levels following the LGM (Barrows et al., 2002; Petherick et al., 2008), genetically isolated populations developed on several of the newly formed islands (Schmitt, 1978; Seddon and Baverstock, 2002; Hinten et al., 2003). In distinct contrast to continental islands off Australia, a combination of mtDNA and microsatellite analyses (Hinten et al., 2003) indicate that the KI population of *R. fuscipes* has a greater level of genetic diversity than the adjacent mainland.

R. fuscipes has experienced population bottlenecking on the mainland (Hinten et al., 2003). Combined pressures imposed by feral cats, foxes and habitat fragmentation may be further contributing to these dwindling populations (Hinten et al., 2003). Given that KI has remained free from foxes, it serves as an ideal place in which to preserve the already diverse and self-maintaining *R. fuscipes* population and to host insurance populations.

4.7. Elapidae

The Pygmy Copperhead (*Austrelaps labialis*) and Eastern Black Tiger Snake (*Notechis ater*, although now believed to be synonymised with *Notechis scutatus*) were common on KI (Robinson and Armstrong, 1999; Houston and Tyler, 2002). Individuals of these species have 200–400 vertebrae that vary morphologically along the vertebral column making them difficult to identify to species when found as fossils. However, because each snake has so many vertebrae they are common in Australian palaeontological excavations. Elapid vertebrae consist largely of thick cortical bone and typically preserve high-quality DNA that can be specifically identified using molecular methods (Polly and Head, 2004).

The bulk-bone method (Murray et al., 2013) allowed rapid initial identification of the snake species present, and then single bone extractions verified the layers in which the species occurred (see Fig. 2). Elapid DNA was amplified from strata that have been radiometrically dated to >20 ka (layer 11: McDowell et al., 2013) and yielded 100% and 99.5% similar matches to the *N. ater* and *Austrelaps labialis* respectively (the only two species recognised from excavated crania). Little genetic difference was observed within species over time, nor when compared to the modern mainland reference sequences. The Pygmy Copperhead, which is widespread and common on KI, has a depauperate mainland range and is found primarily restricted to the southern Mount Lofty ranges and Fleurieu Peninsula. Habitat destruction is probably the main force driving the decline of this species (Robinson and Armstrong, 1999; Houston and Tyler, 2002).

4.8. Overview of plant data from sediments

A total of eight plant families were detected (Table 2) from an initial assessment of ancient DNA preservation in sediment (*sed*aDNA). DNA was screened for 12 sediment sub-samples ranging from >20 ka to approximately 9 ka. Sediment DNA was also tested for mtDNA although no animal DNA was detected. A total of six sediment samples (Fig. 2 and Table 2) successfully yielded DNA using a two-locus approach targeting the chloroplast *trn*L and *rbc*L genes resulting in approximately 100,000 sequences from several HTS runs. Sequences were identified using BLAST (Altschul et al., 1990), and results interpreted with MEGAN (Huson et al., 2007) (Table 2). As with previous sedimentary ancient plant DNA literature using these chloroplast loci (Parducci et al., 2005; Jorgensen et al., 2011), taxonomic resolution is reliant upon comparative database coverage, and within certain families is constrained by the degree of interspecific variation. Nevertheless, the results provide some interesting insights. There are nearly 900 native plant species recorded on KI which currently include 40 endemic species (Holiday et al., 2003). Families which are well represented on KI are those characteristic of the higher rainfall areas of South Australia in general such as Myrtaceae (*e.g.*, eucalyptus), Cyperaceae (e.g., sedges), Liliaceae (e.g., lilies), Ericaceae (formerly – Epacridaceae e.g., heath shrubs and herbs) and Apiaceae (formerly – Umbelliferae *e.g.*, heath shrubs).

The majority of species present in the layers were found on KI and/or the adjacent mainland. However, some of the deeper layers yielded taxa not found on KI or mainland South Australia. For example layer 12 (dated at >20 ka) yielded *Piper* sp. indet. (Piperaceae – pepper family) (Table 2), from *sed*aDNA, but has never been recorded on the island before. Closely related species appear to be common in the eastern states of Australia, particularly all along the coast of Queensland and New South Wales (Atlas of living Australia; http://www.ala.org.au/). Species of the Myrtaceae family, which are considerably common throughout KI, was found in four sedimentary layers, and included a number of the taxa that could be identified to genus (Table 2). *Eucalyptus* sp., *Leptospermum sp.* and *Syzygium* sp., were present in the deeper excavation layers and with the exception of *Syzygium* sp., which remains common on the mainland, all can be found on KI today.

The study of past plant species distribution (palynology), which principally relies on pollen, has a long and venerable history. However, until the discovery of aDNA persistence in sediments

(Willerslev et al., 2003), the absence of macro-fossil remains (seeds, buds or vegetative tissue) and micro-fossils (pollen) was a serious limitation (Jorgensen et al., 2011). Ancient sediments, in particular permafrost deposits, have proven to be an excellent archive for the long-term preservation of environmental ancient DNA (*sed*aDNA) (Willerslev et al., 2003), making them useful for palaeo-reconstructions. However, *sed*aDNA preservation in less favourable climates like that of Australia can be somewhat challenging and so far have only been reproducible in a minority of plant aDNA studies. To date no record has been made on the preservation of plant macro- and micro-remains in the sediments of KHC, therefore any information that can be gathered of past floral assemblages will provide a significant contribution to our understanding of paleo-vegetation and climate changes.

5. Conclusion

This is the first aDNA assessment of vertebrate fossils and plant DNA from KI in South Australia. Outcomes presented here add significant value to the late Pleistocene–mid-Holocene palaeontological record of KI. We demonstrate the utility of aDNA techniques when applied to KI cave deposits and build upon results of morphological studies by confirming existing identifications and revealing additional species not preserved as diagnostic fossil specimens. The excellent preservation of both animal and plant DNA extracted from KHC bones and sediments are fundamental in providing such valuable information about the past biodiversity of KI. This study provides a detailed molecular record of animal and plant species that once lived in the KHC region, including some species that have been extirpated from KI. We reveal an array of taxa ranging from marsupials and reptiles to shrubs and trees, some of which were deposited >20 ka. Our results also highlight the potential of the ongoing excavation that will delve even deeper (>15 layers) into the past and may provide insights into the extinction of megafaunal species previously recorded on KI (Hope et al., 1977) and at the KHC site (McDowell, 2013).

Acknowledgements

Thanks to Grant Gully for modifications of Fig. 1. This study was funded by ARC DP120104435 to GJP and MB. MB is funded by an Australian Research Council Future Fellowship (FT0991741).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.quascirev.2013.11.007.

References

Abbott, I., 1974. The avifauna of Kangaroo Island and causes of its impoverishment. Emu 74, 124–134.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Barrows, T.T., Stone, J.O., Fifield, L.K., Cresswell, R.G., 2002. The timing of the last glacial maximum in Australia. Quat. Sci. Rev. 21, 159–173.
Belperio, A.P., Flint, R.B., 1999. Geomorphology and geology. In: Robinson, A.C., Armstrong, D.M. (Eds.), A Biological Survey of Kangaroo Island, South Australia in November 1989 and 1990. Heritage and Biodiversity Section, Department for Environment, Heritage and Aboriginal Affairs, Adelaide, pp. 19–31.
Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2006. GenBank. Nucleic Acids Res. 34, 16–20.
Bradley, R.S., 1999. Paleoclimatology. Reconstructing Climates of the Quaternary, second ed.. In: International Geophysics Series Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto. Geological Magazine 138, pp. 501–502.
Bunce, M., Oskam, C.L., Allentoft, M.E., 2012. Quantitative real-time PCR in aDNA research. In: Shapiro, B., Hofreiter, M. (Eds.), Ancient DNA. Springer, pp. 121–132.
Butzer, K.W., 2008. Challenges for a cross-disciplinary geoarchaeology: the intersection between environmental history and geomorphology. Geomorphology 101, 402–411.
Chappell, J., Shackleton, N., 1986. Oxygen isotopes and sea level. Nature 324, 137–140.
Chiang, T.Y., Schaal, B.A., Peng, C.I., 1998. Universal primers for amplification and sequencing a noncoding spacer between the atpB and rbcL genes of chloroplast DNA. Bot. Bull. Acad. Sin. 39, 245–250.
Claridge, A., Seebeck, J., Rose, R., 2007. Potoroos and the Musky Rat-kangaroo. CSIRO Publishing, Collingwood, Vic, Australia.
Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.
Cutler, K., Edwards, R., Taylor, F., Cheng, H., Adkins, J., Gallup, C., Cutler, P., Burr, G., Bloom, A., 2003. Rapid sea-level fall and deep-ocean temperature change since the last interglacial period. Earth Planet. Sci. Lett. 206, 253–271.
Darwin, C.R., Wallace, A.R., 1858. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. J. Proc. Linn. Soc. Lond. 3, 46–50.
De Bruyn, M., Hoelzel, A.R., Carvalho, G.R., Hofreiter, M., 2011. Faunal histories from Holocene ancient DNA. Trends Ecol. Evol. 26, 405–413.
Diamond, J.M., 1972. Biogeographic kinetics: estimation of relaxation times for avifaunas of southwest Pacific islands. Proc. Natl. Acad. Sci. 69, 3199–3203.
Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T., Wilson, A., 2011. Geneious V. 6.0.1. Available from: http://www.geneious.com/.
Firestone, K.B., Elphinstone, M.S., Sherwin, W.B., Houlden, B.A., 1999. Phylogeographical population structure of tiger quolls Dasyurus maculatus (Dasyuridae: Marsupialia), an endangered carnivorous marsupial. Mol. Ecol. 8, 1613–1625.
Firestone, K.B., Houlden, B.A., Sherwin, W.B., Geffen, E., 2000. Variability and differentiation of microsatellites in the genus Dasyurus and conservation implications for the large Australian carnivorous marsupials. Conserv. Genet. 1, 115–133.
Foufopoulos, J., Ives, A.R., 1999. Reptile extinctions on land-bridge islands: life-history attributes and vulnerability to extinction. Am. Nat. 153, 1–25.
Frankham, G.J., Handasyde, K.A., Eldridge, M.D.B., 2012. Novel insights into the phylogenetic relationships of the endangered marsupial genus *Potorous*. Mol. Phylogenet. Evol. 64, 592–602.
Fulton, T.L., 2012. Setting up an ancient DNA laboratory. In: Shapiro, B., Hofreiter, M. (Eds.), Ancient DNA: Methods and Protocols. Springer Science, pp. 1–11.
Gielly, L., Taberlet, P., 1994. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus rbcL sequences. Mol. Biol. Evol. 11, 769–777.
Gordon, G., Lawrie, B.C., 1980. The rediscovery of the bridled nailt-tail wallaby, Onychogalea fraenata (Gould) in Queensland. Aust. Wildl. Res. 7, 339–345.
Haile, J., 2012. Ancient DNA extraction from soils and sediments. In: Shapiro, B., Hofreiter, M. (Eds.), Ancient DNA. Humana Press, pp. 57–63.
Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M.T.P., Nielsen, R., Munch, K., Ho, S.Y.W., Shapiro, B., Willerslev, E., 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? Mol. Biol. Evol. 24, 982–989.
Haouchar, D., Haile, J., Spencer, P.B.S., Bunce, M., 2013. The identity of the Depuch Island rock-wallaby revealed through ancient DNA. Aust. Mammal. 35, 101–106.
Hinten, G., Harriss, F., Rossetto, M., Braverstock, P., 2003. Genetic variation and island biogeography: microsatellite and mitochondrial DNA variation in island populations of the Australian bush rat, Rattus fuscipes greyii. Conserv. Genet. 4, 759–778.
Holiday, I., Overton, B., Overton, D., 2003. Kangaroo Island's Native Plants. Finsbury Printing, Adelaide, South Australia.
Hope, J., Lampert, R., Edmondson, E., Smith, M., Van Tets, G., 1977. Late Pleistocene faunal remains from Seton rock shelter, Kangaroo Island, South Australia. J. Biogeogr. 4, 363–385.
Houston, T.F., Tyler, M.J., 2002. Reptiles and Amphibians. In: Davies, M., Twidale, C., Tyler, M.J. (Eds.), Natural History of Kangaroo Island. Royal Society of South Australia, Adelaide.
Huson, D.H., Auch, A., Qi, J., Schuster, S.C., 2007. Megan analysis of metagenome data. Genome Res. 17, 377–386.
Jackson, S., Vernes, K., 2010. Kangaroo: Portrait of an Extraordinary Marsupial. Allen and Unwin, NSW.
Jones, M.E., Oakwood, M., Belcher, C.A., Morris, K., Murray, A.J., Woolley, P., Firestone, K.B., Johnson, B., Burnett, S., 2003. Carnivore Concerns: Problems, Issues and Solutions for Conserving Australasia's Marsupial Carnivores, Predators with Pouches: the Biology of Carnivorous Marsupials. CSIRO Publishing, pp. 422–434.
Jorgensen, T., Haile, J., Moller, P., Andreev, A., Boessenkool, S., Rasmussen, M., Kienast, F., Coissac, E., Taberlet, P., Brochmann, C., 2011. A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. Mol. Ecol. 21, 1989–2003.
Kingsley, L., Goldizen, A., Fisher, D.O., 2012. Establishment of an endangered species on a private nature refuge: what can we learn from reintroductions of the bridled nailtail wallaby *Onychogalea fraenata*? Oryx 46, 240–248.
Kress, W.J., Erickson, D.L., 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS One 2, e508.
Lampert, R.J., 1981. The Great Kartan Mystery. Terra Australis 5. Australian National University, Canberra.

Leonard, J.A., 2008. Ancient DNA applications for wildlife conservation. Mol. Ecol. 17, 4186–4196.
Lister, A.M., 2004. The impact of Quaternary ice ages on mammalian evolution. Philos. Trans. Royal Soc. Lond. Ser. B: Biol. Sci. 359, 221–241.
Lomolino, M.V., Brown, J.H., Davis, R., 1989. Island biogeography of montane forest mammals in the American Southwest. Ecology 70, 180–194.
Losos, J.B., Ricklefs, R.E., 2010. The Theory of Island Biogeography Revisted. Princeton University Press, Princeton, NJ.
MacArthur, R.H., Wilson, E.O., 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, NJ.
McDowell, M.C., 2013. Late Quaternary Fuanal Responses to Environmental Change and Isolation on a Large Australian Land-Bridge Island. School of Biological Sciences Flinders University, South Australia.
McDowell, M.C., Bestland, E.A., Bertuch, F., Ayliffe, L.K., Hellstrom, J.C., Jacobsen, G.E., Prideaux, G.J., 2013. Chronology, stratigraphy and palaeoenvironmental interpretation of a Late Pleistocene to mid-Holocene cave accumulation on Kangaroo Island, South Australia. Boreas 42, 974–994.
McKnight, M., 2008. Onychogalea fraenata. In: IUCN 2013. IUCN Red List of Threatened Species, p. 1. Version 2013.
Murray, D.C., Haile, J., Dortch, J., White, N., Haouchar, D., Bellgard, M.I., Allcock, R.J., Prideaux, G.J., Bunce, M., 2013. Scrapheap challenge: a novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. Sci. Rep 3, 3371. http://dx.doi.org/10.1038/srep03371.
Pääbo, S., 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc. Natl. Acad. Sci. U. S. A. 86, 1939–1943.
Parducci, L., Suyama, Y., Lascoux, M., Bennett, K.D., 2005. Ancient DNA from pollen: a genetic record of plant population history. Mol. Ecol. 14, 2873–2882.
Petherick, L., McGowan, H., Moss, P., 2008. Climate variability during the Last Glacial Maximum in eastern Australia: evidence of two stadials? J. Quat. Sci. 23, 787–802.
Pledge, N.S., 1979. Fossil vertebrates. In: Tyler, M.J., Twidale, C., Ling, J.K. (Eds.), Natural History of Kangaroo Island. Royal Society of Southern Australia, Adelaide.
Polly, P.D., Head, J.J., 2004. Maximum-likelihood Identification of Fossils: Taxonomic Identification of Quaternary Marmots (Rodentia, Mammalia) and Identification of Vertebral Position in the Pipesnake Cylindrophis (Serpentes, Reptilia), Morphometrics: Applications in Biology and Paleontology. Springer, p. 197.
Prideaux, G.J., Gully, G.A., Couzens, A.M.C., Ayliffe, L.K., Jankowski, N.R., Jacobs, Z., Roberts, R.G., Hellstrom, J.C., Gagan, M.K., Hatcher, L.M., 2010. Timing and dynamics of Late Pleistocene mammal extinctions in southwestern Australia. Proc. Natl. Acad. Sci. 107, 22157–22162.
Prideaux, G.J., Long, J.A., Ayliffe, L.K., Hellstrom, J.C., Pillans, B., Boles, W.E., Hutchinson, M.N., Roberts, R.G., Cupper, M.L., Arnold, L.J., 2007. An arid-adapted middle Pleistocene vertebrate fauna from south-central Australia. Nature 445, 422–425.
Ramakrishnan, U., Hadly, E.A., Mountain, J.L., 2005. Detecting past population bottlenecks using temporal genetic data. Mol. Ecol. 14, 2915–2922.
Robinson, A.C., Armstrong, D.M., 1999. A Biological Survey of Kangaroo Island South Australia in November 1989 and 1990. Department for Environment, Heritage and Aboriginal Affairs, Adelaide.
Schmitt, L.H., 1978. Genetic variation in isolated populations of the Australian bush-rat, Rattus fuscipes. Evolution 38, 1–14.
Seddon, J., Baverstock, P., 2002. Variation on islands: major histocompatibility complex (Mhc) polymorphism in populations of the Australian bush rat. Mol. Ecol. 8, 2071–2079.
Short, J., Bradshaw, S.D., Giles, J., Prince, R.I.T., Wilson, G.R., 1992. Reintroduction of macropods (Marsupialia: Macropodoidea) in Australia – a review. Biol. Conserv. 62, 189–204.
Stiller, M., Baryshnikov, G., Bocherens, H., d'Anglade, A.G., Hilpert, B., Munzel, S.C., Pinhasi, R., Rabeder, G., Rosendahl, W., Trinkaus, E., 2010. Withering away – 25,000 years of genetic decline preceded cave bear extinction. Mol. Biol. Evol. 27, 975–978.
Stone, A.C., 2000. Ancient DNA from skeletal remains. In: Katzenberg, M.A., Saunders, S.R. (Eds.), Biological Anthropology of the Human Skeleton. Wiley-Liss, New York, pp. 351–371.
Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res. 35, 14–21.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.
Twidale, C., Bourne, J.A., 2002. The land surface. In: Tyler, M.J., Twidale, C.R., Ling, J.K. (Eds.), Natural History of Kangaroo Island. Royal Society of South Australia, Adelaide.
Van der Geer, A., Lyras, G., De Vos, J., Dermitzakis, M., 2010. Evolution of Island Mammals: Adaptation and Extinction of Placental Mammals on Islands. John Wiley & Sons, London.
Van Dyck, S., Strahan, R., 2008. The Mammals of Australia Sydney. Reed New Holland.
Walshe, K., 2013. Archaeological evidence for a sealer's and wallaby hunter's skinning site on Kangaroo Island, South Australia. J. Isl. Coastal Archaeol. (in press).
White, W.B., 2007. Cave sediments and paleoclimate. J. Cave Karst Stud. 69, 76–93.
Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M.B., Brand, T.B., Hofreiter, M., Bunce, M., Poinar, H.N., Dahl-Jensen, D., Johnsen, S., Steffensen, J.P., Bennike, O., Schwenninger, J.-L., Nathan, R., Armitage, S., de Hoog, C.-J., Alfimov, V., Christl, M., Beer, J., Muscheler, R., Barker, J., Sharp, M., Penkman, K.E.H., Haile, J., Taberlet, P., Gilbert, M.T.P., Casoli, A., Campani, E., Collins, M.J., 2007. Ancient biomolecules from deep ice cores reveal a forested southern Greenland. Science 317, 111–114.
Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A., 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300, 791–795.
Yokoyama, Y., De Deckker, P., Lambeck, K., Johnston, P., Fifield, L., 2001. Sea-level at the Last Glacial Maximum: evidence from northwestern Australia to constrain ice volumes for oxygen isotope stage 2. Palaeogeogr. Palaeoclimatol. Palaeoecol. 165, 281–297.

APPENDIX B – INFERRING KANGAROO PHYLOGENY FROM INCONGRUENT NUCLEAR AND MITOCHONDRIAL GENES.

The following citation resulted from research that is of relevance to this doctoral thesis and was published during my candidature. In context of this multi-author publication I extracted and sequenced the mitochondrial and nuclear DNA from several of the macropod individuals used in this study.

A copy of the main article is attached. The Supplementary material may be viewed online at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0057745#s5

Phillips, M.J., **Haouchar D**., Pratt, R.C., Gibb, G.C., Bunce, M., 2013. Inferring kangaroo phylogeny from incongruent nuclear and mitochondrial genes. *PloS ONE* **8**, e57745.

OPEN ACCESS Freely available online

PLOS ONE

Inferring Kangaroo Phylogeny from Incongruent Nuclear and Mitochondrial Genes

Matthew J. Phillips[1]*, Dalal Haouchar[2], Renae C. Pratt[3], Gillian C. Gibb[4], Michael Bunce[2]

1 School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, Queensland, Australia, **2** School of Veterinary and Life Sciences, Murdoch University, Perth, Western Australia, Australia, **3** Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia, **4** Institute of Agriculture and Environment, Massey University, Palmerston North, New Zealand

Abstract

The marsupial genus *Macropus* includes three subgenera, the familiar large grazing kangaroos and wallaroos of *M. (Macropus)* and *M. (Osphranter)*, as well as the smaller mixed grazing/browsing wallabies of *M. (Notamacropus)*. A recent study of five concatenated nuclear genes recommended subsuming the predominantly browsing *Wallabia bicolor* (swamp wallaby) into *Macropus*. To further examine this proposal we sequenced partial mitochondrial genomes for kangaroos and wallabies. These sequences strongly favour the morphological placement of *W. bicolor* as sister to *Macropus*, although place *M. irma* (black-gloved wallaby) within *M. (Osphranter)* rather than as expected, with *M. (Notamacropus)*. Species tree estimation from separately analysed mitochondrial and nuclear genes favours retaining *Macropus* and *Wallabia* as separate genera. A simulation study finds that incomplete lineage sorting among nuclear genes is a plausible explanation for incongruence with the mitochondrial placement of *W. bicolor*, while mitochondrial introgression from a wallaroo into *M. irma* is the deepest such event identified in marsupials. Similar such coalescent simulations for interpreting gene tree conflicts will increase in both relevance and statistical power as species-level phylogenetics enters the genomic age. Ecological considerations in turn, hint at a role for selection in accelerating the fixation of introgressed or incompletely sorted loci. More generally the inclusion of the mitochondrial sequences substantially enhanced phylogenetic resolution. However, we caution that the evolutionary dynamics that enhance mitochondria as speciation indicators in the presence of incomplete lineage sorting may also render them especially susceptible to introgression.

Citation: Phillips MJ, Haouchar D, Pratt RC, Gibb GC, Bunce M (2013) Inferring Kangaroo Phylogeny from Incongruent Nuclear and Mitochondrial Genes. PLoS ONE 8(2): e57745. doi:10.1371/journal.pone.0057745

Editor: Axel Janke, BiK-F Biodiversity and Climate Research Center, Germany

Received October 31, 2012; **Accepted** January 24, 2013; **Published** February 22, 2013

Funding: This work has been supported by Australian Research Council grants to MJP (DP07745015) and MB (FT0991741). The website for the funder is www.arc.gov.au. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mphillips.biol@gmail.com

Introduction

The family Macropodidae includes more than 60 species of bipedal hopping kangaroos and wallabies living throughout Australia, New Guinea and surrounding islands. The family has Late Oligocene-Early Miocene rainforest origins and its diversification primarily coincides with subsequent aridification, during which woodland and grassland habitats expanded [1,2]. The most iconic and species-rich group of macropodids to exploit these more open mesic to semi-arid habitats is the genus *Macropus*. The 13 extant species are divided into three subgenera, (i) *M. (Macropus)*, including the grey kangaroos, (ii) *M. (Osphranter)*, including the red kangaroo and wallaroos and (iii) the *M. (Notamacropus)* wallabies.

Body size and foraging ecology vary substantially among kangaroos and wallabies. Species are sexually size-dimorphic (e.g. mean adult body mass in the red kangaroo, *M. rufus* is 26 kg ♀/66 kg ♂ [3]), although foraging is broadly similar among the sexes. Sanson [4] characterised macropodid dental grades associated with foraging ecology, contrasting browsers of dicotyledonous plants with grazes feeding primarily on grasses. Predominance of grazing and larger adult body mass (averaged over males and females [3,5]) distinguish *M. (Macropus)* (26–33 kg) and *M. (Osphranter)* (17–46 kg) from the smaller, typically mixed browsing/grazing *M. (Notamacropus)* (4–16 kg).

Cardillo et al. [6] inferred a marsupial supertree from a comprehensive survey of earlier molecular and morphological phylogenies. This summary tree (modified in Figure 1A) closely matches the subsequent study by Meredith et al. [7], which sampled DNA sequences for five nuclear genes, including for 11 of the 13 *Macropus* species (Figure 1B). The two differences from the Figure 1A summary concern the relative affinities of the three *Macropus* subgenera and the placement of the monotypic *Wallabia bicolor*. In the former case, Meredith et al. [7] group *M. (Osphranter)* with *M. (Notamacropus)* to the exclusion of *M. (Macropus)*, though with weak support. The interrelations of these three subgenera have remained opaque to all data sources. Even consistent morphological support for grouping the larger *M. (Osphranter)* and *M. (Macropus)* hinges primarily on dental and palatal characters that may instead reflect correlations with grazing [8].

The more striking difference is Meredith et al. 's [7] placement of *W. bicolor* within *Macropus*, either as sister to *M. (Notamacropus)* or as sister to the more inclusive *M. (Notamacropus)/M. (Osphranter)* clade. On this basis the authors suggested subsuming *Wallabia* into *Macropus*, with subgeneric status for *M. (Wallabia)*. Many early workers [1,9] also grouped *W. bicolor* with members of *M. (Notamacropus)* in the genus *Protemnodon* (which now includes only extinct members). Ride [10] however, noted that parallelism and plesiomorphy could explain anatomical similarities between *W.*

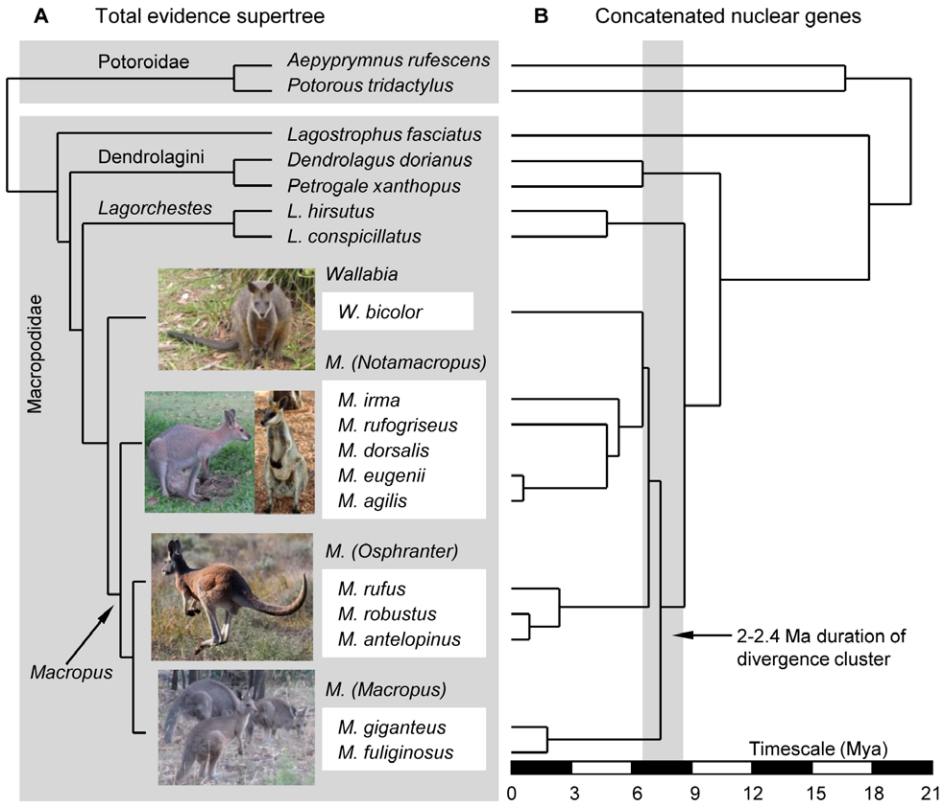


Figure 1. Phylogenetic relationships of *Wallabia* and the three *Macropus* subgenera, *M. (Macropus)*, *M. (Osphranter)* and *M. (Notamacropus)*. (A) The supertree of Cardillo et al. [6] summarizing previous molecular and morphological phylogenies and (B) Meredith et al.'s [7] evolutionary timescale (ave. of four BEAST analyses), showing the 2–2.4 Ma duration divergence cluster. Both trees are modified to include only the taxa sampled in the present study. Dendrolagini was not recovered by Cardillo et al. [6], however its inclusion in the summary tree is warranted on subsequent strong evidence from morphology [2] and all recent molecular analyses. Photos include (from the top) *W. bicolor*, *M. rufogriseus* (left), *M. irma* (right), *M. rufus and M. giganteus*. Photo credits – Matt Phillips, except *M. irma* (Ric Dawson) and *M. rufus* (Daniel Hoops).
doi:10.1371/journal.pone.0057745.g001

bicolor and species now placed in *M. (Notamacropus)*, which overlap in size (*W. bicolor* adult mean, 15 kg [5]) and in foraging habits – although *W. bicolor* is more specialized for browsing. Morphological studies [2,8,11–13] and behavioural analysis [14] have since favoured placing *W. bicolor* outside of *Macropus*, although without identifying characters that provide unambiguous support.

Earlier molecular studies have been similarly indecisive on the relationship of *Wallabia* to *Macropus*. Analyses of allozymes [15] and mitochondrial 12/16SrRNA+tRNA-valine sequences [16,17] favoured *Macropus* monophyly, to the exclusion of *W. bicolor*. Meanwhile, serology [18], microcomplement fixation [19] and DNA-DNA hybridization [20] tended to favour *W. bicolor* falling within *Macropus*, albeit often in different positions. Mitochondrial (mt) *Cytb* and nuclear *Selenocysteine tRNA* [21] did not clearly resolve affinities between the subgenera, while *Protamine P1* [22] favoured grouping *W. bicolor* with *M. rufogriseus*, leaving not only *Macropus*, but also *M. (Notamacropus)* paraphyletic.

This study expands the available 12S/16SrRNA and *Cytb* sequences and adds new *NADH1* and *NADH2* protein-coding sequences. Together, these provide a 5.6 kb mtDNA dataset for *W. bicolor*, nine *Macropus* species and seven outgroup macropodids and potoroids. The new sequences include the first molecular data for *M. dorsalis* and the first mtDNA for *M. irma*. Both of these wallabies are classified on morphology as members of *M. (Notamacropus)* [8].

We provide a more comprehensive examination of species relationships among kangaroos and wallabies by analysing the mtDNA alongside the five published nuclear genes (*BRCA1*, *IRBP*, *RAG1*, *ApoB* and *vWF*) from Meredith et al. [7]. Combining mt and nuclear sequences has previously provided strong statistical power for resolving family and ordinal-level marsupial relationships [23,24]. However, concatenation is expected to mask uncertainty and potentially bias inference of relationships among closely diverged species, where multiple gene lineages persist through speciation events (incomplete lineage sorting, ILS) [25,26].

We employ three "species tree" methods for combined analyses of the mt and nuclear genes in order to account for ILS among gene trees. The first of these, *BEAST [27] is highly parametric, employs the multi-species coalescent model and co-estimates the individual gene trees embedded within the species tree. The second, minimizing deep coalescences (MDC [28]) is a non-parametric alternative that uses a parsimony algorithm to identify the species tree requiring the fewest deep coalescent events among specified gene trees. The third species tree approach, Bayesian concordance analysis within BUCKy 1.4 [29] models gene tree incongruence while accounting for stochastic variation within posterior or bootstrap distributions of gene trees. Importantly, BUCKy does not assume any particular source of gene tree incongruence, unlike *BEAST and MDC, which both assume incongruence derives from ILS.

The potential importance of post-speciation gene flow in the present study is underlined by introgressive hybridization having been identified among natural populations of parapatric rock wallaby species (*Petrogale spp.* [30,31]) and between the grey kangaroos, *M. giganteus* and *M. fuliginosus* [32]. Introgression however, can be difficult to distinguish from ILS [33,34]. We use a simulation approach [35,36] to distinguish these sources of incongruence.

The role of mtDNA for inferring relationships among closely related animal species has been much argued recently [37–41]. In several regards the mt genome should be an excellent marker. In contrast to the high rates of duplication and translocation of nuclear genes, the mt genome offers near-certain orthology for mammals, as long as appropriate practices are employed to avoid nuclear copies of mt genes [42]. Moreover, mitochondrial haploidy and uniparental inheritance confer ~4-fold lower effective population size (N_e) relative to nuclear DNA, such that mtDNA is expected to be a "leading indicator" of speciation [43]. Mitochondrial N_e and coalescent times may often be even further reduced by strong selection [37].

On the flip side of these arguments, population structure can diminish the influence of lower N_e on coalescence times [44]. Furthermore, the lack of recombination tends to lead genomes into fitness traps via a process known as Muller's ratchet [45]. Lower N_e and higher mutation rates can serve to accelerate this ratchet [46]. Resulting differences in mean fitness between populations and species can drive introgression of mtDNA, as demonstrated in *Drosophila* [47].

In this study we examine the utility of mtDNA for complementing nuclear sequences in reconstructing the phylogeny of kangaroos and wallabies. Inclusion of mtDNA substantially improves phylogenetic resolution of clades that have apparently been subject to incomplete lineage sorting among nuclear loci. In turn, it is encouraging that the nuclear signal overwhelms the mitochondrial signal where the latter is discordant with both the nuclear and morphological data.

Materials and Methods

Ethics Statement

DNA and tissue samples were obtained from pre-existing collections as donations from The Australian Centre for Ancient DNA (University of Adelaide), The Research School of Biology (Australian National University), The Department of Environment and Conservation, Western Australia and as loans from The Australian National Wildlife Collection, Canberra – in each case with permission from the relevant authorities within these institutions. One additional frozen tissue sample was purchased from a local butcher (EcoMeats) in Canberra. No live animals were sampled and none of the DNA/tissue collection or handling procedures required either approval or a permit from a review board or ethics committee. Sequences published previously by other groups were obtained from GenBank.

Taxon sampling and DNA sequencing

In order to reconstruct a mitochondrial tree for kangaroos and wallabies we targeted three protein-coding genes, *NADH1*, *NADH2* and *Cytb* along with the 12S and 16S ribosomal RNA genes. Taxon sampling focused on ten *Macropus* species and the monotypic *Wallabia bicolor*. The affinities of the three extant *Macropus* species not included here are uncontroversial [6,8,48] and add little to the sampled diversity. *M. parma* and *M. parryi* are nested within *M. (Notamacropus)* and *M. bernardus* groups with the other wallaroos within *M. (Osphranter)*. Outgroup sampling includes the macropodids, *Lagorchestes (Lagor. hirsutus* and *Lagor. conspicillatus*), Dendrolagini (*Petrogale xanthopus* and *Dendrolagus dorianus*) and *Lagostrophus fasciatus*, in addition to the potoroids, *Aepyprymnus rufescens* and *Potorous tridactylus*.

Mitochondrial DNA was sequenced from DNA previously extracted at University of Adelaide (*A. rufescens*, *D. dorianus*, *P. xanthopus*, *Lagor. conspicillatus*, *W. bicolor* and *M. fuliginosus*) and Australian National University (*M. eugenii* and *M. rufogriseus*). For the remaining new sequences, DNA was extracted from tissue samples. These were provided by collections at Murdoch University (*M. rufus*), The Department of Environment and Conservation, WA (*M. irma*) and The Australian National Wildlife Collection (*M. dorsalis*) or purchased from EcoMeats in Canberra (*M. giganteus*). In addition, we sequenced two nuclear genes (*IRBP* and *ApoB*) from *M. irma* and *W. bicolor* to validate the provenance of our samples, given that their mtDNA placements differed from Meredith et al. [7]. Using our *IRBP* and *ApoB* sequences in place of Meredith et al. 's varied the maximum likelihood bootstrap support on the nuclear data for the placements of *M. irma* and *W. bicolor* by <1.5%. As a default however, we preferentially use the previously available *W. bicolor* sequences, which derive from the same individual as each of the other nuclear loci. For *M. irma* we use our *IRBP* and *ApoB* sequences, which cover a 28 bp sequencing gap and resolve for six ambiguity codes in the previously available sequences.

DNA extraction for *M. rufus* and *M. irma* was carried out at Murdoch University using a Qiagen DNeasy kit (Qiagen Sciences, MD, USA) and at Australian National University for all other taxa, using the salting out method (following [49]). DNA was amplified using standard PCR protocols on a Corbet Research iPAQ thermocycler (NSW, Australia). Primers and amplification conditions are provided in Table S1. All amplicons were sequenced by Macrogen (Seoul, South Korea). Nuclear sequences have been submitted to GenBank for *M. irma* (*IRBP*; JN967008, *ApoB*; JN967009) and for *W. bicolor* (*IRBP*; KC429577, *ApoB*; KC429578). GenBank accession details for the mt sequences are provided in Table S2.

Data matrices

The primary mitochondrial dataset (Mt_{16}) combines the *NADH1*, *NADH2* and *Cytb* protein-coding genes with the 12S and 16S rRNA genes. The sequences were initially aligned in ClustalW2 [50] with penalties of 5 for gap opening and 0.2 for gap extension. Manual adjustments were then made in Se-Al 2.0a [51], where sites with ambiguous homology were excluded, leaving a 5,593 bp mtDNA alignment. An expanded matrix (Mt_{17}) includes *M. dorsalis*; although poor tissue preservation resulted in low DNA yields and the sequence is 72% complete, missing 147 bp of *NADH2*, all of *Cytb* and 269 bp of the rRNA genes. The only mt sequence available on GenBank for *M. antilopinus* was included in a 1,146 bp $Cytb_{18}$ matrix (with the taxa from Mt_{16} and an additional *W. bicolor* sequence) and was sufficient to confidently place *M. antilopinus* as sister to *M. robustus*, in agreement with nuclear genes (Figure 2).

Previous research has demonstrated a requirement to ameliorate nucleotide compositional biases among marsupial mt genomes for phylogenetic inference of ordinal level relationships [23,24]. By contrast, the present focus on closely related genera is relatively shallow. Our composition homogeneity χ^2 testing on Mt_{16} in PAUP* 4.0b10 [52] with uninformative and gapped sites excluded offers little evidence for base compositional non-stationarity among the *Macropus* and *Wallabia* ingroup (protein 1st codons: P = 0.4222, 2nd codons: P = 0.9708, 3rd codons: P = 0.5810, RNA stems: P = 0.9941, RNA loops: P = 0.3400).

We analyse the mt sequences alongside the nuclear dataset of Meredith et al. [7], which includes protein-coding segments from *BRCA1* (exon 11, breast and ovarian cancer susceptibility gene), *ApoB* (exon 26, Apolipoprotein B), *IRBP* (exon 1, interphotoreceptor retinoid binding protein gene), *RAG1* (intronless recombination activating gene-1) and *vWF* (exon 28; vonWillebrand factor gene). Aligning the nuclear sequences followed the procedure described above for the mtDNA. Two 5,988bp matrices were constructed, Nuc_{16}, with taxon sampling matching Mt_{16} and also Nuc_{17}, which further includes *M. antilopinus*. Combined analyses ($MtNuc_{16}$) concatenated the Mt_{16} and Nuc_{16} matrices.

Phylogenetic inference of mitochondrial and nuclear gene trees

Kangaroo phylogeny was inferred under maximum likelihood (ML) and Bayesian inference from the mitochondrial and nuclear sequences separately and concatenated, as well as for the individual nuclear genes. Substitution model categories for each data partition employed the more general of the jModelTest 0.1.1 [53] hLRT or AIC recommendations (Table S3) or the next most general available for each phylogenetic inference program. Substitution was modelled separately among the mt protein-coding codon positions and RNA stem and loop sites.

Our initial efforts to reconstruct kangaroo phylogeny employed Bayesian inference in MrBayes 3.1.2 [54] and ML in RAxML vGUI093 [55]. MrBayes analyses ran two independent sets of two MCMC chains for 6,000,000 (Nuc_{17}, $MtNuc_{16}$) or 4,000,000 (Mt_{16}, Mt_{17}, $Cytb_{18}$ and individual nuclear genes) generations, with trees sampled every 2,500 generations. Burn-ins varied from 500,000 to 1,200,000 generations, and were chosen to ensure that $-\ln L$ had plateaued, clade frequencies had converged between runs and estimated sample sizes for substitution parameters were >200 (using Tracer v1.5 [56]). ML analyses in RAxML carried out 500 full bootstrap replicates. Branch-length multipliers and substitution models were partitioned among protein codons and RNA stems and loops for each of the ML and Bayesian analyses, with $MtNuc_{16}$ further partitioned between mt and nuclear sites.

Support among alternative topologies was further examined with the approximately unbiased (AU) test [57], using the RELL method (100,000 replications) within CONSEL [58]. Site likelihoods employed in CONSEL were inferred in PAUP*, with all substitution parameters and branch-lengths ML optimized separately for each of the protein codon and RNA structural partitions, for each tree hypothesis. Maximum likelihood trees conforming to the alternative Mt_{16} and Nuc_{17} placements of *W. bicolor* and *M. irma* (Figure 2) were identified for each gene in PAUP* with 20 random addition heuristic searches. Support among these individual genes for the alternative placements was compared with SH tests [59], which as pairwise comparisons reduce to equivalency with AU and KH [60] tests. ML bootstrapping (500 replicates) for each gene was also performed in PAUP* with the substitution model parameters estimated in the earlier heuristic searches.

We estimated a mitochondrial timescale for kangaroo evolution using BEAST v.1.6.1 [61] with Mt_{16} partitioned as per the phylogenetic analyses. An uncorrelated relaxed clock model was used with rates among branches distributed according to a lognormal distribution. Note that likelihood ratio tests in PAUP* rejected strict clocks for both the mt and nuclear sequences ($P<0.01$). Four independent runs totalling 40,000,000 MCMC generations ensured estimated sample size values >100 (as estimated in Tracer v1.5) for all node height, prior, posterior, $-\ln L$, tree, and substitution parameters. Chains were sampled every 5,000 th generation after burn-ins of 1,000,000 generations.

Four fossil-based priors were used to calibrate the BEAST analysis. (i) Potoroidae/Macropodidae (15.97–28.4 Ma), with the minimum based on the Early Miocene macropodid, *Ganguroo* [2] and the maximum covering putative Late Oligocene macropodids [62] and potoroids [63]. (ii) Macropodidae (11.6–23 Ma), with the Middle Miocene macropodid, *Wanburoo* [2] providing the mini-

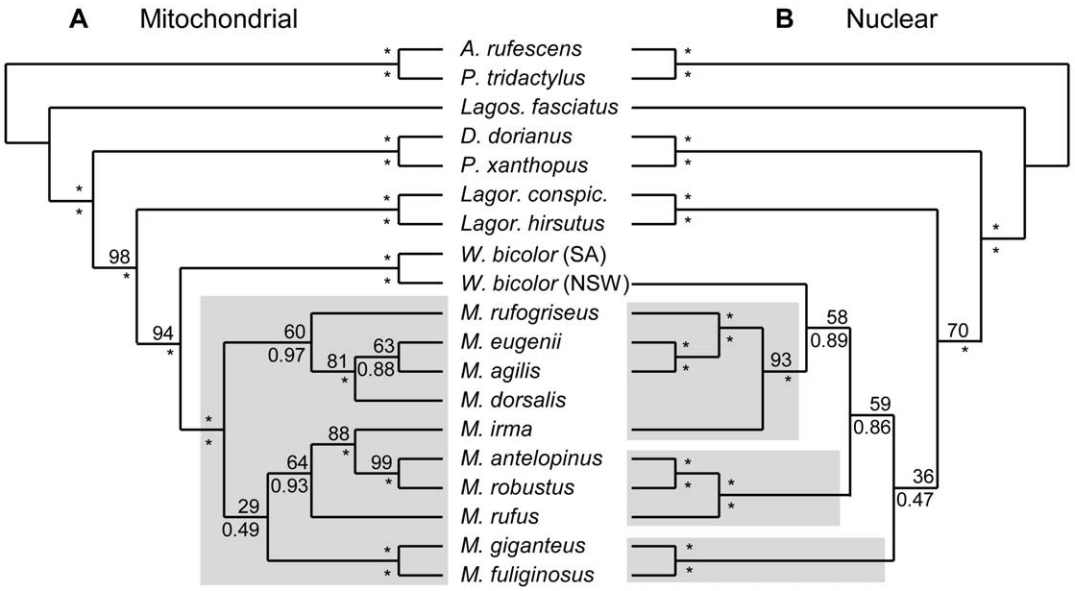


Figure 2. Phylogenetic analysis of kangaroos and wallabies. Maximum likelihood phylogenies inferred from the (A) mitochondrial (Mt_{16}) and (B) nuclear (Nuc_{17}) concatenated datasets, with RAxML bootstrap values (BP_{ML}) above branches and MrBayes Bayesian posterior probabilities (BPP) below branches. The mt placement of *M. dorsalis* is derived from the reduced-length Mt_{17} and the mt placements of *M. antilopinus* and *W. bicolor* (NSW, New South Wales) are derived from the $Cytb_{18}$ alignment. Support for grouping *M. eugenii* and *M. agilis* increases ($BP_{ML}=88$; $BPP=0.98$) for Mt_{16}, which excludes *M. dorsalis*, but increases sequence length. Asterisks indicate full support. Clades including members of *Macropus* are shaded.
doi:10.1371/journal.pone.0057745.g002

mum bound and the maximum acknowledging that earliest Miocene macropods fall outside of the macropodid crown. *Ganguroo* is also a candidate for calibrating Macropodidae, however, its placement within this crown clade is not well resolved [2]. (iii) Dendrolagini (4.46–16.0 Ma), with the minimum based on the Hamilton fauna *Dendrolagus* [64] and the maximum bound recognising that all middle Miocene macropods fall outside the Dendrolagini crown. (v) *Macropus/Lagorchestes* (4.46–16.0 Ma), with the minimum based on Hamilton fauna *Macropus* [2] and the maximum bound recognising that all middle Miocene macropods fall outside of this crown clade.

Palaeontological data do not clearly favour any particular timing within the given bounds for calibrations (i), (iii) and (iv) and hence, flat priors were employed. A normal prior was employed for calibration (ii), in line with the recommendation of Ho and Phillips [65] for when the balance of evidence [2,7,62] suggests crown divergences fall well within the bounds. The normal prior was applied conservatively (90% of prior probability between the bounds).

Partition homogeneity testing

To identify incongruence between partitions we performed the incongruence length difference test [66] in PAUP*. This test however, can be biased in cases where parsimony is statistically inconsistent [67]. To overcome this concern we also perform likelihood-based parametric bootstrap tests. For these we infer one ML score (ML_F) with branch lengths and the models (as shown in Table S3) partitioned across genes, but assuming a single topology (T*) and another ML score (ML_V) for which the topology is allowed to vary across genes. The difference between ML_F and ML_V provides a critical value for testing the null hypothesis that all genes evolved on the same phylogeny. Next we used Seq-Gen 1.3.2 [68] to simulate 200 datasets partitioned into the original mtDNA and five nuclear gene sequence lengths and evolved on topology T* with the original branch lengths and model parameters for each gene. The distribution of ML_F - ML_V differences from the 200 simulated datasets was then compared with the critical value from the original dataset.

Species tree reconstruction

*BEAST analysis [27] within BEAST v1.6.1 employed the multi-species coalescent to infer the species tree underlying the mt and five nuclear gene trees coestimated from $MtNuc_{16}$. The mtDNA was further partitioned into protein codon positions and RNA stems and loops for substitution modelling. Separate mt and nuclear uncorrelated lognormal relaxed clock models were used, with a Yule species tree prior and differential ploidy (autosomal nuclear and haploid mitochondrial). Eight independent runs totalling 80,000,000 MCMC generations ensured estimated sample size values >100 (estimated in Tracer v1.5) for all node height, prior, posterior, $-\ln L$, tree, and substitution parameters. Chains were sampled every 5,000 th generation after burn-ins of between 1,000,000 and 4,000,000 generations. Given our focus on phylogeny rather than dating, we did not calibrate the *BEAST analysis, and so avoid potentially misleading influences of calibration priors on clade posterior probabilities [65]. Instead we provided a nominal mean substitution rate of 1.00 with unspecified time units for the nuclear data and allowed the mt rate to vary relative to this.

Minimizing deep coalescences, MDC [28] trees were inferred under the dynamic programming mode in PhyloNet 2.4 [69] from the mtDNA and five nuclear gene trees (Figure 2A, Figure S1), which were each estimated under ML bootstrapping in PAUP*. We collapsed branches that received <50% bootstrap support in these source trees to reduce the influence of stochastic artefacts among the individual genes on MDC tree building.

Bayesian concordance analysis within BUCKy [29] was run on 500 bootstrap replicate trees inferred in RAxML, for each of the six loci. Bootstrap distributions typically reflect stochastic variation in the gene tree estimates more closely than do Bayesian posterior distributions [24,70]. Otherwise, BUCKy analyses employed default parameters, except where stated.

Coalescent simulation

In order to better understand whether ILS could plausibly account for incongruence among gene trees we simulated the evolution of the five nuclear genes and the mtDNA under a coalescent process in MCcoal [71] within BPP 2.1 [72]. Two alternative guide trees were used for MCcoal, the combined data and nuclear-only *BEAST species trees. In the former case *M. irma* was excluded from the *BEAST analysis because of the concern that its mtDNA derives from introgression (see Discussion), which violates the assumptions of *BEAST. Instead, *M. irma* was grafted onto the tree – its temporal placement along the stem lineage from the other *Notamacropus* members was scaled in proportion to the nuclear-only *BEAST tree. For comparability the combined data and nuclear-only guide trees were both scaled to a root height of 20 Ma, closely matching both the mt BEAST estimate (21.3 Ma) and Meredith et al. 's [7] estimates from the concatenated nuclear genes (ave. BEAST estimate, 20.0 Ma).

The model that MCcoal simulates under (JC+Γ) is less complex than the models selected in jModelTest for each locus. Therefore we used a two-step simulation process (illustrated in Figure 3A). First, MCcoal was run on the species guide tree to provide simulated coalescent trees. Gene sequences were then simulated on these coalescent trees under their respective substitution models (Table S3) in Seq-Gen. All model parameters were estimated from the original data and the simulations maintained the aligned sequence length for the mtDNA and each nuclear gene.

MCcoal requires a population dynamics parameter $\theta = 4N_e\mu$ ($2N_e\mu$ for mtDNA), where N_e is the effective population size and μ is the mutation rate per site per generation. Mutation rates per site per year were obtained by scaling PAUP* ML treelengths for each locus to the *BEAST timetree length. Generation time across macropods is not well studied, but we used an average of 7 years in consideration of life history data from most macropodid species [73]. The influence of effective population size was evaluated with N_e varied from 1,000 to 1,000,000.

Results

Phylogenetic inference from separate mitochondrial and nuclear sequences

Our analyses of the mt and nuclear sequences agree on grouping *Macropus* and *Wallabia* to the exclusion of the consecutive outgroups, *Lagorchestes*, Dendrolagini, *Lagostrophus* and Potoroidae (Figure 2). The inclusion of the mtDNA greatly enhanced phylogenetic resolution. Whereas four clades received 36–70% ML bootstrap support on the nuclear data alone, on the combined data all but one clade received ≥90% ML bootstrap support (Figure 4A). Only the relative affinities of the three *Macropus* subgenera (*Macropus*, *Notamacropus* and *Osphranter*) remained poorly resolved. However, the combined result hides conflict between the mt (Figure 2A) and nuclear (Figure 2B) trees for the placements of *M. irma* and *W. bicolor*.

The nuclear data favours *M. irma* and *W. bicolor* as consecutive sister groups to the wallabies we refer to as core-*Notamacropus*, which here includes *M. rufogriseus*, *M. eugenii*, *M. agilis* and *M.*

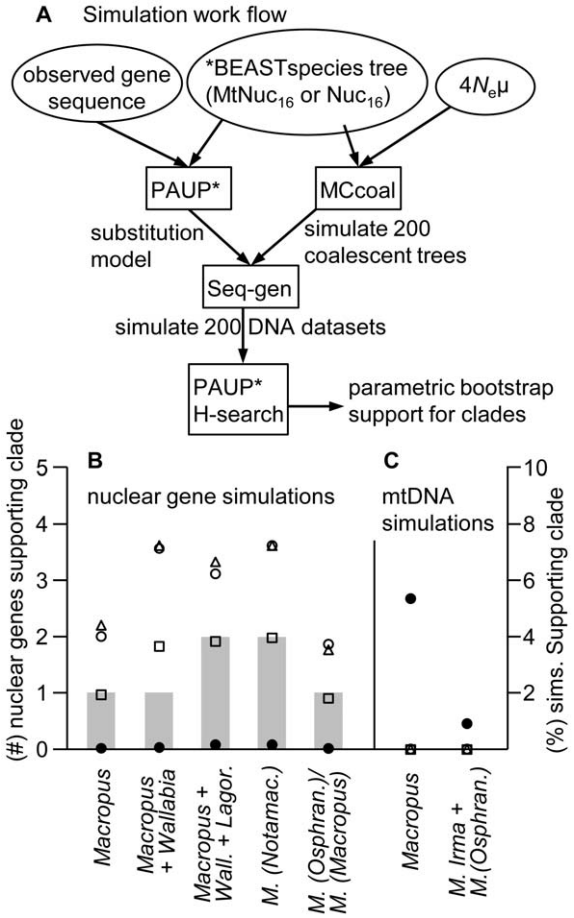


Figure 3. Macropodid clade support from datasets simulated under coalescence. (A) Simulation workflow. (B) Mean number of the five nuclear genes supporting each clade in maximum likelihood analyses of 200 simulations of the combined data *BEAST species tree for N_e values of 1,000 (triangle), 10,000 (open circle), 100,000 (square) and 1,000,000 (filled circle). For comparison, the grey bars show the number of genes supporting each clade on the observed data. (C) Percentage of ML analyses supporting each clade among 200 mtDNA simulations on the nuclear-only *BEAST species tree for N_e values set to mitochondrial equivalency for the same populations (one quarter of the corresponding nuclear values). Abbreviations: *Lagor.*; *Lagorchestes*, *Wall.*; *Wallabia*, *M. (Notamac.)*; *M. (Notamacropus)*, *M. (Osphran.)*; *M. (Osphranter)*.
doi:10.1371/journal.pone.0057745.g003

dorsalis. Placing *M. irma* with core-*Notamacropus* receives high ML bootstrap support (BP_{ML}, 93%) and Bayesian posterior probability (BPP, 1.00). Further expanding this clade to include *Wallabia* is only supported modestly (BP_{ML} 58%, BPP 0.89). These results closely mirror Meredith et al. [7]. Our mitochondrial trees strongly conflict with these placements, instead favouring *M. irma* as sister to the wallaroos (*M. robustus*, *M. antilopinus*) (BP_{ML} 88%, BPP 1.00) and placing *W. bicolor* outside a monophyletic *Macropus* (BP_{ML} 100%, BPP 1.00).

Maximum likelihood AU testing reveals strong incongruence between the nuclear and mt data for the placements of both *W. bicolor* and *M. irma*. Table 1A shows that for Mt_{16} the favoured nuclear placement for *W. bicolor* as sister to the subgenus *M. (Notamacropus)* is rejected at P = 0.008 and reciprocally, AU testing on Nuc_{17} rejects the favoured mt placement for *W. bicolor* as sister to all *Macropus* at P = 0.011. Similarly for *M. irma* (Table 1B), the favoured mt placement as sister to *M. robustus* is rejected with the nuclear data (P<0.001) and reciprocally, the favoured nuclear placement as sister to core-*Notamacropus* is rejected with the mtDNA (P = 0.031).

Turning to the individual nuclear genes, support for the favoured mt versus nuclear placements reveals distinctly different patterns for *M. irma* and *W. bicolor* (Table 2). All of the nuclear genes favour an *M. irma* relationship with *M. (Notamacropus)* over the mt relationship with *M. (Osphranter)* – except *BRCA1*, for which both of these relationships were equally likely. In contrast, the overall nuclear placement of *W. bicolor* as sister to *M. (Notamacropus)* is only favoured over the mt placement by *BRCA1* and *vWF*. Another gene (*IRBP*) instead favours the mt placement of *W. bicolor* outside *Macropus*, while the ML analyses for *ApoB* and *RAG1* find the nuclear and mt hypotheses for *W. bicolor* affinities to be equally likely.

Our analyses of Mt_{17} show that *M. dorsalis* groups with *M. eugenii* and *M. agilis* (BP_{ML} = 81%, BPP = 1.00; Figure 2A), with the latter two wallabies favoured as sister taxa (BP_{ML} = 63%, BPP = 0.88). AU testing (Table 1E) echoes these results, favouring *M. dorsalis* as sister to M. *eugenii* and *M. agilis*, although with other placements of *M. dorsalis* within core-*Notamacropus* rejected only at modest significance levels (P values from 0.079–0.334).

Kangaroo species tree inference

Partition homogeneity testing performed in PAUP* identified significant incongruence between the mt and nuclear datasets (P = 0.027) and between the five nuclear genes (P = 0.003). Nuclear gene trees are shown in Figure S1. These parsimony-based results are in agreement with the likelihood-based parametric bootstrap test. For the latter, the improvement in likelihood of partitioning over concatenation for the mt and nuclear sequences ($MtNuc_{16}$) and among the five nuclear genes (Nuc_{16}) was 30.44 and 82.23 – ln*L* units respectively. In both cases these critical values fall higher than the distribution of likelihood improvements from partitioning for each of the 200 simulated datasets, therefore rejecting homogeneity at P<0.005.

We employed four approaches to inferring the kangaroo species tree from the mtDNA and five nuclear genes. First the data were concatenated, with substitution models and relative rates partitioned between mt and nuclear sequences and within these, between the protein-coding codon positions and RNA stems and loops. The concatenated $MtNuc_{16}$ ML and Bayesian analyses provide a well resolved tree (Figure 4A) that combines the mitochondrial placement of *Wallabia* as sister to *Macropus*, with relationships among the *Macropus* species following the nuclear tree.

The second approach using $MtNuc_{16}$ applied the multi-species coalescent within *BEAST to allow for ILS among the mtDNA and the five nuclear genes. As shown in Figure 4B *BEAST reconstructed the same topology as the concatenated analysis, except with *M. irma* as sister to *M. (Osphranter)* rather than core-*Notamacropus*. We also ran the *BEAST analysis without the putative introgressive hybrid, *M. irma*. *Macropus* monophyly was retained (BPP = 0.99), although among the subgenera, *M. (Osphanter)* grouped (at BPP = 0.86) with *M. (Macropus)*, instead of with *M. (Notamacropus)*.

Among the two other species tree approaches, BUCKy carries out Bayesian concordance analysis, which requires a prior level of discordance (α) to be assigned. We ran separate analyses with α at 0.5, 1, 5 and 10, which provide for a range of prior expectations for the 6 loci representing one or two distinct trees up to representing five or six distinct trees. The same concordance tree was recovered under each of these levels and shares the same topology with both the concatenated analysis (Figure 4A) and the

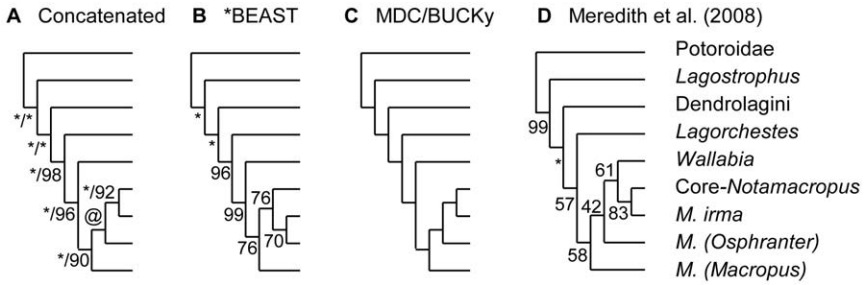


Figure 4. Macropodid species tree estimates from the combined mitochondrial and nuclear sequences ($MtNuc_{16}$). (A) concatenated sequences, showing BPP/BP_{ML} (@ = 0.89/59), (B) *BEAST partitioned between the mtDNA and five nuclear genes, showing BPP values. (C) both MDC and BUCKy, which recovered the same tree from the mt and five nuclear gene trees. (D) Meredith et al. (2008) with BP_{ML} values included for comparison. Several supraspecific clades that were identical across all reconstructions were collapsed for visualization convenience. Relationships within each of the collapsed clades were as inferred in Figure 2. Asterisks indicate full BPP or BP_{ML} support.
doi:10.1371/journal.pone.0057745.g004

MDC tree (Figure 4C). On the nuclear data alone, each of the species tree methods followed the concatenated nuclear tree (Figure 2B), in placing *W. bicolor* with *M. (Notamacropus)*. However, all methods used to combine the mt and nuclear sequences support *Macropus* monophyly.

Table 1. Approximately unbiased (AU) test results.

	Mitochondrial genes		Nuclear genes	
	–lnL	P-value	*–lnL*	P-value
(A) Placement of *Wallabia*				
1. Sister to *Macropus*	25,140.45	best	+20.06	0.011
2. Sister to *M. (Notamacropus)*	+31.505	0.008	13,107.34	best
(B) Placement of *M. irma*				
1. Sister to *M. (Notamacropus)*	+16.90	0.031	13,107.34	best
2. Sister to *M. robustus*	25,140.45	best	+90.16	<0.001
3. Sister to *M. (Osphranter)*	+6.21	0.157	+24.14	0.063
(C) Macropus subgenera relative affinities^				
1. *M. (Notamacropus)+M. (Osphranter)*	+1.39	0.317	13,107.34	best
2. *M. (Osphranter)+M. (Macropus)*	25,140.45	best	+12.95	0.052
3. *M. (Macropus)+M. (Notamacropus)*	+0.62	0.455	+11.39	0.161
(D) Placement of *Lagorchestes*				
1. Sister to *Wallabia+Macropus*	25,140.45	best	13,107.34	best
2. Sister to *Wallabia*	+8.59	0.038	+25.82	0.009
3. With *Macropus*	+7.55	0.103	+0.94	0.462
(E) Placement of *M. dorsalis* (5 highest)				
1. Sister to (*M. agilis+M. eugenii*)	19,771.52	best		
2. Sister to *M. agilis*	+4.47	0.334		
3. Sister to *M. eugenii*	+7.17	0.079		
4. Sister to all other *M. (Notamacropus)*[#]	+10.25	0.134		
5. Sister to all other *Macropus*	+13.35	0.098		

Nuclear sequences are partitioned into protein codon positions and mitochondrial sequences are partitioned into protein codon positions and RNA stems and loops. Comparisons (A) – (D) employ Mt_{16} and Nuc_{17}. Comparison (E) employs Mt_{17}.
^Allowing *W. bicolor* and *M. irma* to float unconstrained on the tree
[#]Not including *M. irma*, which is favoured as sister to *M. robustus* on the mt data.
doi:10.1371/journal.pone.0057745.t001

Table 2. Individual nuclear gene –ln*L* differences and SH test results.

	IRBP	*vWF*	*ApoB*	*BRCA1*	*RAG1*
(A) Placement of *W. bicolor*					
1. Outside monophyletic *Macropus*	**2765.74**	+10.18	1763.63	+17.45	1008.98
2. With *M. (Notamacropus)*	+7.95	**2057.71**	1763.63	**5488.65**	1008.98
	P = 0.117	P = 0.070	—	P = 0.055	—
(B) Placement of *M. irma*					
1. With *M. (Osphranter)*	+14.10	+10.18	+2.85	5504.78	+0.05
2. With *M. (Notamacropus)*	**2763.58**	**2057.71**	**1763.82**	5504.78	**1009.18**
	P = 0.045	P = 0.042	P = 0.226	—	P = 0.646

ML placements in bold.
To ensure relevance of the individual gene results to the overall nuclear phylogeny, the relative positions of the outgroups and placements within *M. (Macropus)*, *M. (Osphranter)* and core-*Notamacropus* were fixed (see Figure 2).
doi:10.1371/journal.pone.0057745.t002

Coalescent simulations

There is agreement between the mtDNA and the majority of nuclear genes on deep and shallow clades with stem lineages ≥1.5 Ma on the combined species tree. Almost all of the discordance arises among a tight cluster of six consecutive divergences covering 4.2–4.9 Ma (see Table 3B,C for divergences) from the Macropodini/Dendrolagini divergence up to the *M. (Notamacropus)* crown divergence. In fact most gene discordance within this cluster (Figure S1) and the lowest bootstrap support (<60%, Figure 2B) involves the sequential divergence of *Lagorchestes*, *Wallabia* and each of the three *Macropus* subgenera covering just 2.0–2.4 Ma (Table 3; Figure 1B, grey strip). Partition homogeneity testing and AU testing (Tables 1,2) suggest that the extent of the incongruence cannot be explained by stochastic error alone. However, another explanation, incomplete lineage sorting is also consistent with the association between short stem length and incongruence.

Simulating the coalescent process over the combined data species tree allows inference of whether ILS provides a plausible explanation for the strong incongruence between the nuclear genes. In Figure 3B the grey bars show the number of nuclear genes supporting each of five short-stem-lineage clades within the diversification cluster. Phylogenetic analyses of the simulated datasets provide estimates for the probability of each gene supporting a given clade for alternative values of N_e. The sum of these values over the five genes is the mean expectation for the number of genes supporting each clade. Coalescent time is very short with $N_e = 1{,}000$ and as a result the coalescent simulations overestimate the number of genes supporting each of the five clades. Under this scenario ILS appears to be a poor explanation for the lack of concordance among the nuclear genes. Increasing N_e to 10,000 makes little difference.

The greater potential for ILS with $N_e = 100{,}000$ provides for a remarkably close match to the observed gene support among the clades (Figure 3B). The sum of squares difference between expected and observed support for the five clades decreases from 13.25 and 12.20 for $N_e = 1{,}000$ and 10,000 respectively, to 0.76 for $N_e = 100{,}000$. This improvement does not continue with N_e being further increased to 1,000,000 (sum of squares = 10.16). Coalescent simulations with such high N_e overestimate the extent of incongruence, with no genes expected to support any of the clades in 80% of the simulated datasets.

MCcoal simulations were also run with the nuclear-only species tree providing the guide phylogeny. Here the aim was to determine whether ILS can potentially explain the mtDNA supporting *Macropus* monophyly or placing *M. irma* distantly from other members of *M. (Notamacropus)*. Effective population size was set to mitochondrial equivalency for the same populations (one quarter of the corresponding nuclear values). All simulated mtDNA sequences favoured the *Wallabia*/*M. (Notamacropus)* and *M. (Notamacropus)* groupings for each of the three lower N_e values (250, 2,500, 25,000). At $N_e = 250{,}000$ these fell to 46% and 66% respectively, although support among the simulated datasets for *Macropus* monophyly remained at 0% and only increased to 1.5% for *M. irma* falling within any grouping outside core-*Notomacropus* that is at least as shallow as *M. (Osphranter)* – where *M. irma* was placed on the observed mtDNA. Hence, the mitochondrial placements of *Wallabia* and *M. irma* are difficult to reconcile with ILS.

Discussion

Mitochondrial sequences provide confirmation and incongruence

Molecular studies have consistently shown that *Lagorchestes*, *Wallabia* and the *Macropus* subgenera *M. (Macropus)*, *M. (Osphranter)* and *M. (Notamacropus)* diverged from each other in rapid succession. We estimate that together, their consecutive divergences cover a temporal window of little more than 2 million years (Figure 1B, Table 3), in agreement with Meredith et al. [7]. The short internal branches may provide low phylogenetic resolution due to stochasticity associated with few substitutions along branches and conflicting signals attributable to incomplete lineage sorting (ILS) among genes. This expectation was borne out for the nuclear dataset, with all relationships among the five groups poorly resolved (Figure 2B, BP_{ML} from 36–59%).

Adding the mitochondrial (mt) sequences to the nuclear dataset substantially enhances resolution (Figure 4A). All groupings on the tree receive ≥90% BP_{ML} and are consistent with the supertree (Figure 1A) modified from Cardillo et al. [6], with the exception of a near-trichotomy among the *Macropus* subgenera. As strong as these results are, there is significant incongruence between the five individual nuclear genes (Figure S1). Moreover, mtDNA discordance with the combined nuclear sequences (see Figure 2) necessitates caution, especially for inferrig the affinities of the swamp wallaby (*W. bicolor*) and the black-gloved wallaby (*M. irma*).

The relationship of *Wallabia* to *Macropus*

All concatenated and species tree analyses of the combined mtDNA and nuclear genes recover *W. bicolor* as sister to *Macropus*

Table 3. Macropodoid divergence time estimates in millions of years before present.

Clade	(A) mtDNA		(B) NucDNA	(C) $MtNuc_{16}$ *BEAST
	Median	95%HPD	Meredith et al. [7]	Species tree
1. Macropodoidea	21.3	(16.0–26.8)	20.0	20.0
2. Potoroidae	16.5	(10.8–23.1)	16.4	17.0
3. Macropodidae	16.2	(12.2–20.6)	17.7	14.4
4. Dendrolagini/Macropodini	11.0	(8.2–14.0)	10.7	7.6
5. *Lagorchestes/Macropus/Wallabia*	9.7	(7.8–12.6)	8.8	6.6
6. *Macropus/Wallabia*	8.9	(6.6–11.5)	7.3	5.3
7. *Macropus*	7.6	(5.5–9.8)	—	4.8
8. *M. (Macropus)/M. (Osphranter)*	7.3	(5.3–9.5)	—	4.4
9. *M. irma/M. robustus*	5.9	(4.3–8.0)	—	
10. *Macropus/Wallabia* except *M. (Macropus)*	—	—	6.8	
11. *M. (Notamacropus)/Wallabia*	—	—	6.7	
12. *M. (Notamacropus)*	—	—	5.8	3.4

(A) BEAST analysis of Mt_{16}, (B) average of four BEAST analyses on the five nuclear gene concatenate from Meredith et al. [7] and (C) *BEAST species tree analysis of $MtNuc_{16}$.

doi:10.1371/journal.pone.0057745.t003

(Figure 4), consistent with recent morphological analyses [2,12,13]. If we accept this relationship, it implies that the nuclear (Nuc_{17}) placement of *W. bicolor* within *Macropus* is an artefact, potentially of ILS among the individual genes. This interpretation is consistent with extreme incongruence among the nuclear genes on both MP and ML-based partition homogeneity tests ($P \leq 0.005$). Moreover, support for the overall nuclear placement with *M. (Notamacropus)* derives from only *BRCA1* and *vWF*, while *IRBP* favours the mtDNA placement as sister to *Macropus* (Table 2A).

We examined the incongruence further with coalescent simulations and show that with small populations the models overestimate support among the nuclear genes for clades within the diversification cluster. However, simulating deeper ILS consistent with $N_e = 100{,}000$ on the species tree closely matches observed support for these clades (Figure 3B). Neaves et al. [74] recently estimated N_e of similar magnitude for *M. fuliginosus*. If indeed *W. bicolor* does fall outside *Macropus*, then it might appear anomalous that only *IRBP* among the five nuclear genes favours *Macropus* monophyly. This result however, matches the coalescent simulations on the species tree under the best fitting N_e; for 74% of the simulated datasets *Macropus* monophyly was recovered for only one (or none) of the nuclear genes. In contrast, none of the corresponding mtDNA coalescent simulations (Figure 3C) on the nuclear-only species tree favour *Macropus* monophyly. Hence, the observed mtDNA support for *Macropus* monophyly is unlikely to be an artefact of incomplete mitochondrial lineage sorting.

Strong mtDNA support for *Macropus* monophyly and apparently extensive ILS among nuclear loci within the diversification cluster around the base of *Macropus* caution against Meredith et al. 's [7] recommendation to subsume *Wallabia* within *Macropus*. Nevertheless, statistical support among our species tree analyses for excluding *W. bicolor* from *Macropus* is not conclusive. Indeed, coalescent simulations (Figure S2) on the species tree suggest that 30 or more nuclear loci may be required to confidently resolve relationships among the diversification cluster. However, considering our present phylogenetic results alongside the distinct browsing ecology and associated morphology/behaviour of *W. bicolor* [2,4,14,75] and its unique $2n = 10(♀)/11(♂)$ karyotype, we believe that *Wallabia* currently warrants separate generic status.

Deep mitochondrial introgression in *Macropus irma*

Incongruence between the mt and nuclear placements for the black-gloved wallaby (*M. irma*) differs in several respects from that concerning *W. bicolor*. It is the nuclear placement of *M. irma* with *M. (Notamacropus)* that concurs with morphology and none of the nuclear genes prefer the mt placement with *M. (Osphranter)* (Table 2). A further point of difference is that the concatenated ($MtNuc_{16}$) analysis and two species tree reconstructions (MDC and BUCKy) support the nuclear placement for *M. irma*.

The mt placement of *M. irma* with wallaroos is difficult to reconcile with incomplete mitochondrial lineage sorting, being nested within *M. (Osphranter)* and because the path-length to its species tree position at the base of *M. (Notamacropus)* is 3–8 million years. Mitochondrial introgression may provide a more plausible explanation. Methods are being developed to directly test for introgressive hybridization, although these are not feasible without many loci or strong priors on the probability of hybridization [33,34]. Nevertheless, our coalescent simulations (Figure 3C) further suggest that the aberrant mt placement of *M. irma* is not an artefact of incomplete lineage sorting. Even with the most extreme deep coalescence (ILS) scenario fewer than 2% of the simulated mtDNA datasets favoured *M. irma* falling outside its species tree grouping and into clades at least as shallow as *M. (Osphranter)*. Hence, on the available evidence the more likely explanation is that *M. irma* obtained its mt genome from introgressive hybridization with an ancestor of the wallaroos, the deepest such event yet hypothesised among marsupials.

Previous examples of hybridization or introgression among wild macropodids (from *Petrogale* and *Macropus*) involve closely related parent species [32,76]. The absence of evidence for introgression among more distantly related macropodids may reflect sparse sampling. Certainly, captive-bred hybrids include more distantly related *Macropus* crosses as well as *Macropus*×*Wallabia* and reports of *Macropus*×*Thylogale* [77,78]. The results of Neaves et al. [33] may also be relevant here. Despite finding evidence for introgression in 17 of 223 grey kangaroos in the *M. giganteus*/*M. fuliginosus* sympatry zone, no F1 individuals were identified. The authors interpreted this result as suggesting a role for selection in

accelerating introgression of some loci into the gene pool at well beyond the actual rate of hybridization.

Adaptive introgression of nuclear loci, as proposed for the grey kangaroos has also been suggested for mitochondria in wild goats [79] and Hares [80]. It is also possible that introgression of *M. (Osphranter)* genes into *M. irma* has adaptive significance and may not be limited to mtDNA. Most notably, Milne and O'Higgins [81] found that skull shape principle components that were correlated with vegetation cropping and mastication grouped *M. irma* within *M. (Osphranter)*, thus matching the mtDNA and further suggesting adaptive convergence. This is consistent with Christensen [82] regarding *M. irma* as somewhat transitional in diet and locomotion between *M. (Notamacropus)* wallabies and the larger wallaroos and kangaroos of *M. (Osphranter)* and *M. (Macropus)*. In particular, *M. irma* favours more open habitats [82] and may rely on grazing and poorer quality plant material more than most of its wallaby relatives [83].

If *W. bicolor* is sister to *Macropus*, then sharing deep gene coalescences with the more ecologically similar *M. (Notamacropus)* wallabies rather than with the larger, grazing *M. (Macropus)* and *M. (Osphranter)* might also point towards adaptive significance. The influence of selection on patterns of ILS and introgression is poorly understood at present and clarification of our intimations concerning *M. irma* and *W. bicolor* requires more thorough genomic sampling and analysis of functional correlations.

Variation among species tree inferences

Each of the $MtNuc_{16}$ concatenated and species tree analyses recovered the same kangaroo phylogeny (Figure 4A–C), except for *BEAST placing *M. irma* as sister to *M. (Osphranter)*, close to the mtDNA relationship. One shortfall in the present usage of *BEAST is that species were represented by only one individual and therefore N_e could only be estimated for internal branches. It is not clear however that this should have any specific impact on the placement of *M. irma*. Indeed, running the analysis without the sequence data indicates that the placement of this taxon was not attributable to any aspect of the tree prior. Furthermore, using single individuals did not promote any other topological differences from the concatenated tree. Instead, the *BEAST result is consistent with mitochondrial introgression being the source of the incongruence concerning *M. irma*. This violates the assumption of *BEAST that ILS is the only source of incongruence. In contrast, another species tree method, BUCKy does not assume any particular source of incongruence and recovered the expected placement of *M. irma* as sister to core-*Notamacropus*.

It is interesting that MDC recovered the expected placement for *M. irma*, despite also assuming that incongruence derives solely from ILS. The explanation may lie in MDC being a consensus method, such that no matter how strong the signal from the mtDNA, it will be overwhelmed by consistent signal among multiple independent nuclear loci. *BEAST is not a consensus method. Instead it models the multi-species coalescent to co-infer gene trees embedded in a species tree, which as Heled and Drummond [27] explain, effectively provides a "reverse auction", where the lowest bidder can set the limit. Consequently, *BEAST is a powerful tool for identifying true species relationships and divergence times when incongruence derives from ILS. However, the reverse auction might often leave *BEAST less robust than consensus methods to introgression or paralogy. Nevertheless, we believe that multi-species coalescent methods such as *BEAST are an important advance for phylogenetics. Allowing for limited post-speciation gene flow [84] will improve their reliability and will provide a valuable test for distinguishing incongruence from ILS.

Our partition homogeneity testing and analysis of coalescent simulations on the species tree (Figure 3B) are consistent with widespread ILS across the cluster of six rapidly diverged lineages, Dendrolagini, *Lagorchestes*, *Wallabia* and the three *Macropus* subgenera. The apparent introgression of *M. (Osphranter)* mtDNA into *M. irma* is the sole instance of significant mitochondrial discord with the species tree. These patterns of incongruence, although too few to draw strong conclusions on, nevertheless fit the expectations set out earlier. Specifically, that ILS will be more common among nuclear loci, consistent with longer coalescence times than for mtDNA, which in turn will be more susceptible to introgressive selective sweeps associated with fitness differences across populations, promoted by Muller's ratchet.

Overall our results suggest that sampling multiple mt genes is well suited to providing a first estimate for species-level phylogenies among marsupials and in combination with the five nuclear genes, substantially enhances phylogenetic resolution. Moreover, concerns that mt signal from three-fold as many parsimony-informative characters would swamp nuclear signal were unfounded. Combined analyses generally favoured the nuclear placements for *M. irma* and among the *Macropus* subgenera over their mt placements. The combined data only favoured the mt placement of *Wallabia*, for which the nuclear loci themselves were incongruent and contradicted morphology. However, larger scale nuclear genomic sampling will ultimately provide a more comprehensive understanding of evolutionary history, including for whether selective advantages contribute to patterns of ILS and introgression in *W. bicolor* and *M. irma*.

Supporting Information

Figure S1 Individual nuclear gene phylogenies. (A) *BRCA1*, (B) *IRBP*, (C) *ApoB*, (D) *vWF* and (E) *RAG1*. MrBayes 3.1.2 Bayesian posterior probabilities (above 0.5) and PAUP* 4.0b10 maximum likelihood bootstrap percentages (>50) are shown above and below branches respectively. Analyses were carried out as per the primary analysis for Nuc_{17}.
(PDF)

Figure S2 Maximum likelihood bootstrap identification of the number of genes required to resolve macropodid phylogeny. (A) for *Macropus* monophyly and (B) for the *M. (Macropus)-M. (Osphranter)* grouping. Simulated gene sequences (1,000 bp) were added in increments of five. Ten independent runs were continued until sufficient sequences were added for ML_{BP} >95%. Seven of 10 simulations reached 95% ML_{BP} with 20 genes for *Macropus* and 35 genes for *M. (Macropus)-M. (Osphranter)*.
(PDF)

Table S1 List of primers and conditions used for amplifying macropodoid DNA.
(PDF)

Table S2 GenBank accession numbers for the sequences included in the mitochondrial data matrices.
(PDF)

Table S3 jModelTest selections for mitochondrial and nuclear data partitions.
(PDF)

Acknowledgments

We thank Alan Cooper, Jenny Graves, Leo Joseph, Peter Spencer and Adrian Wayne for providing tissue samples and DNA extractions as well as Daniel Hoops and Ric Dawson for permission to use photos in Figure 1. Mark Eldridge and Rachel O'Neill supplied provenance information for

previously available wallaby mitochondrial sequences. Axel Janke and three anonymous reviewers provided valuable comments on the manuscript.

Author Contributions

Conceived and designed the experiments: MJP. Performed the experiments: RCP DH GCG. Analyzed the data: MJP. Contributed reagents/materials/analysis tools: MB. Wrote the paper: MJP MB DH RCP.

References

1. Raven HC, Gregory WK (1946) Adaptive branching of the kangaroo family in relation to habitat. Am Mus Novit 1309: 1–15.
2. Prideaux GJ, Warburton NM (2010). An osteology–based appraisal of the phylogeny and evolution of kangaroos and wallabies (Macropodidae: Marsupialia). Zool J Linnean Soc 159: 954–987.
3. Strahan R (1995) The Mammals of Australia. Sydney: Reed New Holland. 756 p.
4. Sanson GD (1989) Morphological adaptations of teeth to diets and feeding in the Macropodoidea. In: Grigg G, Jarman P, Hume I, editors. Kangaroos, Wallabies and Rat-kangaroos. Chipping Norton, NSW: Surrey Beatty & Sons. pp. 151–168.
5. Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90: 2648–2648.
6. Cardillo M, Bininda–Emonds ORP, Boakes E, Purvis A (2004) A species-level phylogenetic supertree of marsupials. J Zool (Lond) 264: 11–31.
7. Meredith RW, Westerman M, Springer MS (2008) A phylogeny and timescale for the living genera of kangaroos and kin (Macropodiformes: Marsupialia) based on nuclear DNA sequences. Aust J Zool 56: 395–410.
8. Dawson L, Flannery T (1985) Taxonomic and phylogenetic status of living and fossil kangaroos and wallabies of the genus *Macropus* Shaw (Macropodidae: Marsupialia), with a new subgeneric name for the larger wallabies. Aust J Zool 33: 473–498.
9. Tate GHH (1948) Results of the Archibald expeditions. No. 59. Studies on the anatomy and phylogeny of the Macropodidae (Marsupialia). Bull Amer Mus Nat Hist 91: 233–251.
10. Ride WDL (1957) *Protemnodon parma* (Waterhouse) and the classification of related wallabies (*Protemnodon*, *Thylogale*, and *Setonix*). Proc Zoo Soc Lond 128: 327–346.
11. Archer M (1984) Origins and early radiations of marsupials. In: Archer M, Clayton G, editors. Vertebrate Zoogeography and Evolution in Australasia. Carlisle: Hesperian Press. pp. 385–425.
12. Flannery T (1989) Phylogeny of the Macropodoidea; a study in convergence. In: Grigg G, Jarman P, Hume I, editors. Kangaroos, Wallabies and Rat–kangaroos. Sydney: Surrey-Beatty and Sons. pp. 1–46.
13. Kear BP (2002) Phylogenetic implications of macropodid (Marsupialia: Macropodoidea) posteranial remains from Miocene deposits of Riversleigh, Northwestern Queensland. Alcheringa 26: 299–318.
14. Ganslosser U (1992) Behavioral data support the currently proposed phylogeny of Macropodoidea (Marsupialia). Aust Mammal 15: 89–104.
15. Richardson BJ, McDermid EM (1978) A comparison of genetic relationships within the Macropodidae as determined from allozyme, cytological and immunulogical data. Aust Mammal 2: 43–52.
16. Burk A, Springer MS (2000) Intergeneric relationships among Macropodoidea (Metatheria: Diprotodontia) and the chronicle of kangaroo evolution. J Mamm Evol 7: 213–237.
17. Westerman M, Burk A, Amrine HM, Prideaux GJ, Case JA, et al. (2002) Molecular evidence for the last survivor of an ancient kangaroo lineage. J Mamm Evol 9: 209–223.
18. Kirsch JAW (1977) The comparative serology of Marsupialia, and a classification of marsupials. Aust J Zool, Suppl Series 52: 1–152.
19. Baverstock PR, Richardson BJ, Birrell J, Krieg M (1989) Albumin immunologic relationships of the Macropodidae (Marsupialia). Syst Zool 38: 38–50.
20. Kirsch JAW, Lapointe F-J, Foeste A (1995) Resolution of portions of the kangaroo phylogeny (Marsupialia: Macropodidae) using DNA hybridization. Biol J Linnean Soc 55: 309–328.
21. Bulazel KV, Ferreri GC, Eldridge MDB, O'Neill RJ (2007) Species–specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. Genome Biol 8: R170.
22. Retief JD, Winkfein RJ, Dixon GH, Adroer R, Queralt R, et al. (1995) Molecular phylogeny and evolution of marsupial protamine P1 genes. J Mol Evol 37: 426–434.
23. Phillips MJ, McLenachan PA, Down C, Gibb GC, Penny D (2006) Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. Syst Biol 55: 122–137.
24. Phillips MJ, Pratt RC (2008) Family-level relationships among the Australasian marsupial "herbivores" (Diprotodontia : Koala, wombats, kangaroos and possums). Mol Phylogenet Evol 46: 594–605.
25. Edwards SV (2008) Is a new and general theory of molecular systematics emerging? Evolution 63: 1–19.
26. Knowles LL, Kubatko LS (2010) Estimating species trees: an introduction to concepts and models. In: Knowles LL, Kubatko LS, editors. Estimating Species Trees: Practical and Theoretical Aspects. Hoboken, New Jersey: Wiley-Blackwell. pp. 1–14.
27. Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Mol Biol Evol 27: 570–580.
28. Maddison WP (1997) Gene trees in species trees. Syst Biol 46: 523–546.
29. Ané C, Larget B, Baum DA, Smith SD, Rokas A (2007) Bayesian estimation of concordance among gene trees. Mol Biol Evol 24: 412–426.
30. Briscoe D, Calaby JH, Close RL, Maynes GM, Murtagh CE, et al. (1982) Isolation, introgression and genetic variation in rock wallabies. In: Groves RH, Ride WDL editors. Species at Risk: Research in Australia. Canberra: Australian Academy of Science. pp. 73–87.
31. Eldridge MDB, Close RL (1992) Taxonomy of rock wallabies, *Petrogale* (Marsupialia: Macropodidae). I. A revision of the eastern *Petrogale* with the description of three new species. Aust J Zool 40: 605–625.
32. Neaves LE, Zenger KR, Cooper DW, Eldridge MDB (2010) Molecular detection of hybridization between sympatric kangaroo species in south-eastern Australia. Heredity 104: 502–512.
33. Meng C, Kubatko LS (2009) Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: a model. Theor Popul Biol 75: 35–45.
34. Chung Y, Ané C (2011) Comparing two Bayesian methods for gene tree/species tree reconstruction: simulations with incomplete lineage sorting and horizontal gene transfer. Syst Biol 60: 261–275.
35. Buckley TR, Cordeiro M, Marshall DC, Simon C (2006) Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada* Dugdale). Syst Biol 55: 411–425.
36. Joly S, McLenachan P, Lockhart P (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. Am Nat 174:E54–E70.
37. Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13: 727–744.
38. Moritz C, Cicero C (2004) DNA barcoding: Promise and pitfalls. PloS Biol 2: 1529–1531.
39. Rubinoff D, Holland BS (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst Biol 54: 952–961.
40. Mueller RL (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. Syst Biol 55: 289–300.
41. Galtier N, Nabholz B, Glémin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol Ecol 18: 4541–4550.
42. Calvignac S, Konecny L, Malard F, Douady CJ (2011) Preventing the pollution of mitochondrial datasets with nuclear mitochondrial paralogs (numts). Mitochondrion 11: 246–254.
43. Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. Mol Ecol 17: 2107–2121.
44. Jesus FF, Wilkins FJ, Solferini VN, Wakeley J (2006) Expected coalescence times and segregating sites in a model of glacial cycles. Genet Mol Res 5: 466–474.
45. Moran NA (1996) Accelerated evolution and Muller's Ratchet in endosymbiotic bacteria. Proc Natl Acad Sci, USA 93: 2873–2878.
46. Lynch M, Blanchard JL (1998) Deleterious mutation accumulation in organelle genomes. Genetica 103: 29–39.
47. Ballard JWO (2000) When one is not enough: introgression of mitochondrial DNA in *Drosophila*. Mol Biol Evol 17: 1126–1130.
48. Westerman M, Meredith RW, Springer MS (2010) Cytogenetics meets phylogenetics: a review of karyotype evolution in diprotodontian marsupials. J Hered 101: 690–702.
49. Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol Biol Evol 13: 510–524.
50. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.
51. Rambaut A (1996) Sequence alignment editor. Available: http://tree.bio.ed.ac.uk/software/seal.
52. Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sunderland, MA: Sinauer Associates, Inc.
53. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253–1256.
54. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
55. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood–based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.
56. Rambaut A, Drummond AJ (2007) Tracer v1.5. Available: http://beast.bio.ed.ac.uk/.
57. Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Syst Biol 51: 492–508.

58. Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17: 1246–1247.
59. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log likelihoods with applications to phylogenetic inference. Mol Biol Evol 16: 1114–1116.
60. Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order Hominidae. J Mol Evol 29: 170–179.
61. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214. Available: http://www.biomedcentral.com/1471-2148/7/214.
62. Kear BP, Pledge NS (2007) A new fossil kangaroo from the Oligocene–Miocene Etadunna Formation of Ngama Quarry, Lake Palankarinna, South Australia. Aust J Zool 55: 331–339.
63. Woodburne MO, Macfadden BJ, Case JA, Springer MS, Plegde NS, et al. (1993) Land mammal biostratigraphy and magnetostratigraphy of the Etadunna Formation (late Oligocene) of South Australia. J Vert Paleont 13: 483–515.
64. Flannery TF, Rich TH, Turnbull WD, Lundelius EL Jr (1992) The Macropodoidea (Marsupialia) of the early Pliocene Hamilton Local Fauna, Victoria, Australia. Fieldiana: Geology, New Series 25: 1–37.
65. Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58: 367–380.
66. Farris JS, Källersjö M, Kluge AG, Bult C (1994) Testing significance of incongruence. Cladistics 10:315–319.
67. Barker FK, Lutzoni F (2002) The utility of the incongruence length difference test. Syst Biol 51: 625–637.
68. Rambaut A, Grassly NC (1997) Seq-Gen: an application for Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput Appl Biosci 13: 235–238.
69. Than C, Ruths D, Nakhleh L (2008) PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinf 9: 322. Available: http://www.biomedcentral.com/1471-2105/9/322.
70. Suzuki Y, Glazko GV, Nei M (2002) Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc Natl Acad Sci USA 99: 16138–16143.
71. Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164:1645–1656.
72. Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. Proc Natl Acad Sci USA 107:9264–9269.
73. De Magalhaes JP, Costa J (2009) A database of vertebrate longevity records and their relation to other life–history traits. J Evol Biol 22: 1770–1774.
74. Neaves LE, Zenger KR, Prince RIT, Eldridge MDB (2012) Impact of Pleistocene aridity oscillations on the population history of a widespread, vagile Australian mammal, Macropus fuliginosus. J Biogeog 39: 1545–1563.
75. Hollis CJ, Robertshaw JD, Harden RH (1986) Ecology of the swamp wallaby (*Wallabia bicolor*) in north-eastern New South Wales. I. diet. Aust Wild Res 13: 355–365.
76. Bee CA, Close RL (1993) Mitochondrial DNA analysis of introgression between adjacent taxa of rock-wallabies, *Petrogale* species (Marsupialia: Macropodidae). Genetical Res 61: 21–37.
77. Van Gelder RG (1977) Mammalian hybrids and generic limits. Am Mus Novit 2635: 1–25.
78. Close RL, Lowry PS (1990) Hybrids in marsupial research. Aust J Zool 37: 259–267.
79. Ropiquet A, Hassanin A (2006) Hybrid origin of the Pliocene ancestor of wild goats. Mol Phylogenet Evol 41: 395–404.
80. Melo-Ferreira J, Boursot P, Carneiro M, Esteves PJ, Farelo L, et al. (2012) Recurrent introgression of mitochondrial DNA among hares (*Lepus spp.*) revealed by species–tree inference and coalescent simulations. Syst Biol 61: 367–381.
81. Milne N, O'Higgins P (2002) Inter-specifc variation in *Macropus* crania: form, function and phylogeny. J Zool 256: 523–535.
82. Christensen P (1995) Western brush wallaby. In: Strahan R, editor. The Mammals of Australia. Sydney: Reed New Holland. pp. 341–342.
83. Wann JM, Bell DT (1997) Dietary preferences of the black-gloved wallaby (*Macropus irma*) and the western grey kangaroo (*M. fuliginosus*) in Whiteman Park, Perth, Western Australia. J Roy Soc West Aust 80: 55–62.
84. Hay J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167: 747–760.

APPENDIX C – MORPHOLOGICAL AND MOLECULAR EVIDENCE SUPPORTS SPECIFIC RECOGNITION OF THE RECENTLY EXTINCT *BETTONGIA ANHYDRA* (MARSUPIALIA: MACROPODIDAE).

The following citation resulted from research that is of relevance to this doctoral thesis and was published during my candidature. In context of this multi-author publication I carried out all genetic analyses, this included; DNA extraction, qPCR screening and bioinformatic work. I also carried out all molecular model tests and Bayesian phylogenetic analyses (Figure 4).

A copy of the main article is attached. The Supplementary material may be viewed online at: http://jmammal.oxfordjournals.org/content/96/2/287#app-1

McDowell, M.C., **Haouchar, D**., Aplin, K.P., Bunce, M., Baynes, A., Prideaux, G.J., 2015. Morphological and molecular evidence supports specific recognition of the recently extinct *Bettongia anhydra* (Marsupialia: Macropodidae). *Journal of Mammalogy* **96**, 287-296.

Journal of Mammalogy, 96(2):287–296, 2015
DOI:10.1093/jmammal/gyv006

Morphological and molecular evidence supports specific recognition of the recently extinct *Bettongia anhydra* (Marsupialia: Macropodidae)

Matthew C. McDowell,* Dalal Haouchar, Ken P. Aplin, Michael Bunce, Alexander Baynes and Gavin J. Prideaux

School of Biological Sciences, Flinders University, Bedford Park, Adelaide, South Australia 5042, Australia (MCM, GJP)
Trace and Environmental DNA (TrEnD) Laboratory, Department of Environment and Agriculture, Curtin University, Perth, Western Australia 6102, Australia (DH, MB)
Division of Mammals, Smithsonian Institution, PO Box 37012, MRC 108, Washington, DC 20013-7012, USA (KPA)
Department of Earth and Planetary Sciences, Western Australian Museum, Locked Bag 49, Welshpool DC, Western Australia 6986, Australia (AB)

* *Correspondent: matthew.mcdowell@flinders.edu.au*

In 1933, geologist and explorer Michael Terry collected the skull of a small macropodid captured by members of his party near Lake Mackay, western Northern Territory. In 1957, this skull was described as the sole exemplar of a distinct subspecies, *Bettongia penicillata anhydra*, but was later synonymized with *B. lesueur* and thereafter all but forgotten. We use a combination of craniodental morphology and ancient mitochondrial DNA to confirm that the Lake Mackay specimen is taxonomically distinct from all other species of *Bettongia* and recognize an additional specimen from a Western Australian Holocene fossil accumulation. *B. anhydra* is morphologically and genetically most similar to *B. lesueur* but differs in premolar shape, rostrum length, dentary proportions, and molar size gradient. In addition, it has a substantial mitochondrial cytochrome *b* pairwise distance of 9.6–12% relative to all other bettongs. The elevation of this recently extinct bettong to species status indicates that Australia's mammal extinction record over the past 2 centuries is even worse than currently accepted. Like other bettongs, *B. anhydra* probably excavated much of its food and may have performed valuable ecological services that improved soil structure and water infiltration and retention, as well as playing an important role in the dispersal of seeds and mycorrhizal fungal spores. All extant species of *Bettongia* have experienced extensive range contractions since European colonization and some now persist only on island refugia. The near total loss of these ecosystem engineers from the Australian landscape has far-reaching ecological implications.

Key words: biodiversity loss, digging, ecological service, environmental degradation, extinction

European colonization has had major environmental repercussions that have fundamentally transformed Australia's biogeography, ecosystems, and landscapes, causing widespread declines in biodiversity (e.g., McDowell et al. 2012). While these impacts have affected all Australian native mammals, few taxa have fared as badly as the potoroines (species of *Aepyprymnus*, *Bettongia*, *Caloprymnus* and *Potorous*)—van Dyck and Strahan (2008). Potoroines typically excavate the majority of their food, and in doing so, perform valuable ecosystem services such as seed and spore dispersal, facilitation of seedling germination and establishment, soil aeration, incorporation of organic matter, and improvement in moisture infiltration (Martin 2003; Fleming et al. 2013). Some exotic species such as rabbits are also fossorial but do not contribute to soil improvement as effectively as potoroines (Vitousek 1990; James et al. 2011). Consequently, in areas where Potoroines have been extirpated or become extinct, soils are likely to have become drier, dustier, more compacted, and less fertile, reducing the productivity of the whole ecosystem.

Potoroines, sometimes known as rat-kangaroos, are small- to medium-sized nocturnal marsupials that occupy a basal branch within the Macropodidae (Prideaux and Warburton 2010). They retain plesiomorphic characteristics such as a prehensile tail, less-reduced forelimbs, well-developed upper canines, large blade-like sectorial premolars, and low-crowned molars (Claridge et al. 2007). Many species of *Bettongia* subsist primarily on excavated hypogeal fungi (Seebeck and Rose 1989; Claridge et al. 2007), which form mycorrhizal associations with the roots of vascular plants and help to maintain soft, friable, well-structured topsoil (Martin 2003; Eldridge and James

2009; Eldridge et al. 2012). Bettongs were once broadly distributed across Australia, but since European colonization each species has been extirpated from much of their former ranges or have become extinct (Short 1998; van Dyck and Strahan 2008).

Four extant and 2 extinct species of *Bettongia* are currently recognized. Taxonomy of the genus was partially revised by Finlayson (1958) and more comprehensively by Wakefield (1967), who raised *B. tropica* from what had hitherto been considered a northern population of *B. penicillata*. Subsequently, 2 fossil species have been added: *B. moyesi* from the Miocene Riversleigh assemblage in northwestern Queensland (Flannery and Archer 1987) and *B. pusilla* from Holocene cave deposits of the Nullarbor Plain (McNamara 1997).

In describing *B. p. anhydra*, Finlayson (1957:553) noted its "remarkable blend of *penicillata* and *lesueuri* [sic] characters" and commented that "if its dual character were confirmed in series, it [*B. p. anhydra*] would merit specific recognition." Wakefield (1967) considered key features of the specimen, including its short rostrum, very large bullae, and proportionately long premolars, resembled *B. lesueur*, and thence declared them synonymous. One author (KPA) examined the holotype of *B. p. anhydra* in 1997 and concluded that it was specifically distinct. More recently, MCM independently reached the same conclusion and observed a Holocene fossil specimen from a Western Australian cave accumulation that demonstrates similar craniodental morphology. In this paper, we recognize *B. anhydra* as a distinct species on the basis of morphological and molecular evidence and consider the ecological implications of its 20th century disappearance.

Materials and Methods

Morphological analysis.—The holotype of *B. anhydra* (SAM M3582) was examined and compared with representative specimens of other species of *Bettongia* to determine taxonomic affinities. The following abbreviations are used in this work: FUR = Flinders University of South Australia reference collection; SAM = South Australian Museum (M: mammal collection; P: palaeontological collection); WAM = Western Australian Museum palaeontological collection; QM = Queensland Museum mammal collection. Dental homology, nomenclature, and family-group taxonomy follows Prideaux (2004) and Prideaux and Warburton (2010). Specimens used for comparison with *B. anhydra* are listed in Appendix I.

Genetic analysis.—The left turbinal bone from the nasal cavity of the *B. anhydra* holotype cranium was sampled using sterile forceps then placed in a labeled sterile vial. The turbinal was chosen because, being inside the nasal cavity, it has been largely protected from contamination due to human handling and its removal did not appreciably alter the appearance of the skull (see Wisely et al. 2004). DNA extraction procedures were carried out in a dedicated ancient DNA (aDNA) laboratory at Murdoch University, minimizing contamination from PCR amplicons and modern DNA. The sample was crushed to powder then stored for DNA extraction and amplification.

The bone digest buffer consisted of: 20 mM Tris pH 8.0 (Sigma, Kansas City, Missouri), 10 mM dithiothreitol (Thermo Fisher Scientific, Waltham, Massachusetts), 1 mg/ml proteinase K powder (Amresco, Solon, Ohio), 0.48 M ethylenediaminetetraacetic acid (Invitrogen, Carlsbad, California), and 1% Triton X-100 (Invitrogen). A total of 1,500 µl of bone digest buffer was added to the bone powder then incubated overnight at 55°C with rotation. After digestion, the solution was centrifuged at 13,000 × *g* for 1 min to pellet undigested material. The supernatant containing the DNA was concentrated to approximately 100 µl in a Vivaspin 500 column (MWCO 30,000; Sartorius Stedim Biotech, Goettingen, Germany) at 13,000 × *g* then combined with 5 volumes of PBi buffer (Qiagen, Valencia, California). DNA was immobilized on silica spin columns (Qiagen) and washed with 700 µl of AW1 and AW2 wash buffers (Qiagen). Finally, the DNA was eluted from the silica in 50 µl of 10 mM Tris pH 8.0 (Sigma).

The DNA extract was screened using specifically designed primer sets targeting the cytochrome *b* gene for *Bettongia*. Primer sets woylie_cytb_139F ACCTTCCAACATTTGAGCCTGATG and woylie_cytb_388R TGAGCCGTAGTAGATTCCTC were used to target a ~ 200 base pair region of cytochrome *b* DNA. Each quantitative polymerase chain reaction (qPCR) was made up to a total volume of 25 µl, containing 12.5 µl ABI Power SYBR master mix (Applied Biosystems, Waltham, Massachusetts), 0.4 µM of forward and reverse primer, 8.5 µM H_2O, and 2 µl DNA extract. Reaction conditions for the specific mammal primer sets were as follows; heat denatured at 95°C for 5 min, followed by 50 cycles of 95°C for 30 s and 56°C for 30 s.

Once initial qPCR screening showed that DNA of sufficient quality and free of inhibition was achieved, it was prepared for Sanger sequencing. To ensure validity of DNA sequences and to overcome ancient DNA damage, multiple sequence data sets were created and an overall consensus was drawn for use in the final analyses. Amplicons were cleaned (Qiagen columns) and prepared for capillary sequencing following Haouchar et al. (2013). Alignments of nucleotide sequences were carried out in GENEIOUS 6.0.1 (Drummond et al. 2011), with any ambiguities resolved by eye. Other *Bettongia* sp. sequences provided by the National Center for Biotechnology Information's GenBank were used to construct an alignment of the new *B. anhydra* sequence, presented in the Bayesian phylogeny.

Phylogenetic analysis was estimated using the Bayesian phylogenetic program BEAST version 1.7.4 (Drummond and Rambaut 2007; Drummond et al. 2012). Sequences were run through JMODELTEST to determine the most appropriate nucleotide substitution model (Posada 2008). A relaxed uncorrelated log-normal clock was employed. Three independent runs of 1×10^6 generations were performed, with every 1,000 generations sampled with 10% burn-in. Analyses were checked in TRACER (Rambaut and Drummond 2009) for convergence and adequate effective sample size. Phylogenetic trees were summarized using TREEANNOTATER version 1.7.4 (Drummond and Rambaut 2007) and visualized in FIGTREE 1.4.0 (Rambaut 2007). The cytochrome *b* gene for *B. anhydra* has been deposited into GenBank under accession number KM974728.

Results

Higher Systematics

Order Diprotodontia Owen, 1866
Superfamily Macropodoidea Gray, 1821
Family Macropodidae Gray, 1821
Subfamily Potoroinae Gray, 1821
Tribe Bettongini Flannery and Archer, 1987
Genus *Bettongia* Gray, 1837

Revised Diagnosis of Bettongini

Tribe Bettongini includes species of *Bettongia*, *Caloprymnus*, *Aepyprymnus*, and *Milliyowi*. Bettongins can be differentiated from members of the tribe Potoroini (containing *Potorous*) based on the following characters: cranium bears postglenoid process and discrete periotic ectotympanic process; I3 short crowned. P3 bears many fine vertical ridgelets; buccal crests of upper molars better developed than lingual counterparts; dentary stout with convex ventral margin; i1 lacks a dorsal and ventral enamel flange; p3 bears many vertical ridgelets; lingual crests of lower molars better developed than buccal counterparts.

Revised Diagnosis of Bettongia

Species of *Bettongia* are united by 1 synapomorphy: jugal extends dorsally to at least level of large lachrymal foramen. However, they can be further differentiated from species of *Caloprymnus* and *Aepyprymnus* by their combined possession of large posterior palatal vacuities, inflated auditory bullae, and P3/p3 with 6 or more vertical ridgelets.

Bettongia anhydra—Finlayson, 1957
Synonyms
Bettongia penicillata anhydra Finlayson, 1957
Bettongia penicillata anhydra Finlayson, 1958
Bettongia lesueur Wakefield, 1967; in part
Bettongia lesueur Calaby and Richardson, 1988; in part

Holotype

Near-complete adult cranium (SAM M3582) with associated left and right dentaries collected from a fresh carcass by Michael Terry in 1933 from the McEwin Hills area, near Lake Mackay, Northern Territory. Cranium lacks the entire occipital complex, both petrosals and ectotympanics, interparietal and part of the left squamosal; the tympanic bullae are broken but enough is preserved to infer shape and degree of inflation. Left dentary complete, though m4 absent. Right dentary articular process not preserved.

Type Locality

McEwin Hills, Lake MacKay area, Northern Territory, approximately 22°02′S, 129°47′E (Fig. 1).

Referred Specimens

Stegamite Cave (5N194), Eucla Basin: WAM 67.10.530, adult left maxilla with P3 and M1–3 (M4 absent). Note: this specimen was referred to *B. anhydra* based on morphological characters only.

Diagnosis

Bettongia anhydra can be distinguished from all other species of *Bettongia* by the following features: steep posterior molar gradient (M/m1 ≤ M/m2 > M/m3 >> M/m4) and highly reduced fourth molars (Table 1); anteroposterior compression of rostrum; marked interorbital constriction; and obscuration of m4, m3 hypolophid, and part of m3 protolophid by ascending ramus in lateral view.

It can be further separated from *B. gaimardi*, *B. penicillata*, and *B. tropica* by its greater degree of inflation of auditory bullae (similar to *B. lesueur*); from *B. pusilla* by its lower-crowned and less lophodont molars, and steeper molar gradient; from *B. gaimardi*, *B. lesueur*, *B. moyesi*, *B. pusilla*, and *B. tropica* by its greater buccal flexion of anterior third of P3; from *B. penicillata* by its lesser buccal flexion of anterior third of P3. Dentary differs from *B. gaimardi*, *B. lesueur*, *B. penicillata*, and *B. pusilla* in robustness of horizontal ramus and anterior occurrence of digastric eminence; from *B. gaimardi*, *B. moyesi*, *B. penicillata*, *B. pusilla*, and *B. tropica* in greater length and breadth of coronoid process and more acute angle between ascending ramus and horizontal ramus.

Morphological Analysis

All features in referred specimen are as for holotype. No juvenile or postcranial specimens are known.

Cranium.—Premaxilla short, with upright portion essentially vertical. Anterior edge of premaxilla very slightly arcuate in lateral view (Figs. 2a and 2e). Diastema very short, straight, and only slightly deflected anteroventrally relative to cheek tooth row; maxilla contributes to most of diastema length. I1 high crowned and peg shaped, I2 and I3 comparatively elongate anteroposteriorly. C1 well developed and close to, but shorter than I3. Anterior palatal foramina broad and short, terminating posteriorly just

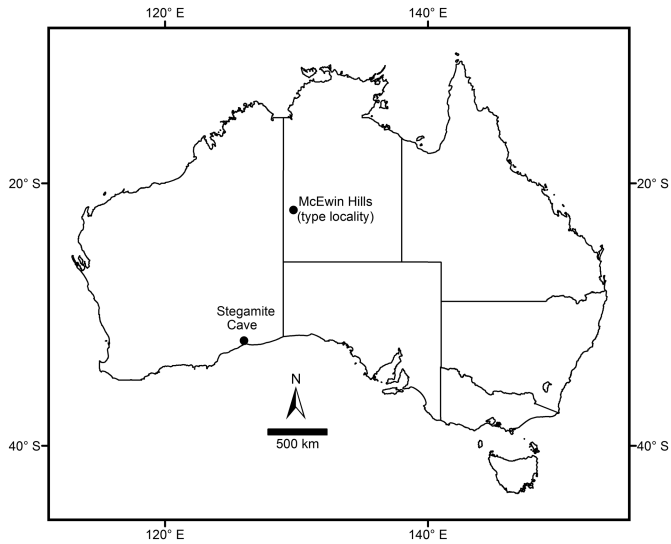


Fig. 1.—Collection locations of the holotype and referred specimen of *Bettongia anhydra*.

Table 1.—Mean cranial and dental measurements (mm) of modern *Bettongia* species. Some data for *B. gaimardi cuniculus*, *B. gaimardi gaimardi*, *B. penicillata ogilbyi*, and *B. tropica* after Wakefield (1967) and data for *B. lesueur* after Finlayson (1958).

	B. anhydra (holotype)	WAM 67.10.530	*B. gaimardi cuniculus*	*B. gaimardi gaimardi*	*B. lesueur*	*B. penicillata ogilbyi*	*B. tropica*
Basal length			70.6	64.4	57.2	66.5	64.2
Zygomatic width	37.4		45.3	42.2	43.0	24.2	41.2
Interorbital width	12.2		19.9	18.5	14.5	17.4	15.8
Nasals length	23.7		34.7	30.8	26.3	32.5	29.6
Nasals, greatest width	9.5		14.5	13.9	12.8	13.5	13.5
Rostrum width	13.5		17.4	15.4	17.9	16	14.3
Nasal opening width	5.1		9.1	8.7	6.8	7.9	7.8
Bulla length	10.4		12.5	12.1	15.9	14.2	13.8
Bulla depth			9.5	9.1		10.6	10.9
P3 length	7.5	6.8	8.1	7.2	8.5	7.3	8.3
M1–3 length	10.5	10.6	13.7	12.8	11.7	12.4	13.2
M4 length	1.4		3.8	3.3	2.5	2.6	2.8

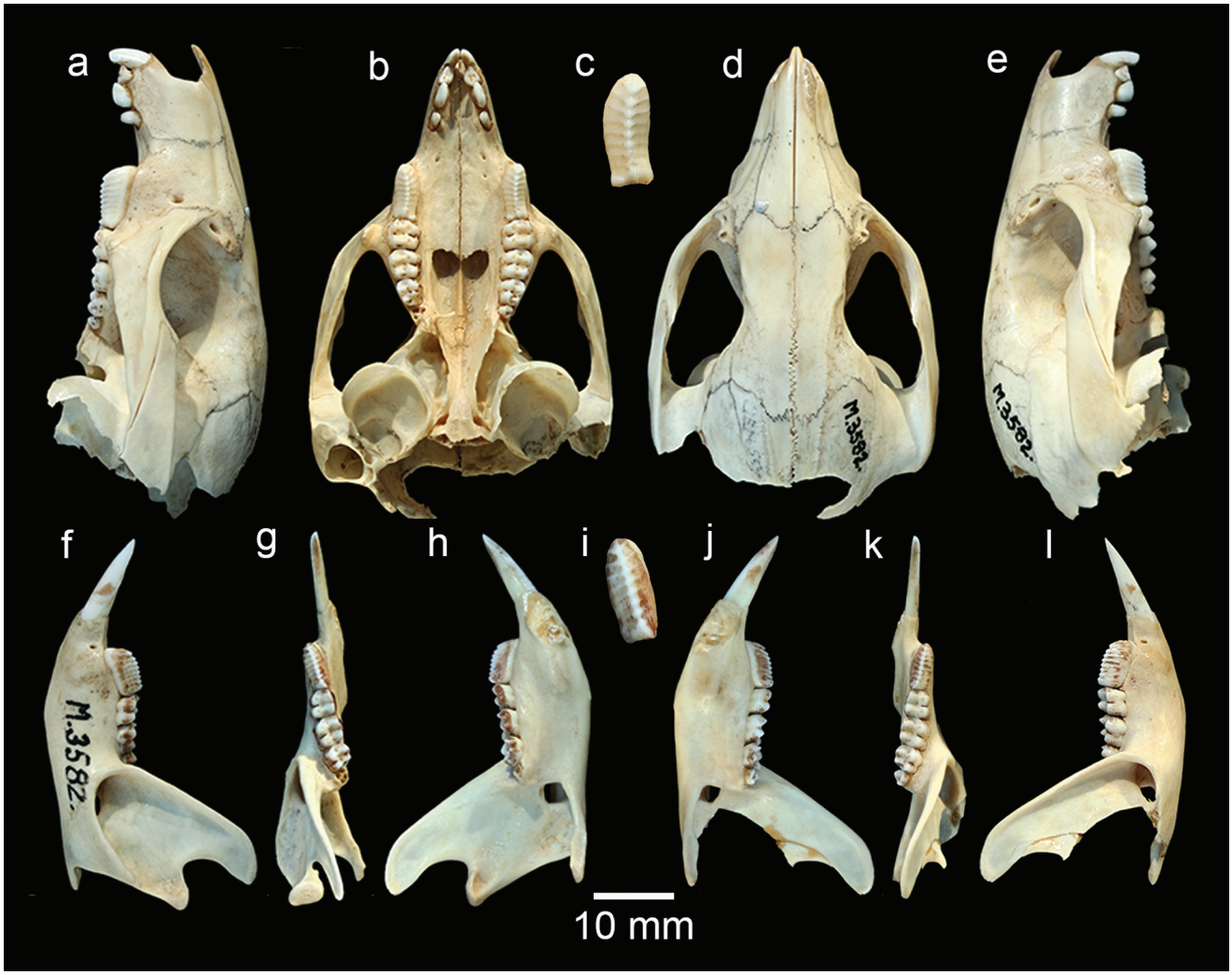


Fig. 2.—The holotype of *Bettongia anhydra* (SAM M3582): a) Left lateral, b) occlusal, c) magnified (2×) view of left P3, d) dorsal, e) right lateral aspect of skull, f) buccal, g) occlusal, h) lingual aspect of left dentary, i) magnified (2×) view of left p3, j) buccal, k) occlusal, l) lingual aspect of right dentary.

past anterior edge of C1 alveolus (Fig. 2b). Very short, deep rostrum strongly tapered anteriorly; lateral edges enclose angle of 20° (Figs. 2b and 2d). Premaxilla contributes approximately half of length of lateral surface of rostrum. Narial aperture deeper than wide (Figs. 2a and 2e). Buccinator fossa shallow, restricted to ventral half of lateral surface of rostrum and extending from anterior edge of P3 anterior root to posterior edge of C1 alveolus (Figs. 2a, 2b, and 2d). Short masseteric process composed entirely of maxilla; positioned adjacent to M1 protoloph (Figs. 2a, 2b, and 2e). Infraorbital foramen opens anteriorly; positioned directly above posterior root of P3 at level of ventral border of orbit. Small posterior (dorsal) lacrimal foramen opens dorsally;

separated from larger anterior (ventral) lacrimal foramen by large lacrimal tuberosity which marks anterodorsal extremity of orbital rim. Anterior nasals narrowly constricted at maxilla–premaxilla suture. Nasofrontal sutures arcuate, extending posterior to anterior edge of orbit. Palatine bones well developed. Large posterior palatal foramina originate adjacent to metaloph of M1 and extend posteriorly along remaining length of palate (Fig. 2b).

Weakly developed temporal (parietal) crests confluent anteriorly with supraorbital crests, extending posteriorly across interparietals (Fig. 2d). Dorsal surface of neurocranium gently curved to posterior terminus of nasals. Zygomatic arch deep; posterior extremity of jugal bears very small ectoglenoid process (Figs. 2a and 2d). Postorbital process of jugal distinct and pointed. Zygomatic process of squamosal arises well anterior of occiput. Very small postglenoid process forms posterior border of glenoid fossa, curves slightly anteriorly at extremity, giving glenoid fossa a semicircular shape when viewed laterally. Auditory bullae highly inflated (Fig. 2b).

Upper incisors.—I1 high crowned, arcuate when viewed laterally (Figs. 2a and 2e). I2 blade like, crown height lower than I3. I3 crown subtriangular in buccal view (Figs. 2a and 2e). Occlusal surface oriented anteroposteriorly in same line as lateral edge of rostrum.

P3.—P3 anteroposteriorly elongate, blade-like, and bears 7 buccal and lingual enamel ridgelets ascending anterodorsally from 7 main crest cuspules. Anterior third flexes slightly buccally. P3 bears moderately developed posterolingual eminence. P3 much longer than all molars, equal in length to M1–2 combined (Fig. 2b).

Upper molars.—Bunolophodont. Holotype with M1 slightly worn, dentine of protocone, paracone, and metaconule breached; M2 slightly worn, dentine of paracone breached; M3–4 unworn (Fig. 2b). M1 protoloph and metaloph of equal width. M2–4 protoloph wider than metaloph. Lingual margin of tooth row virtually straight; buccal margin convex laterally due to marked size reduction of molars posteriorly. Paracrista low but distinct, merges with weaker (worn) protocrista to form protoloph. Preprotocrista unites with preparacrista forming precingulum. Postprotocrista weak, unites with strong premetaconulecrista. Metacone higher crowned than metaconule. Metacrista well developed forming majority of metaloph. Metaconulecrista weak. Premetacrista and postparacrista weakly developed and do not unite. Postmetacrista moderately well developed, terminates in position of stylar cusp E. Weak postmetaconulecrista joins postmetacrista at position of stylar cusp E defining posterior border. M4 highly reduced, protocone well-developed, paracone reduced, metacone highly reduced, metaconule absent. Preparacrista forms a small precingulum, paracrista weak, unites with well-developed preprotocrista. Postparacrista runs posterobuccally and contacts metacone, defining posterior border of M4 (Fig. 2b).

Dentary.—Horizontal ramus stout with a convex ventral margin, digastric eminence deep, occurs below m1 hypolophid (depth 8.7 mm; Figs. 2e–j). Digastric sulcus shallow (Figs. 2e and 2j). Buccinator sulcus straight, shallow, extends beneath posterior third of p3 to protolophid of m1. Anterior root of vertical ascending ramus adjacent to posterior of m3 protolophid (Figs. 2f and 2l). Angular process wide; lingual border thickened, tip pointed posteriorly. Masseteric fossa deep, ventral border extends well below buccinator sulcus to half depth of horizontal ramus. Anterior insertion area for 2nd layer of masseter muscle thin and restricted to rim of masseteric fossa (Figs. 2f and 2l). Masseteric foramen large, anteroventrally oriented and leads into masseteric canal which extends to beneath m1 protolophid. Mandibular foramen oval shaped, opening largely posteriorly (Figs. 2g, 2h, and 2j). Articular process anteroposteriorly wide, articular condyle small, wider laterally. Coronoid process anteroposteriorly wide (viewed laterally), with slight posterior "hook" at dorsal end.

Lower incisor.— Lanceolate i1, bears moderate wear on anterior half of superior border (Figs. 2f and 2l). Diastema short, approximately 2/3 length of p3.

p3.—Blade-like p3 anteroposteriorly elongate, aligned with molar row. Bears 8 buccal and lingual enamel ridges, which descend vertically from 8 crest cuspules. p3 equal in length to m1–2 combined (Figs. 2f–l).

Lower molars.—Bunolophodont. Holotype with m1 slightly worn, dentine of protoconid and hypoconid breached; m2 slightly worn, dentine of hypoconid breached; m3–4 unworn (Figs. 2g and 2k). Lophid faces smooth; m1 protolophid narrower than hypolophid, m2–4 hypolophid narrower than protolophid. Metaconid and entoconid taller than protoconid and hypoconid. Pre- and postmetacristids and pre- and post-entocristids all well developed. Well-developed metacristid forms protolophid, protocristid very weak. Well-developed entocristid forms hypolophid, hypocristid very weak. Low lingual cristid obliqua bisects interlophid valley. Paracristid (buccal) merges with premetacristid enclosing small trigonid basin. Small postcingulid defined by equally developed postentocristid and posthypocristid. Highly reduced m4, metaconid, protoconid, and hypoconid subequal in height, entoconid absent. Weak premetacristid merges with buccal paracristid defining reduced trigonid basin. Weak buccal cristid oblique connects protolophid and hypolophid (Figs. 2g and 2k).

Comparison with Other Species of Bettongia

Cranium.—*Bettongia anhydra* is smaller in overall cranial dimensions than *B. penicillata*, *B. gaimardi*, *B. lesueur*, and *B. tropica*, but larger than *B. pusilla*. Its dentition is larger relative to the size of the cranium than in all other bettongs (although the cranium of *B. pusilla* is unknown; Figs. 3a, 3d, 3g, 3j, and 3m). Compared with the other species of *Bettongia*, *B. anhydra* has a more reduced M4 relative to M3, a shorter rostrum, and narrower frontals and nasals (Table 1; Figs. 2 and 3); interorbital region more constricted than in any other species in the genus (Table 1); braincase narrower and more tapered anteriorly than in other bettongs (Finlayson 1958); squamosal makes greater contribution to the zygomatic arch which is deep and robust than in other *Bettongia* spp. *B. anhydra* shares with *B. lesueur* marked inflation of the tympanic bulla, an attribute that easily distinguishes the crania of these species from those of *B. penicillata*, *B. gaimardi*, and *B. tropica* (Fig. 3). Temporal crests are well developed but instead of following the line of the interorbital ridge they extend dorsally toward the midline of the skull suggesting large temporalis muscles relative to skull size (Fig. 2d); diastema and anterior palatal foramina shorter than in other species of *Bettongia*, the posterior terminus of the latter occurring near the anterior border of the canine; posterior palatal foramina

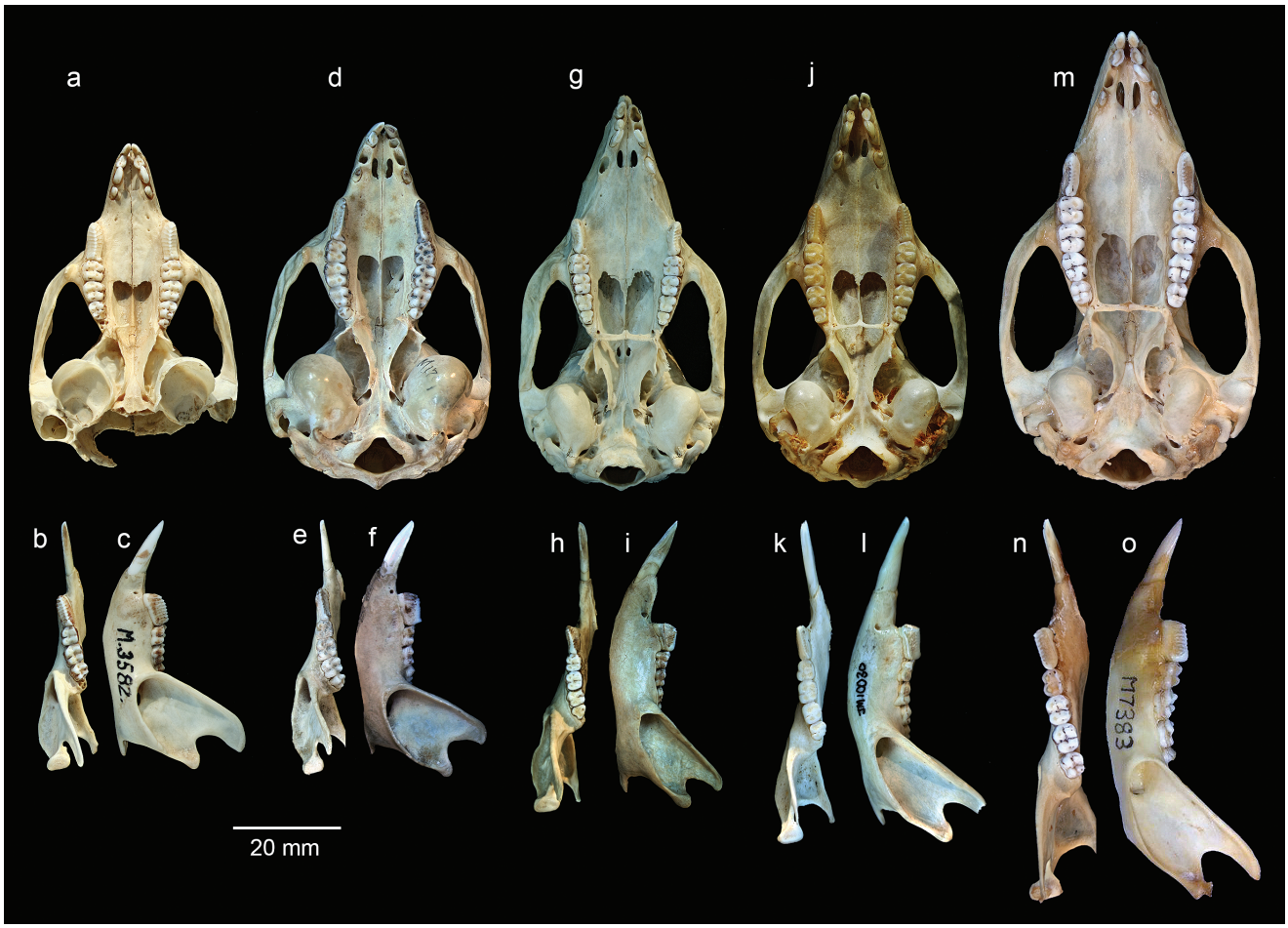


Fig. 3.—Comparison of *Bettongia anhydra* with other extant species of *Bettongia*, showing occlusal view of the skull and left dentary and buccal lateral view of the left dentary of a–c) *B. anhydra*; d–f) *B. lesueur*; g–i) *B. penicillata*; j–l) *B. tropica*; m–o) *B. gaimardi*.

large with anterior margins occurring near the posterior margin of M1 but shorter than in other bettongs (Fig. 3).

P3.—The P3 of *B. anhydra* flexes slightly anterobuccally but remains within the line with the molar row; differs from *B. gaimardi*, *B. lesueur*, *B. moyesi*, *B. pusilla*, and *B. tropica* in which P3 is straight and in line with the molar row (Fig. 2); differs from *B. gaimardi*, *B. lesueur*, and *B. moyesi* in which the lingual face of P3 is convex; differs from *B. penicillata* in which P3 flexes anterobuccally outside the line of the molar row; and differs from that of *B. gaimardi*, *B. lesueur*, *B. pusilla*, and *B. tropica* in which the superior and inferior borders of P3 are subparallel. It is similar to *B. penicillata* in that anterobuccal flexion of P3 increases the depth of the enamel on the anterior buccal face of the tooth such that it is approximately twice as deep as the posterior of the tooth (Wakefield 1967).

Upper molars.—Upper molars of *B. anhydra* differ from all other bettongs in the steepness of the molar gradient and extreme reduction of M4. It further differs from *B. pusilla* in which m1–4 approach equal size and are higher crowned and more lophodont (McNamara 1997); *B. gaimardi* in which M4 is only slightly smaller than M1–3 and in *B. penicillata*, *B. lesueur*, and *B. tropica* in which M1 ≤ M2 > M3 > M4, but M4 is much less reduced (Fig. 3).

Dentary.—Dentaries of *B. anhydra* are short relative to toothrow length and in lateral view the ascending ramus of *B. anhydra* obscures the view of m4 and most of m3 (Fig. 3c). In *B. gaimardi* (Fig. 3o), *B. lesueur* (Fig. 3f), *B. penicillata* (Fig. 3i), and *B. pusilla*, the ascending ramus obscures the view of m4 only and in *B. moyesi*, the ascending ramus obscures only part of m4. The dentary of *B. anhydra* also differs from *B. gaimardi*, *B. lesueur*, *B. penicillata*, and *B. pusilla* in the robustness of the horizontal ramus (the ventral margin of the jaw of *B. moyesi* is unknown), the greatest depth of which occurs quite anteriorly at the digastric eminence below m1 (Fig. 3; Finlayson 1957, 1958). It differs from *B. penicillata*, *B. gaimardi*, *B. moyesi*, *B. pusilla*, and *B. tropica* in the proportions of the coronoid process which is long and broad with subparallel borders, and in the angle between the ascending ramus and horizontal ramus which is more acute (Fig. 3).

p3.—p3 of *B. anhydra* (Figs. 2i, 3b, and 3c) has fewer, more broadly spaced cuspules and grooves than seen in *B. lesueur* (Figs. 3e and 3f) and differs from *B. gaimardi* (Figs. 3n and 3o), *B. lesueur* (Figs. 3e and 3f), *B. moyesi*, and *B. tropica* (Figs. 3k and 3l) in that the anterior portion of p3 deflects slightly buccally, though not as much as in *B. penicillata* (Figs. 3h and 3i); differs from *B. gaimardi*, *B. lesueur*, *B. pusilla*, and *B. tropica* in which the superior and inferior borders of p3 are subparallel but is similar to *B. penicillata* in that anterobuccal flexion of p3 increases the depth of the enamel on the anterior buccal face of the anterior portion of the tooth making it slightly deep than in the posterior portion of the tooth (Wakefield 1967).

Lower molars.—Lower molars of *B. anhydra* can be distinguished from other bettongs by their high posterior molar gradient where m1 ≤ m2 > m3 >> m4 and the extreme reduction of m4 (as in upper molars; Fig. 3).

Genetic Analysis

Approximately 200 bp of aDNA were successfully isolated from the cytochrome *b* gene of the *B. anhydra* holotype. Comparison of its DNA sequence with those species of *Bettongia* on GenBank (*B. gaimardi*, *B. lesueur*, *B. penicillata*, and *B. tropica*) revealed that all species share 92% or fewer identical sites in this DNA fragment—a strong indication that it is a genetically distinct species. Phylogenetic analysis (Fig. 4) of this cytochrome *b* fragment grouped *B. anhydra* as sister to *B. lesueur*, but with poor support. Phylogenetic analysis in BEAST was unable to clearly resolve the branching topology between *B. anhydra*, *B. lesueur*, and the other "surface nesting" bettongs. Kimura-2-p pairwise analyses of sequence data determined by MEGA 5.2.2 (Tamura et al. 2011) indicated the genetic distance between *B. anhydra* and other bettongs is estimated to be between 9.6% and 12% for the cytochrome *b* gene. The depth of this genetic split, taken together with the result of the morphological analyses, provides strong support for the specific status of *B. anhydra.*

Discussion

Taxonomy.—Finlayson (1957, 1958) stated that he would have assigned *B. anhydra* specific status if a series of specimens demonstrating its unique attributes existed. The overriding factor that drove Finlayson's placement of *B. anhydra* within *B. penicillata* seems to have been the similarity of their P3. Wakefield (1967), persuaded more by the overall shape of the cranium, placed the Lake Mackay specimen within *B. lesueur*. Aided by the recognition of an additional specimen and DNA analysis, we clearly show that *B. anhydra* is distinct from all other members of its genus. This is supported by evidence of its former sympatry with both *B. lesueur* and *B. penicillata* which also occupied semi-arid to arid habitats. *B. anhydra* shares several primarily plesiomorphic features with *B. lesueur* and *B. moyesi* suggesting that they lie outside a clade containing the remaining extant species of *Bettongia*, which appear more derived.

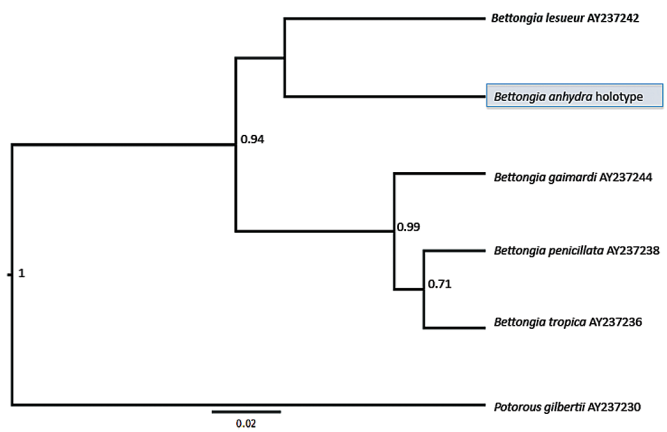


Fig. 4.—Bayesian phylogeny showing relationships between *Bettongia anhydra* (highlighted) and all other extant *Bettongia* species. Species are labeled with GenBank accession numbers. This tree was generated in BEAST using 203 bp of cytochrome *b* gene. A Hasegawa, Kishino, and Yano model and births-deaths tree prior with invariant gamma substitution sites was imposed with a relaxed molecular clock. Posterior probabilities > 70% are shown on selected nodes, scale bar represents the number of substitution sites per year.

Functional adaptations.—Many of the differences distinguishing *B. anhydra* from other bettong species relate to the shortening of the skull, e.g., reduced rostral length, short diastema, short palatal foramina, and highly reduced 4th molars. Mammalian rostrum morphology typically correlates with feeding adaptations (e.g., Mora et al. 2003; Pergams and Lawler 2009; Wilson and Sánchez-Villagra 2010). By shortening the rostrum and dentary but retaining unreduced anterior dentition, *B. anhydra* may have been able to apply greater bite force to its sectorial premolar or anterior molars, potentially allowing it to exploit harder foods such as browse or large seeds. This is consistent with the relatively high positioning of the temporalis origin, which suggests the presence of large temporalis muscles relative to skull size.

Bettongs occupying temperate parts of Australia are primarily fungivores (Johnson and McIlwee 1997; Claridge et al. 2007). However, *B. lesueur*, the most arid-adapted extant bettong, subsists mainly on roots and tubers, also occasionally consumes bulbs, carrion, insects, and seeds, including those of the Quandong and Sandalwood (*Santalum* spp.), which have hard seed-coats (Claridge et al. 2007). *Caloprymnus campestris*, a bettongin which occupied the absolutely lowest rainfall zone in Australia prior to its extinction near the middle of the 20th century, also had a very short robust skull and is reported to have been primarily herbivorous (Finlayson 1932). Although hypogeal fungi occur in the Australian arid zone (Trappe et al. 2008), the morphological similarities shared by *B. anhydra*, *B. lesueur*, and *C. campestris* suggest that *B. anhydra* may have supplemented its diet with browse. A similar diet was inferred for *Borungaboodie hatcheri*, a very large bettong known only from the late Pleistocene of southwestern Australia (Prideaux 1999).

The most striking characteristic shared by *B. lesueur* and *B. anhydra* is the extreme inflation of their auditory bullae. Desert dwelling mammals frequently possess more inflated auditory bullae than similar-sized forest-dwelling relatives. A larger middle ear air volume is often associated with more acute low frequency hearing which may enhance predator detection, thereby conferring advantages to animals foraging in open areas (e.g., Francescoli et al. 2012). Phylogenetic analysis (Fig. 4) identified *B. lesueur* and *B. anhydra* as sister taxa. Therefore, it is possible that inflated auditory bullae occurred in a common ancestor as an early adaptation to aridification of Australia.

Ecological implications.—The near obliteration of bettongs from mainland Australia has likely had serious ecological repercussions. Bettongs, potoroos, and similar ground-foraging small mammals cultivate the soil and in doing so provide important ecological services (Fleming et al. 2013; McDowell 2014). Soil disturbance has implications for incorporation of organic matter, aeration, moisture infiltration, seed germination, and seedling establishment (Martin 2003; Fleming et al. 2013). In addition, it promotes microorganism growth, influences topsoil formation,

and improves water penetration and retention, thereby enhancing soil structure (Fleming et al. 2013). Bettongs also play an integral role in the dispersal of seeds and fungal spores, many of which form symbiotic relationships crucial for the establishment and growth of numerous native plants, particularly eucalypts (Claridge et al. 2007). These ecological services are not replicated by introduced fossorial mammals such as the European rabbit, *Oryctolagus cuniculus* (James et al. 2011). Ecological services performed by bettongs are so important that they may actually determine vegetation succession and facilitate greater biodiversity (Martin 2003). The loss of bettongs and other ground-foraging small mammals from much of mainland Australia has probably compounded soil compaction problems caused by hard-hooved livestock, leaving little doubt that their loss will have far-reaching ecological impacts (e. g., Johnson and McIlwee 1997; Garkaklis et al. 1998; Martin 2003).

Since Europeans colonized Australia every bettong species, and most other potoroines, have either been extirpated from most of their original geographic range or driven entirely to extinction (see van Dyck and Strahan 2008). *Bettongia gaimardi gaimardi* is restricted to Tasmania and the mainland form *B. gaimardi cuniculus* is extinct. *Bettongia lesueur*, which once occupied much of the continents arid and semi-arid zones only a century ago, is now restricted to a few small islands off the Western Australian coast (Burbidge et al. 2007). *Bettongia penicillata ogilbyi* persists in a few small populations in southern Western Australia but *B. p. penicillata* is extinct. *Bettongia tropica* persists in a very small part of Queensland (see van Dyck and Strahan 2008). *Bettongia pusilla*, known exclusively from Holocene Nullarbor Plain cave accumulations (McNamara 1997), may have become extinct prior to European colonization of Australia. However, given the isolated nature of its predicted range, it is possible it persisted unnoticed until the arrival of Europeans before succumbing to the combined impacts of European-led habitat destruction and introduction of exotic predators and competitors. The evidence for such extinction pressures are stronger for *B. anhydra*, given that it survived in central Australia well into the 1930s. Its disappearance, along with many other small- to medium-sized Australian mammals, coincides with the 4th toll of the post-European mammal extinction bell (Johnson 2006:171–172).

All species of *Bettongia* appear to be highly sensitive to anthropogenically driven environmental change. The recognition of yet another recently extinct Australian mammal suggests that the extent of Australian biodiversity loss since Europeans settlement may be greater than previously thought. This research also highlights the potential that numerous cryptic or rare species may remain hidden among their more common, morphologically similar relatives. However, the loss of ecological services that accompanied the extirpation or extinction of bettongs and other digging mammals may be of greater ecological importance. Given the unlikelihood that any extant species of bettong will be restored to its former range and abundance, the loss of these landscape engineers and the ecological services that they once performed will make the restoration of Australia's pre-European ecology all the more difficult.

Acknowledgments

We thank C. Johnson and C. Pavey for constructive comments on the submitted manuscript; C. Kemper, D. Stemmer, and M-A. Binnie for providing access to the mammal and palaeontology collections of the South Australian Museum; M. Siversson for access to the palaeontology collection of the Western Australian Museum; and T. Reardon of the South Australian Museum for providing access to and help with photographic equipment. MB and GJP are supported by Australian Research Council Future Fellowship (FT0991741 and FT130101728, respectively).

Literature Cited

Burbidge, A. A., J. Short, and P. J. Fuller. 2007. Relict *Bettongia lesueur* warrens in Western Australian deserts. Australian Zoologist 34:97–103.

Calaby, J. H., and B. J. Richardson. 1988. Potoroidae. Pp. 53–59 in Zoological catalogue of Australia. 5. Marsupialia (D. W. Walton, ed.). Australian Government Publishing Service, Canberra, Australia.

Claridge, A. W., J. H. Seebeck, and R. Rose. 2007. Bettongs, potoroos and the musky rat-kangaroo. CSIRO Publishing, Collingwood, Australia.

Desmarest, A. G. (1822). Description of Kangurus gaimardi. Mammalogie ou description des espèces de mammifères. Pp. 542 in Encyclopédie Methodique (Dictionnaire Encyclopédique Méthodique) (A. G. Desmarest, ed.). Supplement ii V. Agasse, Paris, France.

Drummond, A., et al. 2011. Geneious v5. 4. http://www.geneious.com/. Accessed 21 January 2012.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969–1973.

Eldridge, D. J., and A. I. James. 2009. Soil-disturbance by native animals plays a critical role in maintaining healthy Australian landscapes. Ecological Management and Restoration 10:S27–S34.

Eldridge, D. J., T. B. Koen, A. Killgore, N. Huang, and W. G. Whitford. 2012. Animal foraging as a mechanism for sediment movement and soil nutrient development: evidence from the semi-arid Australian woodlands and the Chihuahuan Desert. Geomorphology 157:131–141.

Finlayson, H. H. 1932. *Caloprymnus campestris*. Its recurrence and characters. Transactions of the Royal Society of South Australia 56:148–167.

Finlayson, H. H. 1957. Preliminary description of two new forms of *Bettongia* (Marsupialia). The Annals and Magazine of Natural History 10:552–554.

Finlayson, H. H. 1958. On central Australian mammals (with notice of related species from adjacent tracts). Part III – the Potoroinae. Records of the South Australian Museum 13:235–302.

Flannery, T. F., and M. Archer. 1987. *Bettongia moyesi*, a new and plesiomorphic kangaroo (Marsupialia: Potoroidae) from Miocene sediments of northwestern Queensland. Pp. 759–767 in Possums and opossums: studies in evolution (M. Archer, ed.). Surrey Beatty and Sons, Sydney, Australia.

Fleming, P. A., H. Anderson, A. S. Prendergast, M. R. Bretz, L. E. Valentine, and G. E. Hardy. 2013. Is the loss of Australian

digging mammals contributing to a deterioration in ecosystem function? Mammal Review 44:94–108.

Francescoli, G., V. Quirici, and R. Sobrero 2012. Patterns of variation in the tympanic bulla of tuco-tucos (Rodentia, Ctenomyidae, *Ctenomys*). Acta Theriologica 57:153–163.

Garkaklis, M. J., J. S. Bradley, and R. D. Wooller. 1998. The effects of Woylie (*Bettongia penicillata*) foraging on soil water repellency and water infiltration in heavy textured soils in southwestern Australia. Australian Journal of Ecology 23:492–496.

Gray, J. E. 1821. On the natural arrangement of vertebrose animals. London Medical Repository 15:296–310.

Gray, J. E. 1837. Description of some new or little known Mammalia, principally in the British Museum collection. Magazine of Natural History 1:577–587.

Haouchar, D., J. Haile, P. B. S. Spencer, and M. Bunce. 2013. The identity of the Depuch Island rock-wallaby revealed through ancient DNA. Australian Mammalogy 35:101–106.

James, A. I., D. J. Eldridge, T. B. Koen, and K. E. Moseby. 2011. Can the invasive European rabbit (*Oryctolagus cuniculus*) assume the soil engineering role of locally-extinct natives? Biological Invasions 13:3027–3038.

Johnson, C. 2006. Australia's mammal extinctions: a 50,000 year history. Cambridge University Press, Melbourne, Australia.

Johnson, C., and A. McIlwee. 1997. Ecology of the northern bettong, *Bettongia tropica*, a tropical mycophagist. Wildlife Research 24:549–559.

Martin, B. G. 2003. The role of small ground-foraging mammals in topsoil health and biodiversity: implications to management and restoration. Ecological Management and Restoration 4:114–119.

McDowell, M. C. 2014. Holocene vertebrate fossils aid the management and restoration of Australian ecosystems. Ecological Management and Restoration 15:58–63.

McDowell, M. C., A. Baynes, G. C. Medlin, and G. J. Prideaux. 2012. The impact of European colonization on the late-Holocene non-volant mammals of Yorke Peninsula, South Australia. The Holocene 22:1441–1450.

McNamara, J. A. 1997. Some smaller macropod fossils of South Australia. Proceedings of the Linnean Society of New South Wales 117:97–106.

Mora, M. A., I. Olivares, and A. I. Vassallo. 2003. Size, shape and structural versatility of the skull of the subterranean rodent *Ctenomys* (Rodentia, Caviomorpha): functional and morphological analysis. Biological Journal of the Linnean Society 78:85–96.

Owen, R. 1866. On the anatomy of vertebrates. Longmans, Green and Co., London, England.

Pergams, O. R. W., and J. J. Lawler. 2009. Recent and widespread rapid morphological change in rodents. PloS One 4:e6452.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25:1253–1256.

Prideaux, G. J. 1999. *Borungaboodie hatcheri* gen. et sp. nov., a very large bettong (Marsupialia: Macropodoidea) from the Pleistocene of southwestern Australia. Records of the Western Australian Museum Supplement No. 57:317–329.

Prideaux, G. J. 2004. Systematics and evolution of the sthenurine kangaroos. University of California Publications in Geological Sciences 146:1–623.

Prideaux, G. J., and N. M. Warburton. 2010. An osteology-based appraisal of the phylogeny and evolution of kangaroos and wallabies (Macropodidae: Marsupialia). Zoological Journal of the Linnean Society 159:954–987.

Quoy, J. R. C., and J. P. Gaimard. 1824. Zoologie. Pp. 712 in de Voyage Autour du Monde, ... exécuté sur les corvettes de S.M. 1'Uranie et la Physicienne pendant les années 1817, 1818, 1819 et 1820. Vol. 1 iv. (L. C. D. Freycinet, ed.). Chez Pillet Aîne, Paris, France.

Rambaut, A. 2007. FigTree, a graphical viewer of phylogenetic trees. http://tree.bio.ed.ac.uk/software/figtree. Accessed 21 January 2012.

Rambaut, A., and A. Drummond 2009. Tracer version 1.5. 0. http://tree.bio.ed.ac.uk/software/tracer/. Accessed 21 January 2012.

Seebeck, J. H., and R. W. Rose. 1989. Potoroidae. Pp. 716–739 in Fauna of Australia 1B Mammalia (D. W. Walton and B. J. Richardson, eds.). Australian Government Publishing Service, Canberra, Australia.

Short, J. 1998. The extinction of rat-kangaroos (Marsupialia: Potoroidae) in New South Wales, Australia. Biological Conservation 86:365–377.

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

Trappe, J. M., A. W. Claridge, D. L. Claridge, and L. Liddle. 2008. Desert truffles of the Australian outback: ecology, ethnomycology, and taxonomy. Economic Botany 62:497–506.

van Dyck, S., and R. Strahan. 2008. The mammals of Australia. 3rd ed. Reed New Holland, Sydney, Australia.

Vitousek, P. M. 1990. Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies. Oikos 57:7–13.

Wakefield, N. A. 1967. Some taxonomic revision in the Australian marsupial genus *Bettongia* (Macropodidae), with description of a new species. Victorian Naturalist 84:8–22.

Wilson, L. A. B., and M. R. Sánchez-Villagra. 2010. Diversity trends and their ontogenetic basis: an exploration of allometric disparity in rodents. Proceedings of the Royal Society of London, B. Biological Sciences. 277:1227–1234.

Wisely, S. M., J. E. Maldonado, and R. C. Fleischer. 2004. A technique for sampling ancient DNA that minimizes damage to museum specimens. Conservation Genetics 5:105–107.

Submitted 6 June 2014. Accepted 5 October 2014.

Associate Editor was Chris R. Pavey.

Appendix I

Specimens used for comparison with *Bettongia anhydra*. Bold denotes specimens that were sympatric with *B. anhydra* specimens.

Bettongia gaimardi Desmarest, 1822: SAM M7386, M7387, M7388.

Bettongia lesueur Quoy and Gaimard, 1824: FUR 034 (Holocene fossils from Corra-Lynn Cave, Yorke Peninsula, South Australia); SAM M1702, M10769.

Bettongia penicillata Gray, 1837: FUR 011, 031; SAM M6211, M11247; WAM 66.1.7c, 66.6.58, 66.12.4a, 66.12.4b, 66.12.11, 67.3.54, 67.3.98, 67.5.1, 67.5.4, 67.5.41–.44, 67.8.69, 67.10.193,

67.10.194, 67.10.31, 67.10.324, 67.10.360, 67.10.37, **67.10.523-.525**, 68.2.90, 68.2.91, 68.3.17, 69.7.649, 69.7.655, 69.7.661, 70.5.22, 70.5.23, 71.9.36, 72.1.109, 72.1.139, 72.1.199, 72.1.224, 72.1.467, 72.1.489a, 72.1.489b, 72.1.634a, 72.1.634b, 72.1.676, 72.1.776, 72.1.777, 72.1.778, 72.1.800, 72.1.824, 72.1.845-.847, 72.1.898a, 72.1.898b, 72.1.924, 72.1.1104, 72.1.1105, 72.6. 184, 75.12.21, 76.4.35.

Bettongia pusilla McNamara, 1997: SAM P35442, P35446, P35450, P35451; WAM 67.10.227, 67.10.412, 68.3.5, 71.1.29a, 72.1.108, 72.1.822, 72.1.823, 76.10.413.

Bettongia tropica Wakefield, 1967: MV C6870, AMNH 65279, QM M10030.

APPENDIX D – GENETIC DIVERSITY LOSS IN A BIODIVERSITY HOTSPOT: ANCIENT DNA QUANTIFIES GENETIC DECLINE AND FORMER CONNECTIVITY IN A CRITICALLY ENDANGERED MARSUPIAL.

The following citation resulted from research that is of relevance to this doctoral thesis and was published during my candidature. In context of this multi-author publication I constructed Figure 1, showing localities of woylie fossils and museum specimens sampled. I also reconstructed Figure 4, showing the Bayesian phylogeny of *Bettongia penicillata ogilbyi.*

A copy of the main article is attached. The Supplementary material may be viewed online at: http://onlinelibrary.wiley.com/doi/10.1111/mec.13430/full

Pacioni, C., Hunt, H., Allentoft, M.E., Vaughan, T.G., Wayne, A.F., Baynes, A., **Haouchar, D**., Dortch, J., Bunce, M., 2015. Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. *Molecular Ecology* **24**, 5813-5828.

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 5813–5828 doi: 10.1111/mec.13430

Genetic diversity loss in a biodiversity hotspot: ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial

CARLO PACIONI,* HELEN HUNT,* MORTEN E. ALLENTOFT,*† TIMOTHY G. VAUGHAN,‡ ADRIAN F. WAYNE,§ ALEXANDER BAYNES,¶ DALAL HAOUCHAR,* JOE DORTCH** and MICHAEL BUNCE*† †

*Ancient DNA Laboratory, School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia, †Centre for GeoGenetics, Natural History Museum, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen K, Denmark, ‡Department of Computer Science, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand, §Department of Parks and Wildlife, Manjimup, WA 6258, Australia, ¶Western Australian Museum, Locked Bag 49, Welshpool DC, WA 6986, Australia, **Archaeology, M257, The University of Western Australia, 35 Stirling Highway, Nedlands, WA 6009, Australia, ††Department of Environment and Agriculture, Trace and Environmental DNA Laboratory, Kent Street, Bentley, Perth, WA 6845, Australia

Abstract

The extent of genetic diversity loss and former connectivity between fragmented populations are often unknown factors when studying endangered species. While genetic techniques are commonly applied in extant populations to assess temporal and spatial demographic changes, it is no substitute for directly measuring past diversity using ancient DNA (aDNA). We analysed both mitochondrial DNA (mtDNA) and nuclear microsatellite loci from 64 historical fossil and skin samples of the critically endangered Western Australian woylie (*Bettongia penicillata ogilbyi*), and compared them with 231 (n = 152 for mtDNA) modern samples. In modern woylie populations 15 mitochondrial control region (CR) haplotypes were identified. Interestingly, mtDNA CR data from only 29 historical samples demonstrated 15 previously unknown haplotypes and detected an extinct divergent clade. Through modelling, we estimated the loss of CR mtDNA diversity to be between 46% and 91% and estimated this to have occurred in the past 2000–4000 years in association with a dramatic population decline. In addition, we obtained near-complete 11-loci microsatellite profiles from 21 historical samples. In agreement with the mtDNA data, a number of 'new' microsatellite alleles was only detected in the historical populations despite extensive modern sampling, indicating a nuclear genetic diversity loss >20%. Calculations of genetic diversity (heterozygosity and allelic rarefaction) showed that these were significantly higher in the past and that there was a high degree of gene flow across the woylie's historical range. These findings have an immediate impact on how the extant populations are managed and we recommend the implementation of an assisted migration programme to prevent further loss of genetic diversity. Our study demonstrates the value of integrating aDNA data into current-day conservation strategies.

Keywords: ancient DNA, *Bettongia*, biodiversity loss, genetic diversity, population bottleneck

Received 22 May 2014; revision received 7 October 2015; accepted 13 October 2015

Introduction

Species and populations that are adversely affected by anthropogenic activities, including habitat modification, are a global phenomenon (Andrén 1994; Ewers &

Correspondence: Michael Bunce, Fax: +61 8 9266 2495; E-mail: michael.bunce@curtin.edu.au

Didham 2006). The reduction in genetic diversity, often associated with population bottlenecks and habitat fragmentation, can result in the loss of reproductive fitness and reduction in evolutionary potential (Frankham 1996; Frankham *et al.* 1999). Genetic bottlenecks are most commonly studied using modern molecular data and theoretical models are applied to determine their likelihood and severity (Cornuet & Luikart 1996; Garza & Williamson 2001). The incorporation of time-stamped ancient DNA (aDNA) data, for example obtained from museum skins and fossils, can provide direct insight into a population's past diversity (Ramakrishnan & Hadly 2009). The demographic history of bison (Shapiro *et al.* 2004), musk ox (Campos *et al.* 2010) and the collared lemming (Brace *et al.* 2012), for example have been extensively studied using mitochondrial DNA (mtDNA) isolated from preserved fossil material. However, the use of aDNA data to make informed conservation decisions is rare, despite the fact that it can provide vital information regarding past population demography (e.g. bottlenecks), former connectivity and extent of genetic diversity (see Leonard 2008 for a review).

An Australian example of a species adversely affected by population declines is the woylie (or brush-tailed bettong, *Bettongia penicillata ogilbyi*). Woylies are critically endangered, nocturnal marsupials (Wayne *et al.* 2008a; Groom 2010), that feed primarily on hypogeal fungi (native truffles), tubers, bulbs, seeds and invertebrates, and they weigh between 1.1 and 1.6 kg (Van Dyck & Strahan 2008). With the exception of five translocated sites (i.e. not naturally occurring) in South Australia and New South Wales, the distribution of the woylie is currently limited to the south-west region of Western Australia (WA), one of only 34 recognized biodiversity hotspots on the planet (Myers *et al.* 2000). Historically, together with the subspecies *B. p. penicillata*, this species had a widespread distribution across southern Australia, but since European settlement (in the 18th century), it has undergone a substantial population decline (Fig. 1), attributed primarily to the introduction of feral predators, particularly red foxes (*Vulpes vulpes*) (Burbidge & McKenzie 1989; Start *et al.* 1998; Woinarski *et al.* 2015). Habitat fragmentation and changes in fire regimes have also been suggested as contributing factors (Groom 2010). An example of the extent of land clearing in the woylie's former range is the 'Wheat-belt' region, an area of 95 800 km^2 about the size of Portugal, where only ~7% of the original vegetation remains in a number of small isolated pockets (Saunders 1979).

Despite a spectacular recovery in the 1990s due to the intensification of the control of the red fox (Start *et al.* 1998), the woylie has undergone a second, more recent, decline, resulting in a 90% reduction in 7 years from a

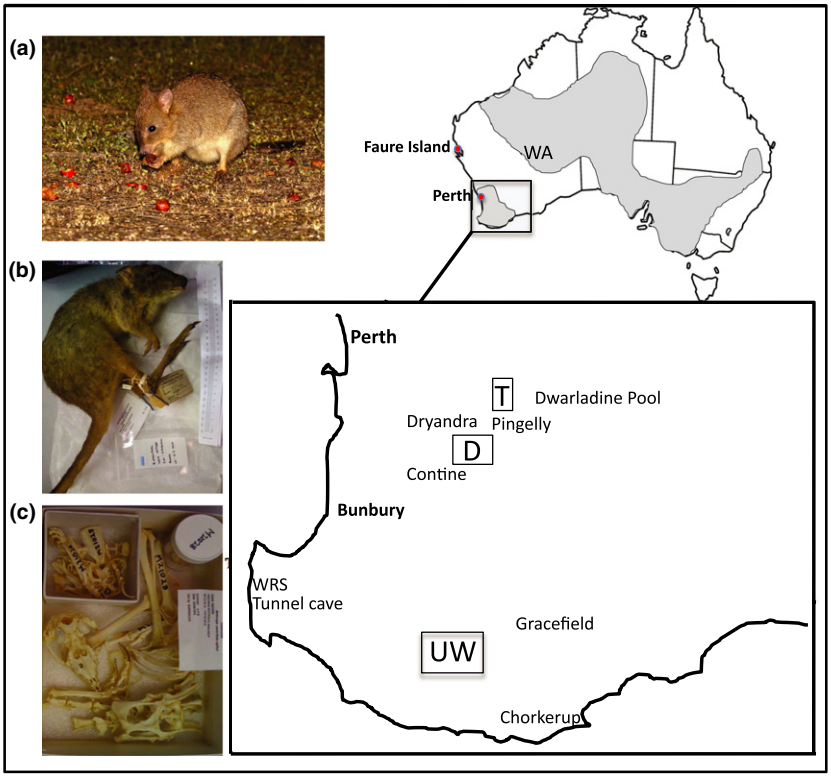


Fig. 1 Localities of historical woylie fossil and museum specimens sampled in south-west Western Australia (WRS: Witchcliffe Rock Shelter). Modern populations are indicated in capital letters within grey squares: UW = Upper Warren, D = Dryandra, T = Tutanning). Grey area in the insert: former *Bettongia penicillata* distribution including both subspecies *B. penicillata ogilbyi* and *B. p. penicillata* (adapted from Wayne *et al.* 2008b) based on contemporary, fossils and Aboriginal knowledge records (Burbidge *et al.* 1988). Images of the woylie (*B. p. ogilbyi*: (a) an adult woylie fitted with a VHF radio collar (Photo: John Lawson), (b) skin and (c) fossil bone remains).

peak of about 200 000 animals in 1999 (Wayne *et al.* 2013). The factors responsible for this decline are still not well defined (Groom 2010; Wayne *et al.* 2015). Recent genetic studies of the extant populations using microsatellite data from 231 live individuals revealed four genetically distinct wild populations (Pacioni *et al.* 2011). Two of these are found in the highly fragmented Wheat-belt region: Dryandra woodland and Tutanning Nature Reserve and the other two within the Upper Warren region; Kingston and Perup (Fig. 1). Current size estimates in these naturally occurring populations (excluding translocated populations) are approximately <10 000, with Tutanning Nature Reserve having gone extinct in the wild in 2011 (Wayne *et al.* 2013; Department of Parks and Wildlife, unpublished data). Woylie translocated populations account for another ~4000 individuals (Wayne *et al.* 2013).

An evaluation of contemporary genetic diversity among 152 modern individuals from the four remnant populations, using the mitochondrial control region (CR), revealed 15 haplotypes and expected heterozygosity (H_E, $n = 231$) values of 63–83% for 12 microsatellite loci (Pacioni *et al.* 2011). However, it remains unknown to what extent genetic diversity has been lost since the collapse of the population. The use of heterozygosity values calculated only from modern samples as a measure of genetic health can be problematic due to the lack of comparative baseline data. A direct assessment of genetic diversity using both pre and postdecline samples is the only way to securely quantify genetic diversity changes over time with any degree of accuracy. An appreciation for temporal effects is often a missing factor when making conservation decisions, the woylies being a case in point where, for example the lack of baseline data makes the establishment of conservation targets difficult (Pacioni *et al.* 2013). In this study, we used the woylie as an example to investigate to what extent the application of aDNA techniques can improve our understanding of the dynamics associated with wildlife declines. We characterized two genetic markers: mitochondrial DNA (CR and cytochrome *b*) and microsatellites, and assessed the possible loss of the genetic diversity and the degree of historical gene flow between regions. We further discuss how the information obtained can inform conservation decisions.

Materials and methods

Skin and bone sampling

A total of 64 historical samples were sourced from the collections of the Western Australian Museum (WAM). Samples included fossil bone fragments from vertebrate archaeological collections ($n = 15$), museum skins ($n = 28$) and bone material from the vertebrate palaeontological collection ($n = 21$) ranging between *c.* 40 and 15 600 years old (Dortch 2004a,b; Table S1, Supporting information). Samples represent a wide geographical area, with coverage across woylie former distribution (Fig. 1). The entire WAM woylie skin collection was sampled for this study. Woylie fossils, from across southern WA, were selected from sites such as caves that are typically conducive to long-term DNA preservation.

DNA extraction

DNA extractions from historical samples were carried out in a dedicated aDNA laboratory located at Murdoch University (Perth, Australia), minimizing contamination from PCR amplicons and modern DNA in accordance with established aDNA guidelines (Gilbert *et al.* 2005). Skin samples with attached hair (~5 × 5 mm) were diced using a sterile scalpel blade and incubated overnight at 55 °C with rotation, in 500 μL of digest buffer (20 mM Tris; pH 8, 1% SDS, 5 mM $CaCl_2$, 10 mM dTT, and 10 mM EDTA) with the addition of 10 mg of Proteinase K (Invitrogen). DNA was bound and eluted from the digest buffer using DNeasy blood and tissue kit according to the manufacturer's recommendations (Qiagen). Samples of bone and teeth were grounded to powder using a Dremel tool (part no. 114; Germany) set at a low speed. Approximately 100 mg of bone powder was incubated in a bone digest buffer (containing; 20 mM Tris; pH 8, 1% SDS, 10 mM dTT, 0.48 M EDTA and 10 mg of Proteinase K) overnight at 55 °C with rotation. Following digestion, all tissue samples were centrifuged briefly to pellet debris and the supernatant transferred to a Vivaspin 500 tube with a 30 000 Dalton molecular weight cut-off membrane (Sartorius Stedim Biotech) and centrifuged at 15 800 *g*, until ~50 μL of supernatant remained. This was then mixed with five volumes of PBI buffer (Qiagen) and DNA was extracted using a DNeasy tissue and blood kit (Qiagen) according to the manufacturer's recommendations.

DNA quantification

Following extraction, each sample was assayed for DNA preservation using quantitative PCR (qPCR) assays. The ability to successfully genotype nuclear aDNA loci is largely dependent on template copy number (Allentoft *et al.* 2011), accordingly qPCR assays represent an important step in determining relative preservation across a set of samples such as collected here.

The mtDNA primers targeting the 12s rRNA region (12s woylie macro 302F, primer sequence

CGTAAAGCGTGTTTAAGCC and 12s woylie macro 425R, primer sequence CTGTAGTGTATTCAGCAAA; PCR product ~120 bp) were tested and shown to function well in a SYBR-bead qPCR assay and therefore used to screen all samples. The PCR reagents and protocol were optimized with a final 25 μL PCR reaction consisting of 2 μL template DNA, 3.0 mM $MgCl_2$ (Fisher Biotech, FB), 1× PCR buffer (FB), 0.5 mM each dNTP's (Astral Scientific), 0.16 μmoles of each primer (12s 302F/425R), 0.6 μL of SYBR Green dye (Stock 1/10 000 dilution; Invitrogen), 0.25 U AmpliTaq Gold polymerase (Life Technologies) and 1 μL (10 mg/mL) bovine serum albumin (BSA) (Fisher Biotech). PCR conditions were an initial hot start at 95 °C for 5 min followed by 40 cycles of 95°C for 45 s, 57 °C for 45 s and 72 °C for 45 s. The same qPCR assay was used to assess the relative quantity of DNA in all the historical DNA extracts, to establish the relative copy number of each sample. While many of the 'failed' samples still yielded DNA (Table S1, Supporting information), they were excluded from the analyses because of problems associated with allelic dropout of microsatellite loci—a phenomenon commonly encountered with degraded DNA (Taylor *et al.* 1994; Bouzat *et al.* 1998; Allentoft *et al.* 2011).

DNA amplification

A ~400 bp fragment of the CR and a 370 bp fragment of the cytochrome *b* gene were amplified using primers designed to target-specific mtDNA regions (Table S2, Supporting information). Reactions typically consisted of a 25 μL reaction containing 2 μL template DNA, 2 mM $MgSO_4$, 1× PCR buffer, 0.25 mM each dNTP's, 8 ρmoles of each primer, 0.25 U HiFi Taq polymerase (Invitrogen), 1 μL (10 mg/mL) BSA and template DNA. PCR conditions were as follows: a hot start at 95 °C for 5 min followed by 40 cycles of 95 °C for 45 s, 56/58 °C (depending on primer set used) for 45 s and 68 °C for 45 s. DNA was sequenced in both directions using dye terminator cycle sequencing chemistry (3730xl sequencer; Applied Biosystems via Macrogen—BigDye ver 3.1). Sequences were aligned in GENEIOUS PRO (BioMatters). PCR repetitions using varying amounts of input DNA were conducted on individuals with unique mutations to check sequence fidelity. In addition, those fragments with unique mutations were cloned using the p-GEM 'TA' vector system (Promega) to verify these rare variants and to investigate whether they could be accounted for by post-mortem DNA damage (primarily the deamination of cytosine, causing a C to T change during PCR amplification) (Willerslev & Cooper 2005; Briggs *et al.* 2007).

The microsatellite loci targeted in this study were previously used to characterize modern woylie genetic diversity (Table S3, Supporting information; Pacioni & Spencer 2010), with the exclusion of Y151 due to large variation in amplicon product size at this locus. All microsatellite amplifications were conducted as described above except that Amplitaq Gold polymerase (Applied Biosystems) was used and 1.25 mM $MgCl_2$. Primers for two loci (Y175 and Bt76) were redesigned closer to the microsatellite repeating unit to shorten the allele length and reduce allelic dropout.

DNA fragment analysis of the amplified microsatellite products was carried out using a 5-dye system on an Applied Biosystems 3730 DNA analyser. The PCR product size was determined using an internal size standard (Genescan TM-500 Liz; Applied Biosystems) and fragments were scored using GENEMARKER software (v1.5; Soft Genetics). Genotyping was repeated at least three times for homozygotes, with the addition of different starting amounts of DNA and were accepted only when all three repeats were homozygous for the same allele and, similarly, heterozygotes were repeated twice and accepted when both were heterozygous for the same alleles. We calculated the frequencies of falsely scoring an individual as homozygous ($P_{(fh)}$, Gagneux *et al.* 1997) and then the probability of falsely scoring an individual (per locus) given our protocols (i.e. $P_{(fh)}{}^3$). Two modern samples were co-run in each PCR to ensure direct comparability between the modern and historical data sets, and the allele scores of microsatellites from redesigned primer pairs were adjusted accordingly.

Mitochondrial DNA data analysis

Haplotypes from the historical samples were identified with DNAsp (Librado & Rozas 2009), which was also used to calculate haplotype (h) and nucleotide diversity (π) (Nei 1987). We applied the Stirling probability distribution to calculate the posterior probability of the expected total number of (historical) haplotypes (Dixon 2006). As a result, we calculated the probability that additional haplotypes were yet to be sampled, and estimated the expected total number of historical haplotypes, based on the observations and sample size of the historical data. Sequences from modern data with complete coverage of the region amplified from historical samples ($n = 146$) were obtained from Pacioni *et al.* (2011), and their h and π values were compared with those obtained from historical samples using a t-test. Rarefaction curves from both historical and modern data sets were calculated using Vegan (Oksanen *et al.* 2013) in R v2.15 (R Development Core Team 2015) and plotted to allow graphical comparison of haplotype richness between the two data sets. The haplotype richness obtained with the rarefaction approach allows for a direct comparison between the 'modern' and 'ancient'

data sets because it standardizes the sampling effort. To further explore the expected number of historical haplotypes, we fitted the two- and three-parameter asymptotic exponential models to the haplotype richness data, compared them with the *F*-test, and then predicted, using the most supported model, the number of historical haplotypes that we would have obtained had we had available the same number of samples as for the modern data. Spatial-temporal changes in the mtDNA genetic diversity were also visualized with a multidimensional parsimony network built in TempNet v1.8 (Prost & Anderson 2011).

The CR and cytochrome *b* sequence alignments were analysed with the Bayesian methods implemented in BEAST v1.5.3 (Drummond & Rambaut 2007) to obtain the phylogenetic relationships. The general time-reversible (GTR) model of nucleotide substitution with invariant and gamma sites (I + G), was used as determined by MODELTEST v3.7 (Posada & Crandall 1998). A strict molecular clock was assumed. The northern bettong (*Bettongia tropica*, GenBank accession numbers AF287895–AF287906) is a sister taxon to *Bettongia penicillata* and was therefore used to root the tree. To place the woylie radiation in a temporal context, the cytochrome *b* sequences, along with other marsupial sequences (GenBank accession numbers in Table S4, Supporting information), were modelled in BEAST using the GTR + I + G substitution model, assuming a strict molecular clock and Yule process speciation, and enforcing two published calibrations (Westerman *et al.* 2004). A calibration of 23.8 million years (95% CI 23.0–25.9), for the divergence between the Macropodinae and the Potoroinae, and of 8.3 million years (95% CI 5–12) for divergence of *Bettongia* were imposed. Three runs of ten million generations (sampling every 1000 trees) were conducted and combined using LOGCOMBINER v1.5.3. After a burn in of 1000 trees per run, a consensus tree was generated using TREEANNOTATOR v1.5.3.

Next, we evaluated changes of woylie population size over time using coalescent-based approaches as these have proven efficient in detecting demographic changes (Peery *et al.* 2012). We analysed the woylie CR mtDNA data using BEAST v.1.8 (Drummond *et al.* 2012) running 200 million iterations and discarding the first 10% as burn in. We used the estimated dates of the aDNA samples and calibrated the root with a lognormal prior using the results of the analysis of the cytochrome *b* data set described above. We used a lognormal prior also for the clock rate while leaving default options for the remaining priors. Using the Akaike information criterion through Markov chain Monte Carlo (AICM, Baele *et al.* 2012) with 1000 bootstraps, we compared three demographic models: constant population size, exponential growth and Bayesian Skyline Plot (BSP). Appropriate mixing and estimated sample size in the analyses was checked with TRACER v.1.6 (Rambaut & Drummond 2007).

We also used a structured coalescent sampler to analyse these data to account for potential problems arising from population structuring (Navascués *et al.* 2010; Ho & Shapiro 2011). The historical data for these analyses were limited to samples that were geographically close to the extant populations ($n = 13$). Using MIGRATE-N v3.6.8 (Beerli 2006), we compared migration models of different levels of complexity: we started from a full model where the four populations (Tutanning, Dryandra, Kingston and Perup) were kept separate. Because of their relative geographical position, direct migration from Tutanning to any of the Upper Warren populations (and vice versa) was prevented (i.e. animals from Tutanning could only migrate to Dryandra and vice versa). We considered two-three-population models where in one the two populations from the Wheat-belt region (Dryandra and Tutanning) were merged into one population and Perup and Kingston kept separate and, in the second, the Perup and Kingston were merged and the Wheat-belt populations were kept separate. We further simplified the migration model by considering only two populations (one for each region) and lastly we simulated a panmictic population. These analyses were conducted running 10 replicates of 400 million iterations, sampling every 2000 trees and discarding 30% of the trees as burn-in. Each analysis ran four chains with the default static heating scheme using slice sampling (Neal 2003) and with a gamma prior for both population parameters (θ: 0.2, 0.06, M: 500, 0.5). Eventually, models were compared using the log Bayes Factor (LBF) calculated with the Bezier marginal likelihoods obtained by thermodynamic integration. Once the migration model was selected, we used MIGRATE-N to generate BSPs. Appropriate mixing and estimated sample size in MIGRATE-N analyses as well as postanalysis data plotting was carried out with the R package mtraceR (https://github.com/carlopacioni/mtraceR). To evaluate the extent of the bias introduced in BEAST where substructuring was not explicitly taken into account, we generated a BSP using BEAST with the same input file used for MIGRATE-N (i.e. limiting the historical samples to geographically close to the extent populations, $n = 13$) and graphically compared the results. Where relevant, we repeated the demographic analysis with modern data only to evaluate whether the addition of aDNA data had substantially changed our results.

Microsatellite data analysis

The microsatellite data do not strictly conform to the requirements for testing the Hardy–Weinberg Equilibrium,

as they do not represent a population in the same space and time (Frankham *et al.* 2002). Too few samples were available from one location during the same time to test meaningfully whether genotype frequencies were in Hardy–Weinberg proportions. Therefore, the data prevented identifying and correcting genotyping errors based on deviations from Hardy–Weinberg proportions, as well as deficiencies and excesses of particular genotypes using these standard approaches (e.g. Kalinowski 2006). However, allelic dropout was identified from PCR replications of historical samples.

Descriptive measures of genetic diversity were calculated in GENALEX v6.2 (Peakall & Smouse 2006), which included measures of allelic frequency per locus, observed heterozygosity (H_o) and observed number of alleles (N_A).

Historical data were compared with modern data obtained from Pacioni *et al.* (2011) to identify alleles that were not present in the modern data (unique alleles). Because differences in sampling intensity can bias comparison of genetic diversity, with data sets containing larger number of samples expected to contain more alleles (Kalinowski 2004), we used rarefaction to compensate for the differences in sampling effort and calculated the number of alleles present in an increasing sample size (from 1 to 20 for the historical data set and from 1 to 210 samples for the modern data) using POPGENKIT (Rioux Paquette 2011) in R v2.15 (R Development Core Team 2015) and plotted the rarefaction curves for each data set. Similar to what we did with the mtDNA data, we fitted the two- and three-parameter asymptotic exponential models to the haplotype richness data, compared them with the *F*-test, and then predicted, using the most supported model, the number of historical haplotypes that we would have obtained had we had available the same number of samples as for the modern data. Moreover, we compared allelic richness, obtained sampling 11 diploid individuals, between the modern and historical data with the non-parametric Wilcoxon signed-rank test (Kalinowski 2004). We also compared the individual *IH* between the two data sets with Mann–Whitney test. Furthermore, to investigate possible differences in connectivity between regions we estimated the number of migrants using the 'private allele' (the alleles found exclusively in one population) method (Barton & Slatkin 1986) and used GENEPOP v4.2 (Rousset 2008) for this purpose. We selected this method because its assumptions are not violated by heterochronous data. However, we acknowledge that the number of migrants estimated with this approach is possibly biased due to the fact that most private alleles would also be rare and, therefore, may be lost during population declines (i.e. the number of migrants calculated from the modern data set would be inflated). For the latter analysis, data from the two woylie populations in the Upper Warren region were merged because records did not specify whether historical samples were collected from the eastern or western compartment of the Upper Warren forest. One locus was omitted from all statistical comparisons: Y112, due to high rates of allelic dropout evident from PCR repetitions (see below). All univariate statistical analyses were performed with SPSS v19 (IBM Corp., Armonk, NY, USA) or R v2.15 (R Development Core Team 2015).

Evidence of population bottlenecks for each extant population was investigated, using modern microsatellite data, by testing for excess in heterozygosity (Cornuet & Luikart 1996) and mode-shift (Luikart & Cornuet 1998), with the programme BOTTLENECK (Piry *et al.* 1999). A mixed model of microsatellite mutation was assumed, with single-step mutations accounting for 95% of all mutation events and a variance among multiple steps of 12 (Piry *et al.* 1999). We used a Wilcoxon signed-rank test to test statistical significance (Piry *et al.* 1999). We also used the *M*-ratio method with M_P_VAL, using recommended mutation parameters: 0.2 for the proportion of one-step mutations (p_s), 3.5 as the average size of non-one-step mutations (Δ_g) and a mutation rate (μ) of 5×10^{-4}/locus/generation (Garza & Williamson 2001). It was assumed that the population census sizes (N_c) prior to European settlement were similar to those estimated at the peak density observed after the commencement of the fox control programme, immediately prior to the most recent declines (i.e. a conservative and best available estimate of pre-European populations) (Groom 2010). Following Frankham (1995), we approximated effective population size (N_e) to be 10% of N_c. Tutanning pre-European settlement N_e was assumed to be similar to the N_e at Dryandra. This test is anticipated to produce positive results for at least 100 generations postreduction (Garza & Williamson 2001).

Lastly, we used MIGRATE-N to estimate demographic changes over time after having determined an appropriate migration model with the same approach as outlined for the mtDNA data. Microsatellite data were analysed with an exponential prior (θ: 300; M: 1000) and running four replicates of 100 million iterations, sampling every 1250 trees and discarding 25% of the trees as burn-in. Each analysis ran four chains with the default static heating scheme using slice sampling (Neal 2003).

Results

DNA recovery and quantification

Of the 64 samples tested, including historical skins and fossils bones, DNA was successfully isolated from 49

samples (76.5% success, Table S1, Supporting information). However, we specifically focused on a subset of 29 well-preserved samples, as determined by relative qPCR cycle threshold (C_T) values (C_T is the number of cycles that are needed to obtained a predetermined fluorescent intensity and are inversely proportional to the initial amount of template used in the reactions), spanning the entire former range of the woylie. There were two reasons to focus on this subset of samples: first, we wanted to minimize the effect of microsatellite allelic drop out often encountered with low copy number DNA (Allentoft *et al.* 2011); second, samples with good DNA preservation are less prone to artefacts such as contamination and DNA damage.

Skin samples had relatively better DNA preservation than fossil bone, with a mean ΔC_T value of 4.24 or 18.9 times more DNA (25.55 ± 5.90 SE compared to 29.77 ± 5.79). Likewise, younger samples yielded more mtDNA quantities than older samples (Table S1, Supporting information; see also Hunt 2010). Although there was substantial variation in preservation, as expected (Allentoft *et al.* 2012), there was a general trend with older specimens having lower relative DNA concentrations, presumably due to a longer exposure to hydrolytic and enzymatic processes (Allentoft *et al.* 2012). Lastly, as previously observed (Allentoft *et al.* 2011), there was an overall trend between the C_T values and the percentage of microsatellite loci that could be typed (Fig. S1, Supporting information).

Mitochondrial DNA analysis

The lengths of the CR sequences obtained from the 29 historical samples varied between 250 and 400 bp (excluding primers). While all these were used for phylogenetic analysis (see below), we limited statistical analysis to 23 historical samples of ~400 bp (Table S1, Supporting information). Among these 23 historical samples (of which ~70% were around 100 years old or less, see Table S1, Supporting information), we identified a total of 18 haplotypes. Six samples possessed four CR haplotypes that are still present in extant populations. The remaining 15 haplotypes have not previously been observed despite intensive sampling of the modern woylie populations (Pacioni *et al.* 2011). These differences are further highlighted in the multidimensional network (Fig. 2).

Haplotype and nucleotide diversity of the sequences obtained from historical data was 0.976 (SD = 0.02) and 0.039 (SD = 0.004), which was substantially higher than any of the values calculated from the extant wild populations (range: h = 0.26–0.82; π = 0–0.022. Pacioni *et al.* 2013), and significantly higher than that obtained from

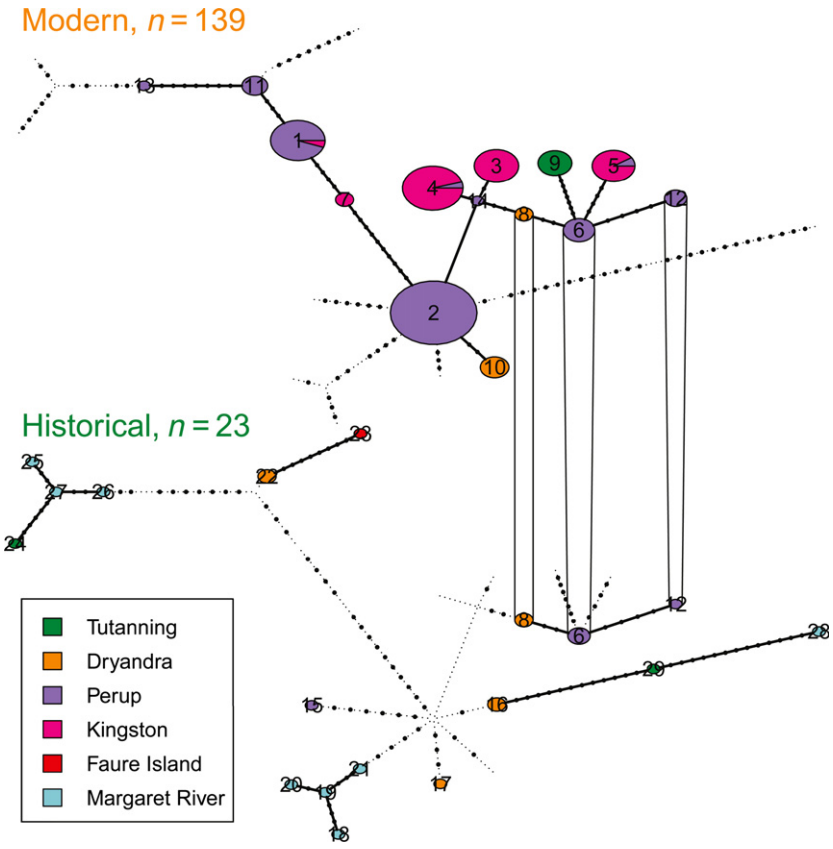


Fig. 2 Multidimensional parsimony network of mtDNA sequences from historical and modern woylie samples (after removal of sequences with ambiguities). Dimensions of the circles are proportional to the number of sequences for each haplotype. Margaret River is a region that includes samples from WRS, Mammoth Cave and Tunnel Cave.

the modern data set (h = 0.735, SD = 0.001, $P < 0.001$; π = 0.022, SD = 0.002, $P < 0.001$).

The rarefaction curve computed from the historical data set had a much higher trajectory and did not plateau, suggesting that there were many unsampled historical haplotypes (Fig. 3a). The three-parameter asymptotic exponential model was a significantly better fit ($P < 0.002$) than the two-parameter model and predicted an expected number of haplotypes equal to 42.4 if 146 historical samples had been used (equation: $n_h = 43.97–0.02\ e^{3.78\ S}$, where n_h is expected number of haplotypes, and S is sample size). Modelling the aDNA data with the Stirling probability distribution suggested that the number of historical haplotypes was 43, with the lower 95% confidence interval being 28 (95% CI: 28–173, Fig. 3b).

The phylogenetic reconstruction of the combined historical (n = 29) and modern data (Haplotypes A–K) revealed three highly supported (>98% posterior probability) clades within the woylie radiation (Fig. 4). These CR clades and haplotypes were also observed in the cytochrome *b* phylogeny confirming topological consistency between mtDNA regions (Fig. S2, Supporting information). Two historical haplotypes (samples MB816 and MB823) constituted a distinct clade, which was not present within the extensively sampled modern populations, suggesting that this particular clade is probably extinct. The other two clades consisted of an admixture of both modern and historical haplotypes and spanned a broad geographical area.

Molecular clock analysis based on the cytochrome *b* data set dated the most recent common ancestor of the western Australian *Bettongia penicillata* as living between 0.59 and 1.75 million years ago (mean = 1.2). This parameter was therefore included in the coalescent-based demographic analysis of CR with a lognormal prior of mean 1 000 000 and SD = 0.32 (in real space). Of the three demographic models analysed, the BSP was the model with the best fit according to differences in AICM estimates (constant population size: 2642.05, SE = 0.08; exponential growth: 2644.47, SE = 0.08; BSP: 2572.52, SE = 0.04). Based on this analysis, the woylie population size for 2006–2008 was 99.86% lower than the historical population size (Fig. 5, range: 99.3–99.97%). The mutation rate was estimated to be approximately 10% per million years, placing the beginning of the population decline between *c.* 4600 and 2300 (95% HPD) years ago.

The comparison of migration models with MIGRATE-N favoured the most complex model with four distinct populations (LBF = 5.9 with the next most likely model; Table 1), therefore we summed the parameter values to obtain the species-wide demographic changes over time. Taking into account migration between populations, woylie numbers declined by a 89.1% compared to historical population size (Fig. 6a, range: 87.1–90.7).

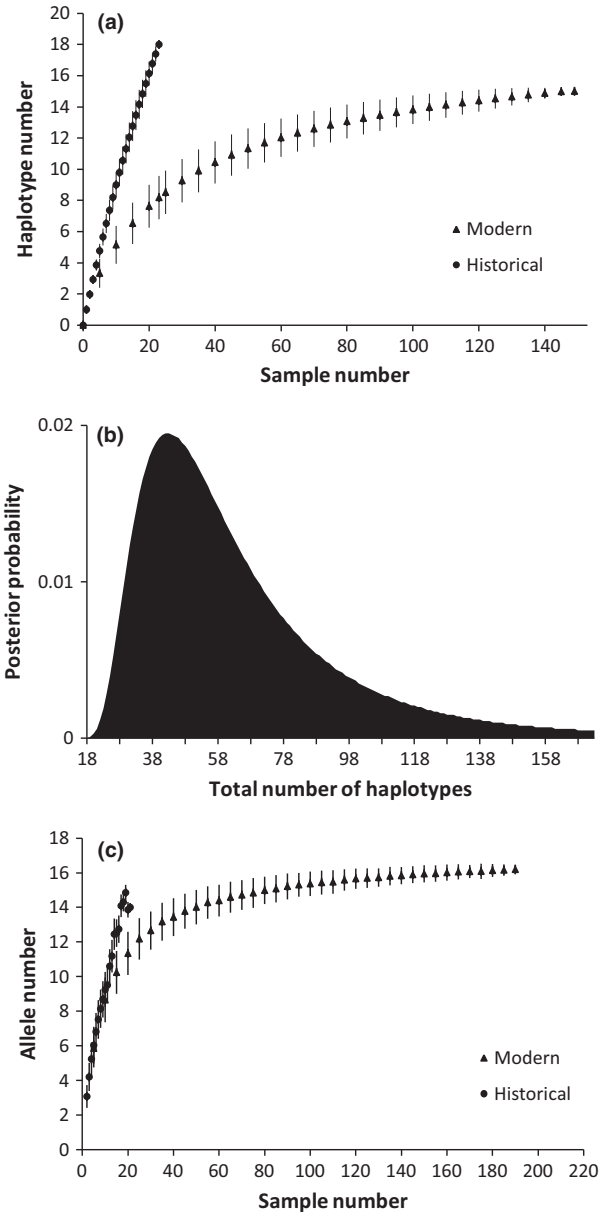


Fig. 3 (a, c) Rarefaction curves of modern and historical woylie data for (mtDNA) haplotype richness and (microsatellite) allelic richness respectively. Error bars indicate SE from re-sampling iterations. (b) Posterior probability distribution for expected total number of haplotypes based on observed number of haplotypes and sampling effort in the historical data set.

Microsatellite data analysis

We implemented several quantitative and qualitative measures to improve the reliability of the microsatellite data set. Following our strict approach, we only

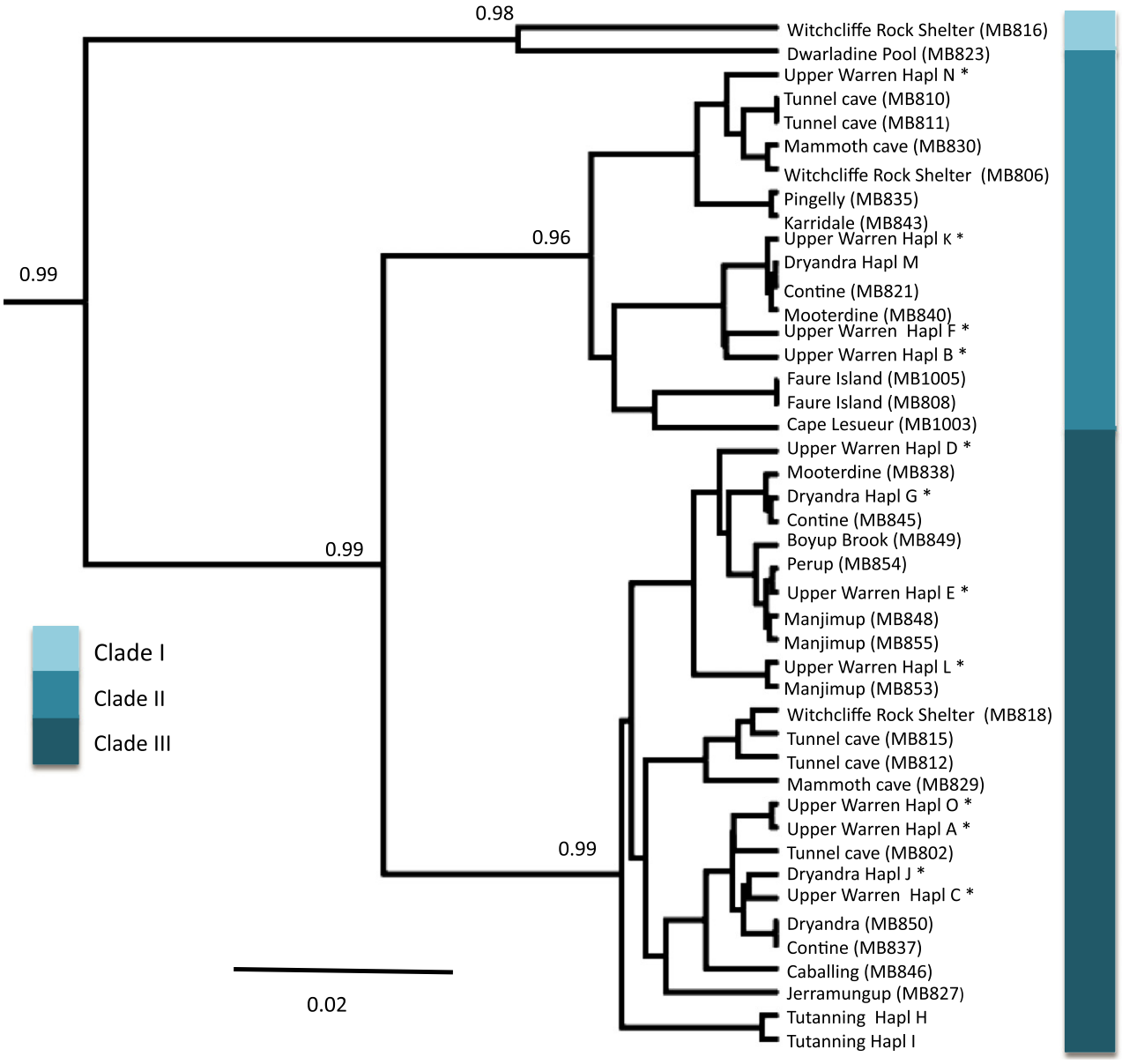


Fig. 4 Bayesian phylogeny of *Bettoniga penicillata ogilbyi*, showing the maximum a posteriori (MAP) tree generated using BEAST. A GTR + G + I substitution model was imposed on an alignment containing the 29 historical control region (CR) sequences (~400 bp) and 15 modern haplotype sequences (identified with an *) from Pacioni *et al.* (2011) detected in 152 modern samples (~600 bp). Posterior probabilities are shown on the selected outer nodes. The scale represents the number of nucleotide substitutions per site. Three distinct clades are indicated on the tree. The tree topology is mirrored at the cytochrome *b* locus (see Fig. S2, Supporting information).

considered 21 historical samples, from which we obtained near-complete 11-loci microsatellite profiles (Table S1, Supporting information). From the PCR repetitions, we calculate a $P_{(fh)}$ for the locus Y112 of 36.4% and it was therefore excluded from all statistical comparisons. The remaining loci had lower $P_{(fh)}$ (range: 0–16.3%, Table 2). Hence, our protocol had <0.5% probability of falsely scoring an individual homozygous (except for the locus Y112, for which it was ~5%). Despite the limited number of historical individuals with full microsatellite profiles we observed nine individuals with at least one 'unique allele' (i.e. alleles that were not present in modern data, Table 2). Levels of heterozygosity and average allelic richness (N_{AR}) varied across loci (Table 2) but were significantly higher in the historical data set (N_{AR}: Wilcoxon tests: $Z = -2.191$, $P = 0.028$; *IH*: Mann–Whitney test: $Z = -2.45$, $P = 0.014$, Table 2, Fig. 3c). The three-parameter

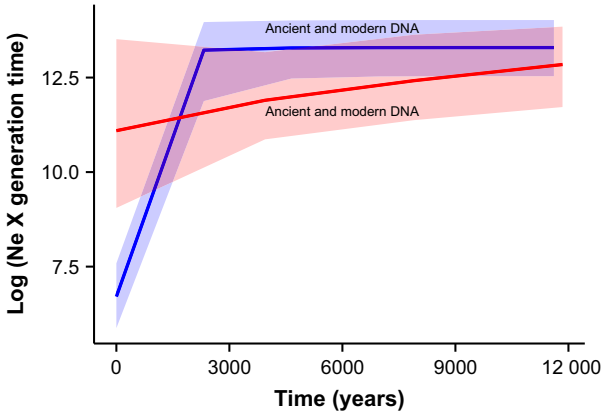


Fig. 5 Reconstruction of the woylie (*Bettongia penicillata ogilbyi*) demographic history based on control region of mtDNA using a Bayesian Skyline Plot analysis implemented in BEAST v1.8 (Drummond *et al.* 2012). Shaded areas represent 95% Highest Probability Density (HPD). Solid lines represent the medians. Population size is expressed as log of the effective population size times the generation times.

asymptotic exponential model was a marginally better fit ($P = 0.044$) to the rarefaction data than the two-parameter model and predicted an expected mean number of alleles equal to 20.7 if 210 historical samples had been used (equation: $n_a = 20.74–0.88\ e^{2.85\ S}$, where n_a is expected number of alleles, and S is sample size). The numbers of migrants were consistently higher when compared with estimates calculated with the modern data alone (Table 3).

No evidence of a bottleneck was detected using the heterozygosity excess and the mode-shift analytical approaches. The *M*-ratio method, however, detected a reduction in population size in all modern populations except for that from Perup, the eastern population in the Upper Warren region.

MIGRATE-N strongly favoured a panmictic population (LBF > 50 000 with the next most likely model). Because the commencement of the decline is less clear in MIGRATE-N microsatellite BSP (Fig. 6a), we calculated the extent of the decline from 2θ because accuracy may decrease beyond this point (P. Beerli, personal communication) and detected a 56.5% population decline in recent time.

Discussion

The aim of this study was to overlay data generated from historical samples on those from extensively sampled modern populations (Pacioni *et al.* 2011), and evaluate whether application of aDNA techniques enhances our understanding of the possible genetic impact of declines in wildlife. Specifically, we set out to characterize ancient mitochondrial and microsatellite DNA to address two key questions for conservation management, namely quantifying the potential loss of genetic diversity and to assess the degree of former connectivity between extant populations remaining within the species' highly fragmented distribution. While modern data alone can be used to model past demographic events, it is not as powerful as using direct measurements of genetic diversity based on ancient DNA. The woylie exemplifies a common situation in that it is a critically endangered species where genetics has been largely ignored in its management for the past 20 years, despite specific recommendations in the woylie recovery plan (Start *et al.* 1995). This exclusion has prevented the establishment of management practices that seek to maximize the level of genetic diversity from amongst remaining populations. Lack of the genetic component also makes it impossible to evaluate whether assisted migration of individuals from different populations would be a viable management strategy.

Genetic diversity

The inclusion of aDNA data allowed us to conclude that the woylie underwent a dramatic genetic loss. This finding is somewhat in contrast with a previous assessment based only on modern data. In fact, modern levels of genetic diversity were considered relatively high (Pacioni *et al.* 2011). However, after the addition of data from historical samples, the rarefaction analysis and the estimation of the most likely number of haplotypes ($n = 43$, 95% CI 28–173, Fig. 3b) confirmed a mtDNA genetic loss of ~65% (range: 46–91%, considering that only 15 haplotypes are still detectable in modern populations compared to a range of 28–173 expected historical haplotypes). These findings are also supported by

Table 1 Mode (2.5 and 97.5 percentile) of number of migrants between woylie populations (from row to column)

	Kingston	Perup	Tutanning	Dryandra
Kingston		1.1 (0–6.1)	n/a	4.8 (0–23.5)
Perup	0.3 (0.1–0.5)		n/a	5.1 (0–23.5)
Tutanning	n/a	n/a		0.1 (0–12.9)
Dryandra	0.1 (0.002–0.5)	2 (0.2–6.8)	0.5 (0–10.3)	

comparisons of the historical and modern microsatellite data where we found a reduction of at least 20%.

The use of microsatellite loci in ancient DNA presents a challenge because of the potential problem of data fidelity associated with DNA degradation (Morin *et al.* 2001; Bourke *et al.* 2010). Amplification of microsatellite data from historical skins has been limited (Bourke *et al.* 2010), and even less frequent from fossil bones (Nyström *et al.* 2012; Allentoft *et al.* 2014). In our study, the inclusion of microsatellite data from historical samples proved to be very informative, demonstrating a significant reduction in allelic richness and heterozygosity across the loci examined. Furthermore, seven of the nine individuals with unique alleles were from within or close to the woylie's current geographical distribution. Considering the large number of modern samples that were profiled ($n = 231$, Pacioni *et al.* 2011), it is very likely that these alleles have been 'lost' from the modern populations. The three-parameter asymptotic exponential model predicted at least 20.7 historical alleles, suggesting a loss of at least 21.3%. Our results indicate how carefully one should interpret results based on modern data alone. In fact, based on the latter data set, only limited concerns were raised on this aspect of woylie conservation (Pacioni *et al.* 2011), whereas this study clearly demonstrates a dramatic overall reduction in genetic diversity. Another example of the potential bias generated by the lack of baseline data, which can result in misleading management targets, is that the mean genetic diversity (from modern data) of the wild woylie populations considered in this study was used as a reference term to evaluate whether the genetic diversity of translocated populations was satisfactory (Pacioni *et al.* 2013). The characterization of the historical genetic diversity, as conducted in this study, indicates that, while in principle that approach is valid, the conclusions are possibly rather optimistic.

It was only possible to quantify these estimations by including ancient DNA data and, considering the occurrence of allelic drop out that it is inevitable when working with aDNA, these should be considered conservative.

Demographic analyses

Data from historical samples also allowed a better understanding of the demographic history of the woylie compared with that obtained from modern samples only. Specifically, the historical data helped to quantify the extent of the decline and its consequences on the genetic diversity. Even with modern data alone, MIGRATE-N correctly detected the overall woylie decline. However, in these analyses, the estimated historical population size was substantially reduced as was the extent of the decline (Fig. 6). Differences in the results are even more dramatic in BEAST analyses. In fact, when all historical samples for which we had time-stamped data ($n = 22$) were included in the mtDNA analysis, the 95% HPD intervals were reduced and this analysis indicated that the woylie experienced more than a 99% decline in population size in recent times. When this analysis was repeated including only modern data ($n = 146$, Fig. 5), it was not possible to discriminate between a constant population and a declining trend in the demography of the woylie. MIGRATE-N analysis with microsatellite data detected the decline even with modern data alone. However, in this analysis too, the confidence intervals are slightly wider and the population size in recent times is mildly underestimated. These results exemplify the potential problems in reconstructing past demographic trajectories from modern data alone. It is evident that the inclusion of aDNA data can be highly informative and it is probably to have important implications for the management of threatened species.

The lack of significant results using common bottleneck tests highlights the limitations of these methods. The *M*-ratio tests detected genetic bottleneck in three out of four remnant wild populations, whereas the heterozygosity excess and mode-shift approaches did not detect a bottleneck at all. Various factors can influence the statistical power of the bottleneck tests, including the time since the bottleneck, the nature of the postdecline recovery and the effective population size prebottleneck (Cornuet & Luikart 1996; Garza & Williamson 2001; Williamson-Natesan 2005; Busch *et al.* 2007). We argue that the timing of the modern population sampling was too close to or contemporary with the most recent declines (i.e. after 1999), for them to be detected (despite the substantial loss of genetic diversity) by the heterozygosity excess and mode-shift methods (Cornuet & Luikart 1996; Garza & Williamson 2001; Williamson-Natesan 2005). Possibly, the demographic postreduction growth, consequent upon the fox control programmes started in the late 1990s, created a heterozygosity deficiency, which balanced out the heterozygosity excess following the bottleneck subsequent to the European settlement. On the other hand, we consider the results of the M-ratio analyses indicating a genetic bottleneck to be the outcome of substantial declines in the 20th century as opposed to the most recent declines. The *M*-ratio approach can detect a bottleneck up to 100 generations post decline (Garza & Williamson 2001), which would correspond to *c.* 250 years in the woylie (Groom 2010; Pacioni 2010). This is also supported by the observation that the new alleles found in historical samples were within the allele frequency distributions of the modern

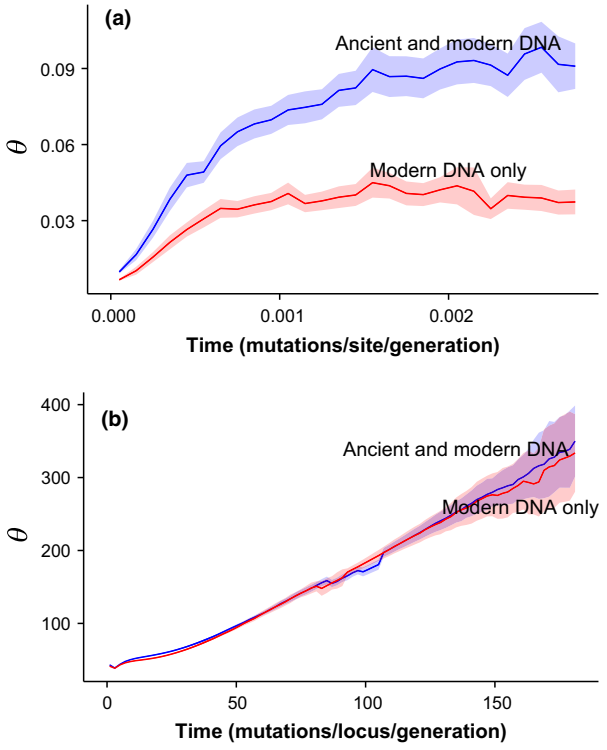


Fig. 6 Reconstruction of the woylie (*Bettongia penicillata ogilbyi*) demographic history based on control region of mtDNA (a) and microsatellite data (b) using a Bayesian Skyline Plot analysis implemented in MIGRATE-N v3.6.8 (Beerli 2006). In red analysis with modern data alone and in blue analysis including modern and ancient DNA data. Shaded areas represent 1.96 standard deviations from parameter values.

populations (i.e. their loss is responsible for a decrease of the *M*-ratio).

Gene flow

With the exception of the pair Kingston–Perup, the other modern populations are, nowadays, completely isolated (Pacioni *et al.* 2011). The coalescent-based analyses, demonstrated that historically these sites were connected, a finding that was only the subject of speculation in Pacioni *et al.* (2011). Analysis of microsatellite data suggests that these were actually part of a large panmictic population. These results are also supported by the private allele analysis, despite the limitations of this approach mentioned above. In this analysis, the estimates of the number of migrants between localities increased when data from the limited number of historical samples were included. This demonstrates a wider spatial distribution of alleles, just a hundred years ago (when most historical samples were collected), than that found at present. It is intriguing to note that the last hundred years is a time consistent with the substantial habitat loss and fragmentation in the WA Wheat-belt region (Saunders 1979).

The differences between the migration models supported by mtDNA and microsatellite data are not surprising. In fact, woylies have a male-biased dispersal (Christensen 1980; Pacioni 2010), with females normally settling near to their mother's home range boundaries (Pacioni 2010). Males are, therefore, primarily responsible for long-distance gene flow, which is not detected in the mtDNA data as mtDNA is maternally inherited.

Timing of the decline

We used BEAST to attempt to establish the time of commencement of the decline, because BEAST can take into account uncertainty around the evolutionary rates and calibration points while reconstructing the phylogeny and the demographic history. While our sampling

Table 2 Descriptive statistics of genetic diversity of microsatellite loci in the woylie (*Bettongia penicillata ogilbyi*) for modern (n = 231)/historical (n = 21) data sets. Between brackets SE

Locus	H_o (%)	N_A	N_{AR}	U_A	$P_{(fh)}$
Bt76	85.3/85.7	19/14	10.5/11.57	2	2.2
Bt64	73.8/100	24/19	11.08/14.29	2	9.4
T17-2	68.7/90.5	16/14	10.07/11.13	1	6.5
Y175	82.2/78.9	19/13	10.57/10.5	0	14.3
Pa593	83.5/88.9	16/13	10.36/11.27	1	8
Y105	40.9/23.1	11/5	5.4/4.4	0	0
Y112*	73.7/90	29/11	12.46/11	1	36.4
Pl2	70.6/68.8	10/8	6.69/6.82	1	5.8
Y170	81/90	18/14	10.51/11.56	0	16.3
Bt80	76.9/58.3	12/10	7.92/9.54	1	8.5
Pl26	42.1/54.5	5/4	3.46/4	0	0
Mean	70.8 (0.5)/ 75.3 (0.7)	16.3 (2.0)/ 11.4 (1.3)	9.0 (0.8)/ 9.6 (1)	0.8 (0.2)	9.8 (3)

H_o, observed heterozygosity; N_A, number of alleles; N_{AR}, allelic richness; U_A, historical unique alleles; $P_{(fh)}$, frequencies of falsely scoring an individual as homozygous (%).
*Excluded from statistical tests because of evidence of substantial allelic dropout.

Table 3 Number of migrants per generation between localities where modern woylie populations are located, calculated with the private allele method using 11 microsatellite loci (bold: ancient and modern DNA data, normal font: modern DNA data alone)

	Upper Warren	Tutanning	Dryandra
Upper Warren		0.52	0.91
Tutanning	**0.69**		0.68
Dryandra	**1.25**	**0.73**	

regime (which was mid way between the 'pooled' and 'scattered' sampling strategies) should not be prone to substantial biases (Heller *et al.* 2013), we acknowledge a possible bias in the inference of the demographic history in the BSP, because there is a strong indication of reduced gene flow between different regions within Western Australia in recent times. Habitat fragmentation and isolation of the studied populations generally lead to an underestimation of population size (Heller *et al.* 2013), and consequently the extent of the demographic decline may be overestimated. The effect of fragmentation on calculations of evolutionary rates is less predictable. Contrary to expectations, when using the same input data in BEAST as for the MIGRATE-N analysis (limiting data to samples geographically representative of the four modern populations), BEAST estimation of historical population size appears to be more conservative (although not substantially different) and the timing of the decline is only slightly shifted towards the present (Fig. S3, Supporting information). It should be stressed that such estimates are subject to uncertainties because the inter-specific mutation rates are probably lower than the intra-specific rates (Ho *et al.* 2008) and, accordingly, caution the reader to interpret these results with care. We argue that, while it is difficult to pinpoint exactly when the decline commenced its downward trajectory, it is clear that it started recently (i.e. less than ~4000 years ago) and has continued to the present. A possible catalyst for the start of this decline could have been the arrival of the dingo (*Canis lupus dingo*) in Australia (3500–5000 years ago, Savolainen *et al.* 2004; Fillios *et al.* 2012), but based on the observation of the substantial loss of genetic diversity compared to samples from as recently as 100 years ago, we also argue that the decline has been intensified since European settlement. Additional time-stamped samples that increase the time span covered by the aDNA data may help to more accurately estimate the timing of woylie decline, as it remains uncertain whether European factors (foxes, cats, disease), pre-European factors (e.g. dingo or fire regimes) or a combination of both, were the trigger for the decline in woylie genetic diversity and population numbers.

Implications for woylie conservation

The findings of this study have direct implications for the conservation and management of the critically endangered woylie in south-western Australia. Notably, evidence of a much wider former connectivity between the now fragmented habitats where the modern populations reside is highly relevant. Managing small isolated populations is challenging due to their susceptibility to stochastic events and greater effects of genetic drift (Frankham 1996; Frankham *et al.* 1999). Our findings show that there is a precedent for re-establishing gene flow between the currently isolated populations. Strategies such as natural corridors and translocations could help restore gene flow to historical levels and assist in the maintenance of the remaining genetic diversity.

Through direct measurement of ancient genetic variability, we uncovered the extent of genetic diversity loss in the woylie that occurred along with a substantial demographic reduction. A very conservative estimate of genetic loss would be of 65% for mtDNA and 20% for nuclear DNA, but in reality the number is likely to be far greater given the rate at which new genetic variants were detected in the ancient gene pool. This loss of genetic diversity can be attributed directly to the massive population decline detected with the BSP analysis. The severity and tempo of the decline should act as a catalyst for more active management strategies that seek to maintain the existing species-wide genetic diversity. The translocated woylie populations that were established in the late 1980s (e.g. populations in South Australia, Delroy *et al.* 1986), despite being at lower genetic diversity than indigenous populations (Pacioni *et al.* 2013), may have captured some of the genetic diversity that it is now lost in Western Australia and may play a role in the future genetic management of the woylie. Indeed, a haplotype found in South Australian translocated populations was not found in Western Australian populations, but it was identified from an historical sample from Dryandra.

A further conservation outcome is the characterization of woylie genetics at the boundaries of the species' distribution. For example our profiling of fossil material from Faure Island, at the northern limit of the former distribution, demonstrated that the woylies that once inhabited this region were genetically very similar to extant populations despite being more than 1000 km away. Faure Island and Dirk Hartog Island are important sites in the Shark Bay World Heritage Area for ecosystem restoration (DEC 2005), which involves translocations of previously present species to re-establish past biodiversity in the region. Our data suggest that the use of the available genetic stock is appropriate and it is not unreasonable to expect that woylies might adapt easily to these relatively more arid environments.

Conclusion

The mode and tempo of genetic diversity loss is often difficult to quantify in species that have already undergone significant declines. The south-west of Western Australia is a biodiversity hotspot and therefore a

critical region for conservation purposes, yet baseline data for past biodiversity are limited. Moreover, many marsupials are still in decline in Australia, but it is difficult to quantify changes in the genetic profile of these species over time. We set out to investigate how knowledge of the past might assist best-practice conservation in the present. The south-west of Western Australia used to provide a rich, continuous habitat to numerous, now endangered species and while we focused on only one species, it is probably that the degree of genetic diversity loss is directly relevant to many endemic species in the region. Whilst detailed screening of available fossils and skins was conducted in the case of the woylie, the number of samples available in these kinds of studies will always be limiting. The lack of samples for key localities, sampling restrictions on skin and bone voucher specimens and constraints imposed by DNA preservation are all limitations when trying to quantify past diversity using ancient DNA techniques. Despite these limitations, this study demonstrates that the inclusion of historical material in genetic studies is beneficial for wildlife management and conservation. We demonstrated that the inclusion of aDNA data improves the quantification of genetic diversity changes and inference of historical population dynamics. In addition, we advocate that, when technically feasible, it is important to expand ancient DNA studies past the use of mtDNA in an attempt to quantify what has occurred at the nuclear DNA level; both microsatellite data and single nucleotide polymorphisms can provide these insights and such approaches are aided by next-generation DNA sequencing platforms. This becomes especially relevant in species where male-biased dispersal may skew mtDNA interpretation of diversity, connectivity and phylogeography.

Acknowledgements

We thank the Wardandi people of south-western Australia for their support of the original archaeological excavations that provided many of the samples analysed. Thanks to Ric How and the WA museum for historical samples and assistance. Computer simulations were supported by PAWSEY through the use of advanced computing resources provided by the Informatics Facility located at Murdoch University and CIPRES (http://www.phylo.org/sub_sections/portal/cite.php). We would like to thank Stefan Prost and Christian Anderson for promptly integrating our suggestion of a 'pie chart' feature in Tempnet. Simon Ho, Martyna Molak and Alexei Drummond provided useful comments regarding analysis of aDNA data and two anonymous reviewers whose suggestions greatly improved this manuscript. We are very grateful to Peter Beerli, who provided advice and assistance in MIGRATE-N analyses. This research was funded by an ARC Future Fellowship (MB) to explore past biodiversity in WA (FT0991741).

References

Allentoft ME, Oskam C, Houston J *et al.* (2011) Profiling the dead: generating microsatellite data from fossil bones of extinct megafauna—protocols, problems, and prospects. *PLoS ONE*, **6**, e16670.

Allentoft ME, Collins M, Harker D *et al.* (2012) The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society of London B: Biological Sciences*, **279**, 4724–4733.

Allentoft ME, Heller R, Oskam CL *et al.* (2014) Extinct New Zealand megafauna were not in decline before human colonization. *Proceedings of the National Academy of Sciences of the USA*, **111**, 4922–4927.

Andrén H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. *Oikos*, **71**, 355–366.

Baele G, Li WLS, Drummond AJ, Suchard MA, Lemey P (2012) Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, **30**, 239–243.

Barton NH, Slatkin M (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity*, **56**, 409–415.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Bourke BP, Frantz AC, Lavers CP *et al.* (2010) Genetic signatures of population change in the British golden eagle (*Aquila chrysaetos*). *Conservation Genetics*, **11**, 1837–1846.

Bouzat JL, Paige KN, Lewin HA (1998) The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *The American Naturalist*, **152**, 1–6.

Brace S, Palkopoulou E, Dalén L *et al.* (2012) Serial population extinctions in a small mammal indicate Late Pleistocene ecosystem instability. *Proceedings of the National Academy of Sciences of the USA*, **109**, 20532–20536.

Briggs AW, Stenzel U, Johnson PLF *et al.* (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the USA*, **104**, 14616–14621.

Burbidge AA, McKenzie NL (1989) Patterns in the modern decline of Western Australia's vertebrate fauna: causes and conservation implications. *Biological Conservation*, **50**, 143–198.

Burbidge AA, Johnson KA, Fuller PJ, Southgate RI (1988) Aboriginal knowledge of the mammals of the central deserts of Australia. *Wildlife Research*, **15**, 9–39.

Busch JD, Waser PM, DeWoody JA (2007) Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Molecular Ecology*, **16**, 2450–2462.

Campos PF, Willerslev E, Sher A *et al.* (2010) Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (*Ovibos moschatus*) population dynamics. *Proceedings of the National Academy of Sciences of the USA*, **107**, 5675–5680.

Christensen PES (1980) *The Biology of Bettongia penicillata Gray, 1837, and Macropus eugenii (Desmarest, 1817) in Relation to Fire*. Forests Department of Western Australia Bulletin, No. 91, i–ix, 1–90, Perth, Western Australia.

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

DEC (2005) *Shark Bay Terrestrial Reserves and Proposed Terrestrial Reserve Additions Management Plan*. Department of Environment and Conservation, Perth.

Delroy LB, Earl J, Radbone I, Robinson AC, Hewett M (1986) The breeding and re-establishment of the brush-tailed bettong, *Bettongia penicillata*, in South-Australia. *Wildlife Research*, **13**, 387–396.

Dixon CJ (2006) A means of estimating the completeness of haplotype sampling using the Stirling probability distribution. *Molecular Ecology Notes*, **6**, 650–652.

Dortch J (2004a) Late quaternary vegetation change and the extinction of black-flanked rock-wallaby (*Petrogale lateralis*) at Tunnel Cave, Southwestern Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **211**, 185–204.

Dortch J (2004b) *Paleo-environmental change and the persistence of human occupation in south-western Australian forests*. British Archaeological Reports, International Series 1288.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Ewers RM, Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews*, **81**, 117–142.

Fillios M, Crowther MS, Letnic M (2012) The impact of the dingo on the thylacine in Holocene Australia. *World Archaeology*, **44**, 118–134.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genetics Research*, **66**, 95–107.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, **10**, 1500–1508.

Frankham R, Lees K, Montgomery ME *et al.* (1999) Do population size bottlenecks reduce evolutionary potential? *Animal Conservation*, **2**, 255–260.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology*, **6**, 861–868.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. *Trends in Ecology & Evolution*, **20**, 541–544.

Groom C (2010) Justification for continued conservation efforts following the delisting of a threatened species: a case study of the woylie, *Bettongia penicillata ogilbyi* (Marsupialia: Potoroidae). *Wildlife Research*, **37**, 183–193.

Heller R, Chikhi L, Siegismund HR (2013) The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS ONE*, **8**, e62992.

Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.

Ho SY, Saarma U, Barnett R, Haile J, Shapiro B (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS ONE*, **3**, e1615.

Hunt H (2010) *A temporal assessment investigating the effects of population declines on genetic diversity, in the critically endangered woylie (*Bettongia penicillata ogilbyi*)*. Honours Thesis, Murdoch University, Perth.

Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*, **5**, 539–543.

Kalinowski ST (2006) hw-quickcheck: an easy-to-use computer program for checking genotypes for agreement with Hardy–Weinberg expectations. *Molecular Ecology Notes*, **6**, 974–979.

Leonard JA (2008) Ancient DNA applications for wildlife conservation. *Molecular Ecology*, **17**, 4186–4196.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, **12**, 228–237.

Morin P, Chambers K, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology*, **10**, 1835–1844.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Navascués M, Depaulis F, Emerson BC (2010) Combining contemporary and ancient DNA in population genetic and phylogeographical studies. *Molecular Ecology Resources*, **10**, 760–772.

Neal RM (2003) Slice sampling. *Annals of statistics*, **31**, 705–741.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, New York.

Nyström V, Humphrey J, Skoglund P *et al.* (2012) Microsatellite genotyping reveals end Pleistocene decline in mammoth autosomal genetic variation. *Molecular Ecology*, **21**, 3391–3402.

Oksanen J, Blanchet GF, Kindt R *et al.* (2013) *Vegan: community ecology package*. R package version 2.0-8.

Pacioni C (2010) *The population and epidemiological dynamics associated with recent decline of woylies (*Bettongia penicillata*) in Australia*. PhD, Murdoch University, Perth.

Pacioni C, Spencer P (2010) Capturing genetic information using non-target species markers in a species that has undergone a population crash. *Australian Mammalogy*, **32**, 33–38.

Pacioni C, Wayne AF, Spencer PBS (2011) Effects of habitat fragmentation on population structure and long distance gene flow in an endangered marsupial: the woylie. *Journal of Zoology*, **283**, 98–107.

Pacioni C, Wayne AF, Spencer P (2013) Genetic outcomes from the translocations of the critically endangered woylie. *Current Zoology*, **59**, 294–310.

Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peery MZ, Kirby R, Reid BN *et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, **21**, 3403–3418.

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

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Prost S, Anderson CNK (2011) TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. *Methods in Ecology and Evolution*, **2**, 663–667.

R Development Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Ramakrishnan UMA, Hadly EA (2009) Using phylochronology to reveal cryptic population histories: review and synthesis of 29 ancient DNA studies. *Molecular Ecology*, **18**, 1310–1330.

Rambaut A, Drummond AJ (2007) *TRACER*. Available from http://beast.bio.ed.ac.uk/Tracer.

Rioux Paquette S (2011) *PopGenKit: useful functions for (batch) file conversion and data resampling in microsatellite datasets*. R package, version 1.0.

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

Saunders D (1979) The availability of tree hollows for use as nest sites by White-tailed Black Cockatoos. *Wildlife Research*, **6**, 205–216.

Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E, Lundeberg J (2004) A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proceedings of the National Academy of Sciences of the USA*, **101**, 12387–12390.

Shapiro B, Drummond AJ, Rambaut A *et al.* (2004) Rise and fall of the Beringian steppe bison. *Science*, **306**, 1561–1565.

Start T, Burbidge AA, Armstrong D, Woylie Recovery Team (1995) *Woylie Recovery Plan*, 2nd edn. Department of Conservation and Land Management, Como.

Start AN, Burbidge AA, Armstrong D (1998) A review of the conservation status of the woylie, *Bettongia penicillata ogilbyi* (Marsupialia: Potoroidae) using IUCN criteria. *CalmScience*, **2**, 277–289.

Taylor AC, Sherwin WB, Wayne RK (1994) Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii*. *Molecular Ecology*, **3**, 277–290.

Van Dyck S, Strahan R (2008) *The Mammals of Australia*, 3rd edn. New Holland Publishers, Sydney, New South Wales.

Wayne AF, Friend T, Burbidge AA, Morris K, Van Weenen J (2008a) *Bettongia penicillata*. Available from www.iucnredlist.org.

Wayne AF, Ward C, Maxwell M *et al.* (2008b) Diagnosing the recent collapse of the woylie in southwestern Australia. 21st Australian Wildlife Management Society Conference, p. 140.

Wayne AF, Maxwell M, Ward C *et al.* (2013) The importance of getting the numbers right: quantifying the rapid and substantial decline of an abundant marsupial, *Bettongia penicillata*. *Wildlife Research*, **40**, 169–183.

Wayne AF, Maxwell M, Ward C *et al.* (2015) Suddenand rapid decline of the abundant marsupial *Bettongia penicillata* in Australia. *Oryx*, **49**, 175–185.

Westerman M, Loke S, Springer MS (2004) Molecular phylogenetic relationships of two extinct potoroid marsupials, *Potorous platyops* and *Caloprymnus campestris* (Potoroinae: Marsupialia). *Molecular Phylogenetics and Evolution*, **31**, 476–485.

Willerslev E, Cooper A (2005) Review paper. Ancient DNA. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**, 3–16.

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conservation Genetics*, **6**, 551–562.

Woinarski JCZ, Burbidge AA, Harrison PL (2015) Ongoing unraveling of a continental fauna: decline and extinction of Australian mammals since European settlement. *Proceedings of the National Academy of Sciences of the USA*, **112**, 4531–4540.

C.P. and M.B. conceived the experiments. A.W., A.B. and J.D. assisted with sample selection, dating and conservation implications. C.P., H.H., M.A., D.H. and T.V. generated and analysed the data and prepared figures/tables. C.P., H.H., M.A. and M.B. wrote the manuscript with input from all co-authors.

Data accessibility

DNA sequences: GenBank accession numbers KT876917–KT876973.

Microsatellite data, and BEAST and MIGRATE-N input and parameter files: Dryad Digital Repository: doi: 10.5061/dryad.26s85.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 The success rate of positive genotyping associated with the corresponding sample C_T value, obtained from the qPCR 302/425 assay.

Fig. S2 Bayesian phylogeny of *Bettoniga penicillata ogilbyi*, showing the maximum a posteriori (MAP) tree generated using BEAST.

Fig. S3 Comparison between reconstruction of the woylie (*Bettongia penicillata*) demographic history based on control region of mtDNA using a Bayesian Skyline Plot analysis implemented in BEAST v1.8 (Drummond *et al.* 2012)—black—and MIGRATE-N (Beerli 2006)—purple.

Table S1 Summary of information of historical woylie (*Bettongia penicillata ogilbyi*) samples, including DNA preservation, age, and their use for statistical analyses.

Table S2 Mitochondrial cytochrome *b* and control region primer information used to genetically characterise *Bettongia*.

Table S3 Details of the microsatellite loci amplified in the woylie (*Bettongia penicillata ogilbyi*) including the species where these were first developed.

Table S4 List of cyt b sequences used for divergence time estimation.

APPENDIX E – SCRAPHEAP CHALLENGE: A NOVEL BULK-BONE METABARCODING METHOD TO INVESTIGATE ANCIENT DNA IN FAUNAL ASSEMBLAGES.

The following citation resulted from research that is of relevance to this doctoral thesis and was published during my candidature. In context of this multi-author publication I developed the 12s rRNA region primer sets targeting reptilian DNA in the bulk bone samples which contributed to the HTS data generated.

A copy of the main article is attached.

Murray, D.C., Haile, J., Dortch, J., White, N., **Haouchar, D**., Bellgard, M.I., Allcock, R.J., Prideaux, G.J., Bunce, M., 2013. Scrapheap challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Scientific Reports* **3**, 3371.

SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
PALAEONTOLOGY
ARCHAEOLOGY
BIODIVERSITY

Received
3 May 2013

Accepted
5 November 2013

Published
28 November 2013

Correspondence and requests for materials should be addressed to M.B. (michael.bunce@curtin.edu.au)

* Current address: Trace and Environmental DNA laboratory, Department of Environment and Agriculture, Curtin University, Perth, Western Australia, 6845, Australia.

Scrapheap Challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages

Dáithí C. Murray[1]*, James Haile[1]*, Joe Dortch[2], Nicole E. White[1]*, Dalal Haouchar[1], Matthew I. Bellgard[3], Richard J. Allcock[4], Gavin J. Prideaux[5] & Michael Bunce[1]*

[1]Ancient DNA Laboratory, School of Veterinary and Life Sciences, Murdoch University, South Street, Murdoch, WA, 6150, Australia, [2]Eureka Archaeological Research and Consulting, School of Social Sciences, The University of Western Australia, Crawley, Perth, WA, 6009, Australia, [3]Centre for Comparative Genomics, Murdoch University, South Street, Murdoch, WA, 6150, Australia, [4]LotteryWest State Biomedical Facility: Genomics, School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, WA, 6009, Australia, [5]School of Biological Sciences, Flinders University, Bedford Park, SA, 5042, Australia.

Highly fragmented and morphologically indistinct fossil bone is common in archaeological and paleontological deposits but unfortunately it is of little use in compiling faunal assemblages. The development of a cost-effective methodology to taxonomically identify bulk bone is therefore a key challenge. Here, an ancient DNA methodology using high-throughput sequencing is developed to survey and analyse thousands of archaeological bones from southwest Australia. Fossils were collectively ground together depending on which of fifteen stratigraphical layers they were excavated from. By generating fifteen synthetic blends of bulk bone powder, each corresponding to a chronologically distinct layer, samples could be collectively analysed in an efficient manner. A diverse range of taxa, including endemic, extirpated and hitherto unrecorded taxa, dating back to c.46,000 years BP was characterized. The method is a novel, cost-effective use for unidentifiable bone fragments and a powerful molecular tool for surveying fossils that otherwise end up on the taxonomic "scrapheap".

Fossil assemblages offer insights into past biodiversity, palaeoecology and human activities[1–3]. However, the accuracy of fossil identifications relies on the preservation of taxonomically significant morphological features, which are often lacking in highly fragmented remains. Over the past decade, analyses of ancient DNA (aDNA) have developed in sophistication and the breadth of contexts in which they are applied. Ancient DNA has been used to address questions of speciation, extinction and disease[4–7] using a variety of substrates, including bone[8], hair[9] and eggshell[10]. However, to date, no study has attempted to use aDNA from taxonomically diverse fossils to map faunal assemblage data from a single site, largely due to the time and cost associated with generating aDNA sequences from each bone fragment.

The destructive nature of sampling also means researchers and collection managers may be reluctant to analyse valuable specimens. At the same time, most archaeological and palaeontological excavations also collect large numbers of small, morphologically indistinct bone fragments (Figure 1a). Such material is of limited use in species identifications, although it may be important for some taphonomic analyses. Taxonomically, however, it is usually destined for the analytical "scrapheap".

It is now possible, largely due to second generation high-throughput DNA sequencing (HTS) methodologies, to genetically profile complex, heterogeneous samples (Figure 1b) in parallel, both cheaply and quickly[11,12]. This DNA metabarcoding[13] approach to genetically unravel complex substrates via HTS, as opposed to cloning, has transformed the analysis of substrates such as sediment[14,15] and faecal material[16,17]. To explore large HTS-generated genomic datasets from environmental samples researchers use tools that are either: 1) taxonomy-dependent, which involves searching DNA reference databases for query and reference sequence matches[18,19], or 2) taxonomy-independent, which involves taxonomy-independent measures of sequence diversity and clustering such as Operational Taxonomic Unit (OTU) analysis or UniFrac-based methods[20–22].

This study seeks to employ HTS technology to sequence and identify aDNA obtained from thousands of morphologically unidentifiable archaeological bone fragments freshly excavated from deposits at Tunnel Cave (115° 02′ E, 34° 05′ S) and Devil's Lair (115° 04′ E, 30° 09′ S), two archaeologically and culturally significant sites

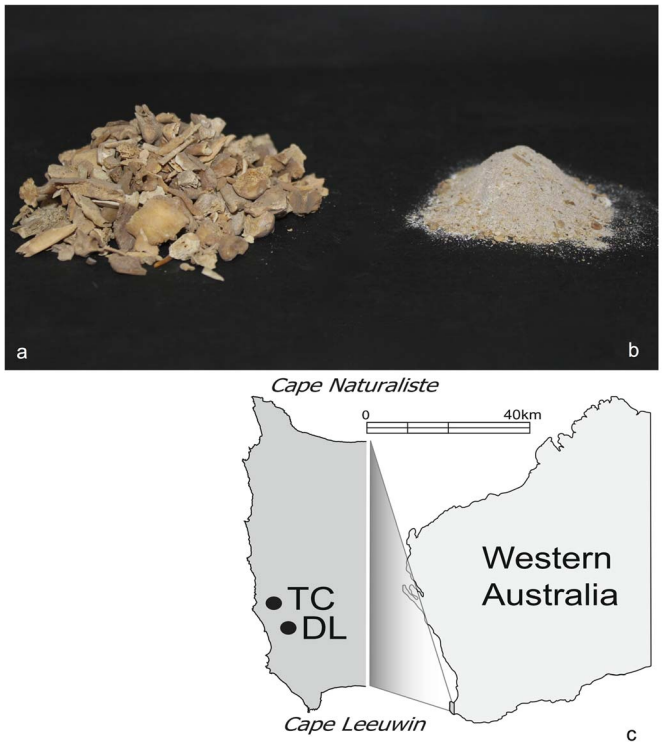


Figure 1 | Bulk-bone fragments ground to form a bulk-bone powder at two archaeological sites. Morphologically indistinct bulk-bone fragments (a) were ground to form single bulk-bone powder samples (b). Bulk-bone fragments were excavated from Devil's Lair (DL) and Tunnel Cave (TC), two archaeologically significant sites in southwest Western Australia (c). The map used in (c) was sourced from www.openclipart.org and was modified by J.H in Adobe Illustrator.

in southwestern Australia (Figure 1c). Taken together, these sites, used to explore this methodological approach, span the last c.50,000 years[23] and provide an unparalleled opportunity to study past Australian biodiversity and Aboriginal occupation[23] located within an internationally recognised biodiversity "hotspot"[24]. A new method for the bulk sampling of fragmented bone material that would otherwise remain an untapped taxonomic resource is presented. By grinding multiple bones (Figure 1a) into an artificial "bulk-bone powder" (Figure 1b), thus producing a single bulk-bone powder sample, a large amount of highly informative genetic data can be quickly extracted. Such an approach should become commonplace in archaeological and palaeontological practice as it enables rapid assessment of DNA preservation and effectively maps zooarchaeological and palaeontological assemblages without destructive sampling of more valuable fossils.

Results

Overview of data generated. In a 2012 excavation, thousands of small bone fragments were collected by dry-sieving sediment from 15 well-dated stratigraphic units or layers at Devil's Lair and Tunnel Cave (Figure 1c). Around 50–150 bone fragments from within each layer were each drilled for 10–15s to form 15 bulk-bone powder samples representing the 15 layers (Figures 1a, 1b). DNA was extracted from each bulk-bone powder sample using established extraction methods (described in Methods) as if the bulk-bone sample were a single-source sample. The DNA extracts were screened for amplifiable mitochondrial DNA (mtDNA) using generic primers (tagged with HTS adaptors and unique barcodes) and subsequently sequenced using two HTS platforms: the GS-Junior (Roche) and the Ion Torrent PGM (Life Technologies).

Ancient DNA was successfully extracted from all bulk-bone powder samples, including a layer dated c.44,260–46,890 years BP (uncalibrated). The successful amplification and sequencing of DNA from all 15 layers was a rapid, cheap and effective way to assess DNA preservation at the sites (Figure 1c).

Amplicon DNA sequences (hereafter referred to as sequences) obtained from collective GS-Junior and Ion Torrent PGM sequencing runs were analysed for quality and possible chimeras. Except for ubiquitous human DNA sequences, control reactions throughout the process (described in Methods) were negative for contaminating DNA arising from laboratory processing.

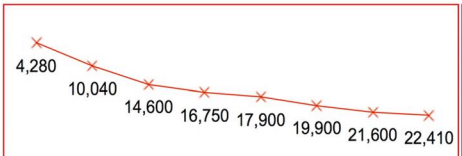
Short regions within the mammalian mitochondrial 12S and 16S rRNA genes were amplified generating products of 100–104 bp and 90–96 bp respectively[25]. Amplification and sequencing of avian mtDNA was successful for some samples, producing either a 106–121 bp or 227–239 bp region of the avian mtDNA 12S gene[25]. Some cross-species reactivity was observed when using both 12S and 16S mammalian primer sets, resulting in the amplification and sequencing of avian and reptilian DNA. A targeted quantitative PCR (qPCR) and HTS approach to identify snake species was successful for a single sample.

Taxonomic identification. Mammalian 12S and 16S assays identified eight mammalian families representing 16 genera, using assignment filters chosen for this study (see Methods; Figure 2). The increase in sequencing depth afforded by the Ion Torrent PGM, as compared to the GS-Junior, did not increase the diversity of taxa identified. Mammalian taxa endemic to Australia were detected in multiple samples, in addition to taxa that have undergone significant range contraction and extirpation. The macropodid genus *Thylogale* (pademelon), provided the closest BLAST matches for many sequences across multiple samples, but to date no member of the genus has been recorded in this region. It was not possible to provide accurate taxonomic identifications for most of the Muridae sequences and for many *Macropus* sequences. While many sequences could be assigned with high confidence to a genus level, others could not be assigned beyond family or genus. A number of birds and reptiles were also identified and these have been collated at the family and genus level (Figure 2). While assignment to the species level is certainly possible in many instances a conservative approach is adopted here to showcase the approach.

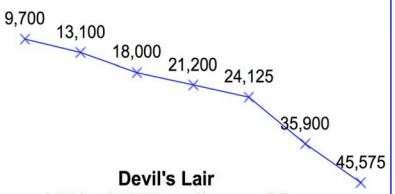
Genetic biodiversity analysis. A largely taxonomy-independent approach was adopted to examine fluctuations in observed genetic diversity over time at both sites. While the taxa identified using the GS-Junior and Ion Torrent PGM were mostly congruent, coverage dependent OTU inflation, arising from homopolymer sequencing error (see Methods; Discussion) was observed. A modified OTU analysis filter was designed to reduce the influence of HTS homopolymer sequencing error[26,27], by employing distance-based metrics obtained from sequence alignments, giving rise to a new method referred to here as Distance-based Taxonomic Units (DTUs).

A total of 72 DTUs were identified across all 15 samples, 23 of which were shared across multiple samples, and in some instances both archeological sites (Figure 3). The number of DTUs fluctuates noticeably with time (Figure 4). The number of DTUs shows a notable decrease that roughly coincides with the last glacial maximum (LGM), whilst also showing an increase post-LGM. The composition of DTUs also varies over time. For instance, Potoroidae (potoroids) DTUs appear around the LGM and show an increase in numbers, whilst numbers of Macropodidae (macropodids) DTUs show a decline post-LGM.

With obvious variation in DTU composition, macropodid sequences were selected to examine DTU number flux at a finer scale to examine whether or not this reflected the overall trends in biodiversity change. Macropodids exhibit a declining trend in DTU diversity post-LGM (Figure 5) that marginally increases near the Holocene/Pleistocene transition 11,700 years ago.



Tunnel Cave
4,280 - 22,410 median yrs BP



Devil's Lair
9,700 - 45,575 median yrs BP

	Family	Genus	Tunnel Cave								Devil's Lair						
Mammalia	Dasyuridae		✓		✓						✓					✓	✓
		Antechinus	✓														
		Dasyurus		✓	✓	✓					✓						
		Sarcophilus	✓												✓		
	Macropodidae		✓†	✓†	✓†	✓	✓†	✓†	✓	✓	✓†	✓†	✓†	✓†	✓†	✓	
		*Dendrolagus**						✓									
		*Dorcopsis**									✓				✓		
		Macropus		✓†	✓†	✓	✓	✓	✓†	✓†	✓†	✓	✓†	✓†			
		Petrogale		✓		✓	✓	✓	✓	✓							
		Setonix	✓				✓		✓		✓	✓	✓	✓	✓	✓	
		*Thylogale**						✓				✓	✓	✓	✓	✓	
	Muridae			✓†								✓†	✓			✓†	✓†
		Rattus#	✓†	✓		✓†	✓†	✓†	✓†	✓					✓	✓†	✓†
	Peramelidae				✓		✓							✓			
		Isoodon			✓†	✓	✓†	✓†			✓	✓		✓†	✓		
	Phalangeridae																
		Trichosurus				✓		✓†	✓	✓	✓	✓†	✓†				
	Potoroidae			✓							✓		✓				
		Bettongia		✓†	✓	✓†	✓†	✓	✓		✓†	✓†	✓†				
		Potorous	✓†								✓†	✓†					✓
	Pseudocheiridae	*Pseudocheirus*	✓		✓†			✓†	✓	✓	✓†			✓	✓		
	Tarsipedidae	*Tarsipes*														✓	✓
Aves	Acanthizidae											✓†					
	Anatidae	*Anas*								✓							
		Cygnus				✓											
	Cardinalidae*											✓					
	Corvidae											✓†					
	Dasyornithidae	*Dasyornis*										✓†					
	Diomedeidae		✓														
	Meliphagidae				✓	✓†											
		Phylidonyris				✓											
	Phalacrocoracidae	*Phalacrocorax*						✓									
	Psittaculidae		✓									✓					
	Timaliidae				✓	✓					✓	✓					
Reptilia	Elapidae																
		Pseudonaja									✓						
	Scincidae																
		Tiliqua											✓				

Figure 2 | Taxa identified in bulk-bone powder samples. Mammals, birds and reptiles identified in each sample are listed. Samples are grouped according to site from youngest to oldest in years BP (uncalibrated), which is plotted on the same scale for both sites. The criteria used in taxonomic assignment are detailed in the Methods. Note that there is uncertainty surrounding taxonomy with regards to both Timaliidae and Cardinalidae (See Discussion). Key: †Detected using multiple primer sets; *Taxa not historically known to occur in the study region; #Sequences assigned to *Rattus* aligned closest to native *Rattus Fuscipes* (bush rat).

Discussion

This study presents a novel HTS method using aDNA characterised from bulk-bone powder samples. It represents a powerful new approach to analyse unidentifiable fragments excavated from fossil deposits. Ancient DNA extracted from bones within a layer dated between 44,260–46,890 years BP (uncalibrated), is the oldest aDNA recovered from Australia to date. These HTS results and the initial exploration of this technique show promise for larger scale bulk-bone analyses of fossil deposits. Rapidly analysing a bulk bone sample to determine if a site is conducive to DNA preservation will be

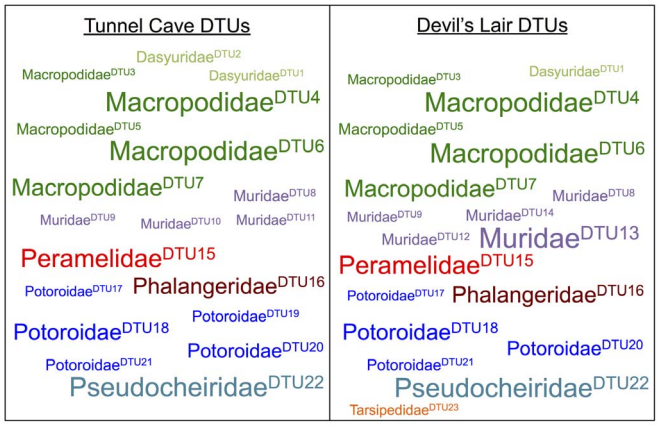


Figure 3 | DTUs shared across bulk-bone powder samples. The DTUs shared between bulk-bone powder samples, and across both Tunnel Cave (left) and Devil's Lair (right), are shown. DTUs have been labeled with the closest BLAST family matches. Each DTU has been assigned a numeric identifier following the acronym 'DTU', shown in superscript. Font size is indicative of the total number of samples a DTU was detected in.

valuable in excavations and test pits as DNA becomes increasingly incorporated into archaeological and palaeontological practices.

Even with the limited sampling, this first foray into bulk-bone analyses, has uncovered a significant amount of biological information that adds substantially to previous knowledge of the sites and surrounding biodiversity. Analysing these data in the context of DNA damage, sequencing error, incomplete reference databases and the necessary use of short DNA sequences raises numerous challenges that must be systematically addressed[17,28–30]. Nevertheless, when appropriate protocols and sequence filters are applied (see Methods) the method affords new insights into past biodiversity (Figure 2) and its temporal and spatial variation (Figures 3, 4 and 5).

Raw DNA sequences obtained from HTS platforms can be sorted and screened using a combination of filters that collectively exclude low-quality reads (Q-scores), sequences with errors in known flanking regions (adaptors, primers, and barcodes), artificial chimeric sequences and low abundance reads (see Methods). However, even sequences that pass these filters need to be interpreted with caution: the bird family Cardinalidae, which is not known to occur in Australia, is a case in point. The identification of birds also serves to illustrate the pitfalls associated with taxonomic revision. The taxonomy of the family Cardinalidae has been revised on a number of occasions, as has that of Timaliidae, which was also identified in some samples. Timaliidae has been regarded as a family consisting of Old World passerine birds, however the Australasian babblers (family: Pomatostomidae) were once within this family and the typical white-eyes (*Zosterops*) are disputably within this family also[31]. The families and genera identified (Figure 2) within each of the 15 samples require further investigation to identify taxa to the species level. Nevertheless, most of the genera identified at both sites from fossil morphology were again successfully detected in

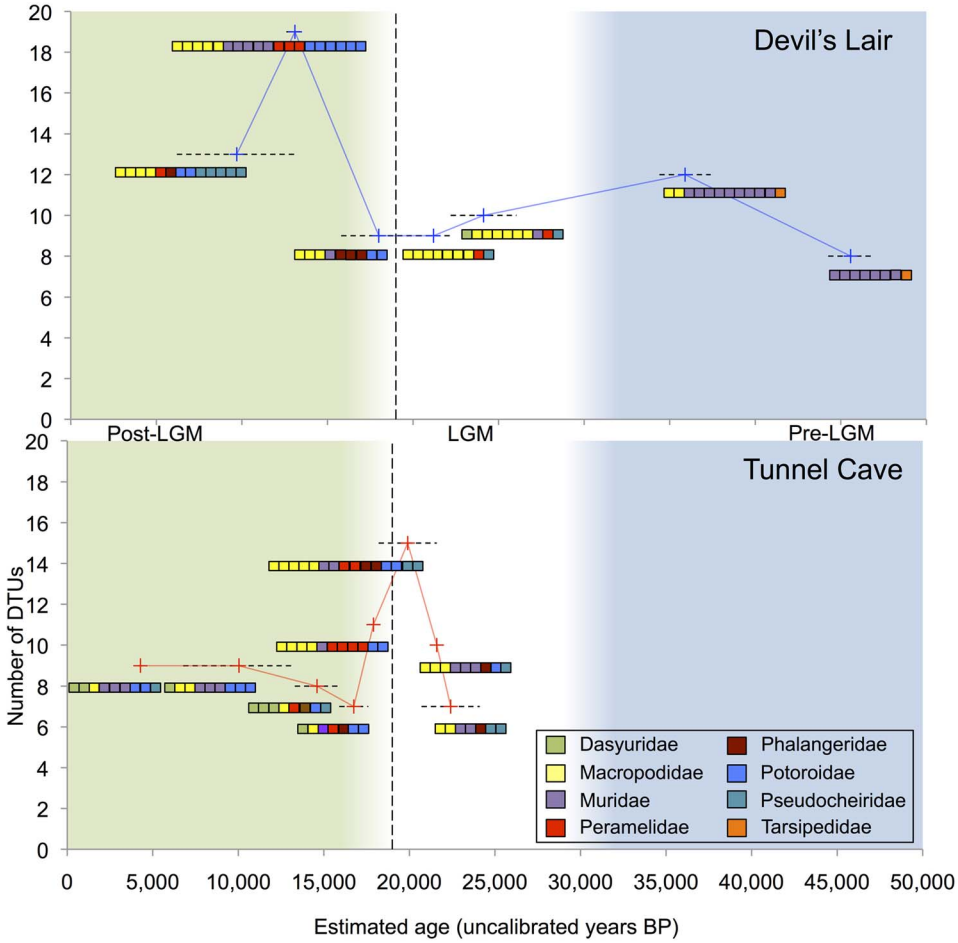


Figure 4 | Change in DTU number and composition over time at Tunnel Cave and Devil's Lair. The fluctuation in DTU number and the change in DTU composition across samples and at both sites are plotted against the backdrop of the major climatic shift around the end of the Last Glacial Maximum (LGM). Dashed vertical line - approximate end of the LGM; Blue background – Pre-LGM; White background – LGM; Green background – Post-LGM. Median ages are plotted for each sample; dashed horizontal line indicates minimum and maximum accepted date range for each layer.

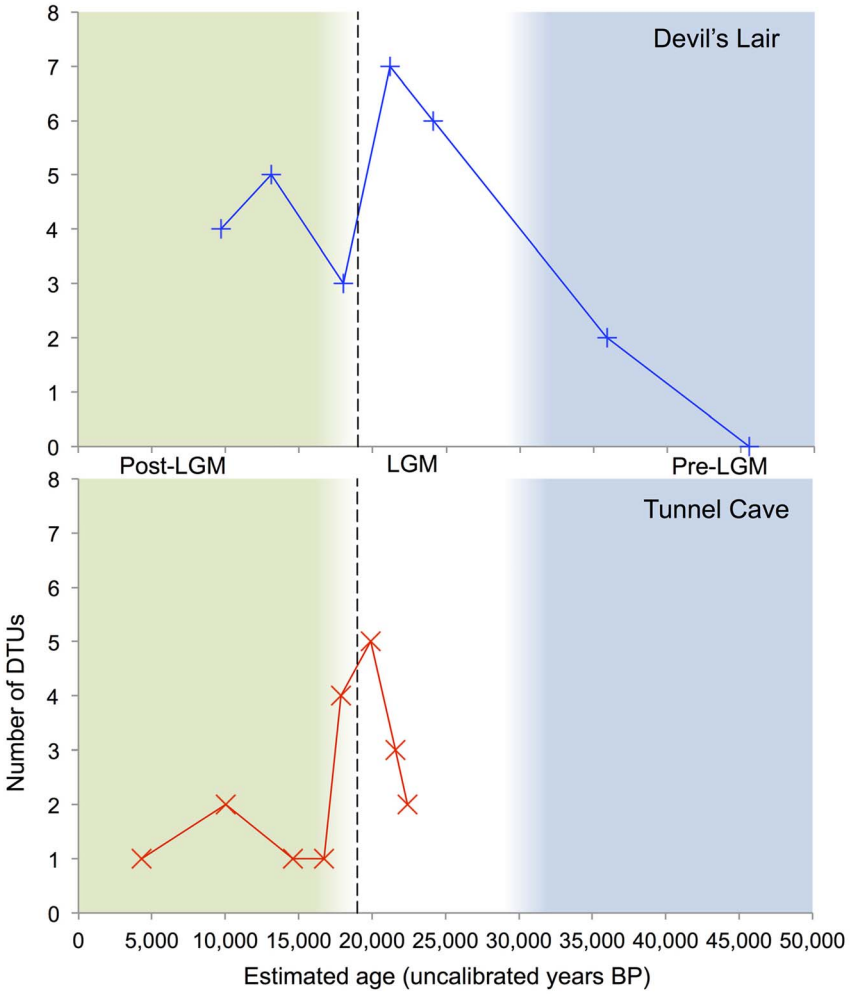


Figure 5 | Change in Macropodidae DTU number over time at Tunnel Cave and Devil's Lair. The fluctuation in Macropodidae DTU number across samples and at both sites is illustrated. Dashed vertical line - approximate end of the LGM; Blue background – Pre-LGM; White background – LGM; Green background – Post-LGM. Median ages are plotted for each sample.

the bulk-bone[23]. The absence of some morphologically identified taxa from the genetically-determined faunal assemblage list is most likely due to sampling bias, as the present analysis derives from deposits representing less than one percent of the volume of the original excavations. Additionally, the possibility of primer binding bias contributing to the discontinuities between both aDNA and fossil assemblage datasets cannot be excluded. *In silico* analysis of variation in binding sites and the use of the multiple markers attempts to identify and minimize the impact of amplification bias. Finally, inherent differences between bones in terms of the preservation and quantum of mtDNA per unit biomass may also skew results between both methods of analysis causing artifactual over-representation of some taxa relative to others. However, taxa were also identified that were not detected in any previous morphology-based analyses, particularly small mammals, birds and reptiles, all of which require highly-specialised taxonomic skills to identify, are less likely to preserve diagnostic remains, and may be poorly represented in reference collections.

A high level of confidence surrounds the bulk of the taxonomic identifications; for instance, the majority of mammalian taxa identified are locally extant or known from the fossil record. The same generally holds true for avian and reptilian taxa identifications. The detection of sequences endemic to southwest Australia, such as a 100% match to *Tarsipes rostratus* (honey possum), further supports the *bona fide* nature of the sequences obtained. Moreover, the detection of extirpated taxa, such as *Setonix* (quokka) and *Sarcophilus* (Tasmanian devil), as far back as c.24,000 years BP (uncalibrated) illustrates the antiquity and authenticity of the sequences, as does the detection of species whose ranges have contracted and are no longer documented at the sites, e.g. *Bettongia* (bettongs). There appears to be little or no environmental contamination as evidenced by the absence of any sequences from highly abundant invasive taxa including *Mus musculus* (house mouse) or *Rattus rattus* (black rat). Whereas downward contamination may be an issue at some sites[32], Devil's Lair contains several stratigraphical layers capped with calcite "flowstone"[33] preventing the movement of fossils, and likely DNA[5,23]. Whilst it is acknowledged that contamination can be cryptic and sporadic[34–36], the strict adherence to aDNA protocols[37], the use of sequence quality filters and the plausibility of the data (see Methods), greatly reduces the likelihood that contamination contributed to the data presented here.

Although most taxonomic assignments from DNA sequences confirmed previous morphological identification[23], some unexpected sequences resulted in distinct DTUs that were more difficult to assign. The issue is best exemplified by indeterminate Macropodidae sequences. It is unlikely that poor database coverage is the cause of this family-level assignment, as the Macropodidae database is nearly complete for both 16S and 12S rRNA mtDNA. In such cases sequencing error or DNA damage is also unlikely as the sequences are abundant and present across numerous samples at both sites, have passed all

quality filters, form distinct DTUs and are unlikely to be nuclear copies (Figures 2, 3 and 4). It is possible therefore that these sequences may arise from extinct lineages of present-day macropodids or indeed from extinct taxa. In some cases sequences mapped closest to species of the New Guinea forest wallaby (*Dorcopsis*) and the east Australian restricted pademelon (*Thylogale*). The presence of such 'indeterminate' DNA sequences in bulk-bone samples is intriguing. For example, two extinct tree-kangaroo species (genus *Bohra*[38,39]), have been described in caves along the Nullarbor Plain, yet tree-kangaroos of the genus *Dendrolagus* are only currently present in northeastern Queensland and New Guinea and were previously not thought to have occurred so far south[38]. It is a tantalizing prospect that 'indeterminate' DNA sequences could represent previously unknown species from southwest Western Australia, but it is also a problematic finding, as there is no easy way to uncover the fossils that contributed the DNA. It is likely that bulk-sampling methods such as this will generate genetically plausible taxa that lack morphological identifications. Arguably a similar result has already occurred with the single Denisovan finger bone from "X-woman" used to postulate a new lineage of archaic humans in Siberia[40,41].

When dealing with past biodiversity and aDNA sequences from fossil assemblages, analyses that are largely independent of taxonomy will likely be crucial to mapping temporal and/or spatial variation in genetic signatures. Such an approach facilitates the use of sequences that would otherwise be labeled "indeterminate", which will be commonly encountered when employing the bulk-bone HTS methodologies advocated here. While it is not possible to comprehensively analyse changes in biodiversity over time presented here from only a handful of samples such an analysis serves to illustrate how bulk-bone data could be approached. The data presented in Figures 3–5 should therefore be viewed tentatively, as further extensive replication and investigation is required to confirm any significant patterning over time.

Owing to the difficulties of definitively assigning sequences to a defined taxonomy, a modified OTU analysis (referred to as DTU), has been introduced to examine biodiversity change over time. It was clear from the initial analysis that OTU numbers were artificially inflated primarily by homopolymer error. When dealing with short sequences homopolymer errors can create a distinct OTU whereby the only difference between it and its closest OTU match is a base within a homopolymer stretch. It was observed that homopolymer-derived OTUs were more common in those samples with greater depth of sequencing coverage. To overcome this issue, an OTU alignment and Kimura 2-parameter distance matrix was adopted whereby errors in homopolymer stretches appear as gaps and homopolymer-derived OTUs collapse into a single DTU (See Methods). Whilst at these particular sites, it is a challenge to disentangle the roles of climate, DNA decay and past anthropogenic influences; shifts in DTU composition appear at the LGM and at the Holocene-Pleistocene transition (Figures 4 and 5). Furthermore, specific Macropodidae DTU analysis showed a reduction in DTU diversity and abundance over time, with a drop in diversity around the LGM (Figure 5). With these tentative patterns of biodiversity being derived from only 15 DNA extractions it is easy to conceptualize how, with adequate sampling and appropriate genetic markers, a bulk-bone sampling method will facilitate detailed mapping of faunal changes over time. Moreover, the method is cheaper than single bone approaches[42,43] while augmenting traditional morphological analysis.

The bulk-bone aDNA metabarcoding method used in this study presents a new, cost effective approach to identifying bulk quantities of morphologically indistinct bone fragments that otherwise end up in the taxonomic scrapheap. From modest amounts of sieved material across multiple layers at two study sites it was possible to detect equivalent diversity as described in previous morphological analyses[23]. While some taxa previously identified were not detected (most noticeably *Macropus* species), the converse was also true. This method is by no means an attempt to supplant traditional morphological approaches to taxonomic identification and analysis. Rather, it complements these approaches and by means of DTU analysis indicates changes in genetic diversity through time. Besides improving the identification of fossil assemblages the method allows researchers to rapidly assess the DNA preservation potential of freshly excavated material, which will vary from site to site. The approach will be equally applicable to archaeological and palaeontological sites, providing snapshots of past faunal diversity and human subsistence in both taxonomic dependent and independent ways. As such, it is anticipated that a bulk-bone approach will become a valuable part of the archaeological and palaeontological toolkit.

Methods

Sample collection and processing. Thousands of indistinct bone fragments were collected from both Tunnel Cave and Devil's Lair during excavations in February 2012. Approximately 150 L (0.15 m^3) of sediment was analysed at both sites. Sediment was dry-sieved on site, using 2 mm and 5 mm sieves, and bagged according to well-defined and dated stratigraphical layers[23]. Each bagged sample was screened for bone fragments off-site, which were kept in groupings according to the layers in which they were found. Fifteen bulk-bone samples representing fifteen layers were processed: eight from Tunnel Cave, covering a period from 4,160–24,110 years BP (uncalibrated)[23], and seven from Devil's Lair, covering a period from 6,200–46,890 years BP (uncalibrated)[23]. Small sections of the bones within each layer (typically 50–150 bones) were drilled (Dremel 114 drill bits) for a few seconds each and approximately equal amounts of drilled material from each bone fragment within a single layer was combined to form a "bulk-bone powder". Owing to inherent differences in the amount of DNA per unit of biomass between species and differential DNA preservation between individual bones, over-representation of certain bone material in terms of DNA amplicon sequences is unavoidable.

DNA extraction and screening. All laboratory work was conducted in keeping with standard aDNA protocols[28]. Approximately 1 g of bulk-bone powder from each sample, including a blank extraction control, was digested overnight on a lab rotator at 55°C in 5 mL of digestion buffer containing: 2.5 mL EDTA (0.5 M), 0.1 mL Tris-HCL (1 M), 5 mg Proteinase K powder, 50 μL DTT (1 M), 50 μL SDS and made up to a final volume of 5 mL using EDTA. DNA digests were centrifuged at 6,000 rpm for 2 mins and the supernatant was concentrated to 50 μL using AMICON 30,000 MWCO columns (Millipore) as per the manufacturer's instructions. Each concentrate was transferred to a clean 2 mL eppendorf tube and PBi buffer (Qiagen) totalling 250 μL (i.e. 5× the volume of concentrate) was added. Each 300 μL PBi/concentrate mix was subsequently transferred to Qiagen silica spin columns and centrifuged at 13,000 rpm. Columns were washed with 700 μL of AW1 followed by AW2. A final dry spin at 13,000 rpm for 1 min followed. DNA was eluted from the columns in 60 μL EB with a 1 min incubation at room temperature prior to centrifugation at 13,000 rpm for 1 min.

Extracts were screened for amplifiable mtDNA using multiple primer sets via qPCR at three concentrations - undiluted, 1/10 and 1/50. Extracts were screened for mammalian mtDNA using 12SA/O and 16Smam primer sets, designed to amplify a small region within mammalian 12S and 16S mitochondrial genes respectively[25,44]. Extracts were also tested for avian mtDNA using 12SA/E and 12SA/H primer sets, designed to amplify a short and slightly longer overlapping region of the avian mitochondrial 12S gene respectively[44]. Finally, extracts were tested for snake mtDNA using the following primers: 12s_tRNA_F1_S AAAGTATAGCACTGAAAATGC TAA and 12s_R1_Snake GTTAGCCTGATACCGGCTCCG, designed to amplify a short region within the mitochondrial 12S gene. Each qPCR reaction was made up to a total volume of 25 μL, containing 1× PCR Gold Buffer (Applied Biosystems), 2.5 mM $MgCl_2$ (Applied Biosystems), 0.4 mg/mL BSA (Fisher Biotech, Aus), 0.25 mM of each dNTP (Astral Scientific, Aus), 0.4 μM forward primer, 0.4 μM reverse primer, 0.25 μL AmpliTaq Gold (Applied Biosystems), 0.6 μL SYBR Green (1 : 2,000, Life Sciences gel stain solution) and 2 μL DNA extract. Quantitative PCR cycling conditions for the 12SA/O and snake 12S qPCR assays were as follows: initial heat denaturation at 95°C for 5 mins, followed by 50 cycles of 95°C for 30 s; 55°C for 30 s (annealing step); 72°C for 45 s followed by a 1°C melt curve and final extension at 72°C for 10 mins. Cycling conditions for 16Smam, 12SA/E and 12SA/H assays were the same as for the 12SA/O assay, except the annealing temperature, which was 57°C in each case. For each qPCR assay, DNA extraction, negative PCR reagent and positive DNA template controls were included.

DNA sequencing. DNA extracts that successfully yielded DNA of sufficient quality, free of inhibition, as determined by initial qPCR screening[45], were prepared for amplicon sequencing. DNA extracts successful for all primer sets were sequenced on Roche's GS-Junior. Additional, separate, amplicon sequences were generated for extracts using mammalian 12SA/O and 16Smam primer sets for sequencing on Life Technologies' Ion Torrent Personal Genome Machine (PGM).

For each primer set, DNA extracts were assigned a unique DNA tag[11]. Each sample was tagged at both the 5′ and 3′ end of the target sequence using separate tags at both

ends, resulting in a unique forward and reverse tag combination for each sequence. Independent tagged qPCRs for all samples, across all primer sets, were carried out in 25 μL reactions with reaction components and cycling conditions as described in 'Methods: DNA extraction and screening'. Tagged qPCR amplicons were generated in triplicate and combined, thus minimizing the effects of PCR stochasticity on low-template samples, purified using Agencourt AMPure XP PCR Purification Kit (Beckman Coulter Genomics, NSW, Aus), as per manufacturer's instructions and eluted in 40 μL H_2O. Purified amplicons were pooled to form separate sequencing libraries according to primer set used and sequencing platform. GS-Junior libraries were quantified using qPCR to determine an appropriate volume of library for sequencing (described in Murray *et al.* 2011). Each 25 μL reaction contained 12.5 μL ABI Power SYBR master mix (Applied Biosystems), 0.4 μM A-adapter primer, 0.4 μM B-adapter primer, 8.5 μL H_2O and 2 μL pooled library, with the following cycling conditions: 95°C for 5 mins; 40 cycles of 95°C for 15 s, 56°C for 1 min followed by a 1°C melt curve. The appropriate library volume for use on the Ion Torrent PGM was determined using a Bioanalyser 2100 (Agilent). For each tagged qPCR assay, negative qPCR controls were included and if found to contain amplifiable DNA these qPCR amplicons were incorporated into the appropriate pooled sequencing library. All sequencing was performed as per manufacturer's instructions, with the use of 200 bp reagents and a 314 chip on the PGM.

Sequence identification. Amplicon sequence reads (hereafter referred to as sequences) were sorted into sample batches based on unique DNA tags. Identification tags and primers were trimmed allowing for no mismatch in length or base composition using Geneious v6.0.5 (created by Biomatters, available from http://www.geneious.com/). Batched and trimmed sequences from both GS-Junior and Ion Torrent PGM sequencing runs were combined according to sample and primer used. Each combined file was dereplicated, thus grouping sequences of exact identity and length, using USEARCH[46]. Dereplicated sequence files were searched for artificial chimeric sequences using the UCHIME *de novo* method[47] in USEARCH and were removed, in addition to sequences occurring only once (i.e. singletons). The remaining sequences in each sample were subsequently clustered at an identity threshold of 97% using USEARCH with the most abundant sequence within each cluster selected as the representative sequence. To reduce noise associated with sequencing error, low abundant clusters, classed as those that occur at less than 1% of the total number of unique sequences when clustered at 100% sequence identity, were removed from the dataset. While the selection of a 1% cut-off is somewhat arbitrary, it should negate the possibility of clusters remaining that are the result of sequencing error. Additionally, the decision to class clusters as being in low abundance with respect to the total number of unique sequences (as opposed to total number of sequences or total number of sequences within the most abundant cluster) was made to minimize the effects of preferential DNA preservation and/or amplification. For each sample, every sequence assigned to the remaining clusters were queried against the NCBI GenBank nucleotide database using BLASTn[48] in YABI[49], enabling taxonomic identification. Sequences were searched without a low complexity filter, with a gap penalties existence of five and extension of two, expected alignment value less than 1e-10 and a word count of seven. The BLASTn results obtained were imported into MEtaGenome Analyzer v4 (MEGAN), where they were mapped and visualised against the NCBI taxonomic framework (min. bit score = 35.0, top percentage = 5%, min. support = 1)[50]. Sequences that were obviously the result of contamination (primarily human and cow) were eliminated from all subsequent downstream analysis steps.

Sequences that were truncated when queried against the NCBI GenBank nucleotide database were discarded from taxonomic analysis. Sequences with percentage similarity to a reference below 90% were discarded. Where sequence similarities were between 90–95% these were assigned to a family level, while those between 95–100% were assigned to a genus. Owing to the difficulties in assigning taxa beyond the genus level for some families, in addition to issues associated with characterizing past biodiversity that has been lost, species identifications were avoided in this particular study. Sequences that provided high percentage similarity to query references at a species level may or may not be bona fide, however with current insufficient data it is prudent to categorise these sequences cautiously. Where multiple taxa had equal percentage similarity scores to a query sequence, such sequences were moved higher up the taxonomic rankings.

While the validity of filters and hard percentage cut-offs are always debatable, those chosen in the analysis of this dataset seemed to afford the best balance when accounting for low template amounts and post-mortem damage on short aDNA fragments.

Genetic biodiversity analysis. Cognisant of the difficulties associated with assigning sequences to lower taxonomic levels, a modified form of OTU analysis was applied to the 16Smam sequences obtained in this study. This allowed changes in observed genetic diversity over time at both sites to be investigated independently of the above taxonomic classifications. Sequences within each sample were clustered at 97% identity, filtered and representative sequences were selected as detailed in Methods: Sequence Identification. Representative sequences within each sample were aligned in Geneious using MAFFT's G-INS-I algorithm and default parameters[51]. MAFFT alignments were imported into MEGA5[52] where a distance matrix between OTUs within a sample was calculated using a Kimura 2-parameter model[53], with all positions containing gaps and missing data ignored. OTUs less than 3% divergent from each other were collapsed into a single DTU. This serves the purpose of reducing the influence of HTS homopolymer sequencing error[26,27] by collapsing multiple homopolymer-derived OTUs into a single DTU, as errors in homopolymer stretches appear as gaps and are not included in the calculation of the distance matrix. Whilst this is first and foremost a largely taxonomic-independent analysis it is still nonetheless useful to identify coarsely to which family each DTU belongs, as this gives an idea of the diversity of DTUs within specific families. As such, all DTUs were searched against the NCBI GenBank nucleotide database using BLASTn[48] to identify the family to which each DTU could be easily assigned. For the faunal specific Macropodidae DTU analysis the same method as above was followed except that only sequences assigned to Macropodidae were selected.

1. Archibald, S. B., Greenwood, D. R. & Mathewes, R. W. Seasonality, montane beta diversity, and Eocene insects: Testing Janzen's dispersal hypothesis in an equable world. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **371**, 1–8 (2013).
2. Colonesea, A. C. *et al.* Holocene snail shell isotopic record of millennial-scale hydrological conditions in western Mediterranean: Data from Bauma del Serrat del Pont (NE Iberian Peninsula). *Quat. Int.* **303**, 43–53 (2013).
3. Dortch, J. & Wright, R. Identifying palaeo-environments and changes in Aboriginal subsistence from dual-patterned faunal assemblages, south-western Australia. *J. Archaeol. Sci.* **37**, 1053–1064 (2010).
4. Rohland, N. *et al.* Genomic DNA Sequences from Mastodon and Woolly Mammoth Reveal Deep Speciation of Forest and Savanna Elephants. *PLoS Biol.* **8**, e1000564 (2010).
5. Haile, J. *et al.* Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22352–22357 (2009).
6. Worobey, M. *et al.* Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature.* **455**, 661–664 (2008).
7. Raoult, D. *et al.* Molecular identification by "suicide PCR" of *Yersinia pestis* as the agent of Medieval Black Death. *Proc. Natl. Acad. Sci. USA.* **97**, 12800–12803 (2000).
8. Smith, C. I. *et al.* Neanderthal DNA. Not just old but old and cold? *Nature.* **410**, 771–772 (2001).
9. Bonnichsen, R. *et al.* Methods of the study of ancient hair: Radiocarbon dates and gene sequences from individual hairs. *J. Archaeol. Sci.* **28**, 775–785 (2001).
10. Oskam, C. L. *et al.* Fossil avian eggshell preserves ancient DNA. *Proc. R. Soc. Biol. Sci. Ser. B.* **277**, 1991–2000 (2010).
11. Binladen, J. *et al.* The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS ONE.* **2**, e197 (2007).
12. Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* **21**, 1794–1805 (2012).
13. Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* **21**, 2045–2050 (2012).
14. Jørgensen, T. *et al.* A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Mol. Ecol.* **21**, 1989–2003 (2012).
15. Jørgensen, T. *et al.* Islands in the ice: detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA Meta-barcoding. *Mol. Ecol.* **21**, 1980–1988 (2011).
16. Deagle, B., Chiaradia, A., McInnes, J. & Jarman, S. Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conserv. Genet.* **11**, 2039–2048 (2010).
17. Murray, D. C. *et al.* High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens. *Quat. Sci. Rev.* **58**, 135–145 (2012).
18. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
19. Little, D. P. DNA barcode sequence identification incorporating taxonomic hierarchy and within taxon variability. *PLoS ONE.* **6**, e20552 (2011).
20. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods.* **7**, 335–336 (2010).
21. Schloss, P. D. & Handelsman, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**, 1501–1506 (2005).
22. Hamady, M., Lozupone, C. & Knight, R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *The ISME Journal.* **4**, 17–27 (2010).
23. Dortch, J. *Palaeo-environmental Change and the Persistence of Human Occupation in South-western Australian Forests.* (Archaeopress, Oxford, 2004).
24. Myers, N., Mittermeier, R. A., Mittermeier, C. G., de Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature.* **403**, 853–858 (2000).
25. Taylor, P. G. Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Mol. Biol. Evol.* **13**, 283–285 (1996).
26. Loman, N. J. *et al.* Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* **30**, 434–439 (2012).
27. Quail, M. A. *et al.* A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics.* **13**, 341 (2012).
28. Cooper, A. & Poinar, H. N. Ancient DNA: Do it right or not at all. *Science.* **289**, 1139–1139 (2000).

29. Coissac, E., Riaz, T. & Puillandre, N. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Mol. Ecol.* **21**, 1834–1847 (2012).
30. Taylor, H. R. & Harris, W. E. An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. *Mol. Ecol. Resour.* **12**, 377–388 (2012).
31. Jønsson, K. A. & Fjeldså, J. A phylogenetic supertree of oscine passerine birds (Aves: Passeri). *Zoologica Scripta.* **35**, 149–186 (2006).
32. Haile, J. *et al.* Ancient DNA chronology within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor? *Mol. Biol. Evol.* **24**, 982–989 (2007).
33. Turney, C. & Bird, M. I. Early human occupation at Devil's Lair, southwestern Australia 50,000 years ago. *Quatern. Res.* **55**, 3–13 (2001).
34. Champlot, S. *et al.* An Efficient Multistrategy DNA Decontamination Procedure of PCR Reagents for Hypersensitive PCR Applications. *PLoS ONE.* **5**, e13042 (2010).
35. Erlwein, O. *et al.* DNA Extraction Columns Contaminated with Murine Sequences. *PLoS ONE.* **6**, e23484 (2011).
36. Tuke, P. W., Tettmar, K. I., Tamuri, A., Stoye, J. P. & Tedder, R. S. PCR Master Mixes Harbour Murine DNA Sequences. Caveat Emptor! *PLoS ONE.* **6**, e19953 (2011).
37. Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M. & Barnes, I. Assessing ancient DNA studies. *Trends. Ecol. Evol.* **20**, 541–544 (2005).
38. Prideaux, G. J. & Warburton, N. M. A new Pleistocene tree-kangaroo (Diprotodontia: Macropodidae) from the Nullarbor Plain of south-central Australia. *J. Vert. Paleontol.* **28**, 463–478 (2008).
39. Prideaux, G. J. & Warburton, N. *Bohra nullarbora* sp. nov., a second tree-kangaroo (Marsupialia:Macropodidae) from the Pleistocene of the Nullarbor Plain, Western Australia. *Rec. West. Aust. Mus.* **25**, 165–179 (2009).
40. Krause, J. *et al.* The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature.* **464**, 894–897 (2010).
41. Meyer, M. *et al.* A high-coverage genome sequence from an archaic Denisovan individual. *Science.* **338**, 222–226 (2012).
42. Shapiro, B. *et al.* Rise and fall of the Beringian steppe bison. *Science.* **306**, 1561–1565 (2004).
43. Lorenzen, E. D. *et al.* Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature.* **479**, 359–364 (2011).
44. Cooper, A. *et al.* Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature.* **409**, 704–707 (2001).
45. Bunce, M., Oskam, C. & Allentoft, M. in *Ancient DNA: Methods and Protocols Methods in Molecular Biology* (eds Shapiro, B. & Hofreiter, M.) Ch. **16**, 121–132 (Humana Press, New York City, 2011).
46. Edgar, R. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* **26**, 2460–2461 (2010).
47. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics.* **27**, 2194–2200 (2011).
48. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. GenBank. *Nucleic Acids Res.* **34**, D16–D20 (2006).
49. Hunter, A. A., Macgregor, A. B., Szabo, T. O., Wellington, C. A. & Bellgard, M. I. Yabi: An online research environment for grid, high performance and cloud computing. *Source Code Biol. Med.* **7**, 1 (2012).
50. Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **17**, 377–386 (2007).
51. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
52. Tamura, K. *et al.* MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
53. Kimura, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).

Acknowledgments

We thank the Wardandi people, and the Webb family in particular, as traditional owners and custodians of the region, for supporting our excavations and analyses. We thank student volunteers for their assistance in the field, Nina Kresoje (UWA) and Vanessa Atkinson (Pathwest Laboratory Medicine WA) and Frances Brigg for sequencing assistance. We also thank the WA Museum and the Curator of Anthropology Moya Smith for access to museum resources, and iVEC for computational support. Australian Research Council grants DP120103725 (to M.B., J.D. and J.H.) and FT0991741 (to M.B.) funded the research.

Author contributions

D.C.M., M.B. and J.H. designed the experiments. D.C.M., J.H., N.W., D.H. and J.D. excavated and prepared samples. D.C.M., J.H., M.I.B., D.H. and R.A. contributed to HTS data generation and bioinformatics. J.D. provided stratigraphic interpretations and G.P. and J.D. provided fossil and taxon interpretations. D.C.M. and M.B. wrote the paper with assistance from all co-authors.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Murray, D.C. *et al.* Scrapheap Challenge: A novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Sci. Rep.* **3**, 3371; DOI:10.1038/srep03371 (2013).