Tracking dogs across the Pacific: an archaeological and ancient DNA study

Karen Greig

A thesis submitted for the degree of Doctor of Philosophy
University of Otago, Dunedin, New Zealand

December 2016
Abstract

The dispersal of dogs across the Pacific region is inseparably linked to the relationships between dogs and people. Unlike movement across continental landmasses, dogs must have been transported by people across the waters that separate islands. The purpose of this PhD research is to investigate the introduction and dispersal of dogs throughout the Pacific. The approach uses dog remains from archaeological sites to generate ancient mitochondrial genomes using next generation sequencing technology. This molecular data is interpreted in relation to regional archaeological models for human and animal movements and human-dog interactions. In order to understand as fully as possible what the nature of the dog-human relationship may have been like, ecological, ethnographic and historical information about dog populations and their interactions with people in the region are also considered.

The outcomes of the ancient DNA analyses of dog specimens from colonisation era archaeological sites in New Zealand demonstrates the impact that the interactions between people and dogs can have on mitochondrial molecular diversity, and hence the usefulness of studies using only maternal markers. The results, nonetheless, are informative about the Pacific dogs brought to New Zealand and the colonisation process. The homogeneity of mtDNA lineages from this first introduction suggests a single, closely-related founding population, and supports the current archaeological model of rapid and strategic colonisation.

Complete and partial ancient mitogenomes were generated from archaeological specimens throughout Southeast Asia and the wider Pacific to investigate the place of dogs in the Lapita migrations. These results indicate the introduction of at least three different dog lineages to the Pacific region, each with a different dispersal history. An association between Late Lapita dogs and modern Taiwanese dogs was found that suggests the possibility that dogs may have been part of the expansion of Austronesian language speakers associated with the Lapita Cultural Complex, but were not successfully transported by groups moving beyond the Bismarck Archipelago. A major Pacific dog clade was observed, which was a relatively late but highly successful introduction. This lineage was found in archaeological specimens
across the Pacific, including several islands in Polynesia. There appears to be a discontinuity between Lapita era mtDNA lineages and later East Polynesian lineages.

This research has resulted in new data generated from ancient DNA analyses of complete mitochondrial genomes of dogs from archaeological samples from Island Southeast Asia and the Pacific. The results demonstrate the complexity of dog introductions and dispersals in the region, with implications for understanding human colonisation processes and the ways in which dogs may have been moved around the Pacific.
Acknowledgements

This research was made possible by an University of Otago Doctoral Scholarship. I am also thankful for funding from the Skinner Fund of the Royal Society of New Zealand, the Department of Anthropology’s Postgraduate Research Fund, and for the opportunities to attend conferences provided by Divisional conference funding and the Department of Anatomy’s Biological Anthropology Research Group. Being able to meet people from around the world and discuss shared research interests has been a stimulating and valuable part of my thesis journey.

I would like to acknowledge Rangitane ki Wairau who gave permission for me to study the dog bone assemblage from Wairau Bar. I would also like to acknowledge Associate Professor Ian Smith for access to archaeological material from Shag Mouth and Pleasant River, Dr Louise Furey, Auckland Museum, and the Department of Anthropology at Auckland University, for access to the Houhora collection, and Professor Lisa Matisoo-Smith for enabling me to work on archaeological material from Southeast Asia and the Pacific.

I gratefully acknowledge Professors Lisa Matisoo-Smith and Richard Walters’ supervision of this thesis. I would also like to acknowledge the encouragement I received, particularly at the early stages of this project, from Foss Leach, Janet Davidson, Ian Smith and Helen Leach. This thesis would not have been possible without the support of the staff and students of the Matisoo-Smith and Gemmell lab groups in the Department of Anatomy, especially at the start of my laboratory work when I was making the transition from working with bones to molecules. I am particularly indebted to Ann Horsburgh, Olga Kardailsky and Anna Gosling for assistance in the laboratory, and Stefan Prost and James Boocock for their bioinformatic wizardry. The support staff in the Department of Anthropology have been very helpful and kind, and Les O’Neill produced the excellent illustrations for my publications.

I would also like to acknowledge the friendship and support I have shared along the way with fellow post-graduate students, especially my PhD officemates in the Anatomy and Anthropology Departments, Sally and Lara who shared the highs and lows of combining a
PhD with family life, and Emma who finished the journey with me. Finally, I would like to thank my family, especially Richard, Rosa, Jamie, Jim and Esther, and friends outside the university for their unfailing support and faith in me for which I am deeply grateful.

The dog on the front page is based on a sketch of a Pacific dog in a previously unpublished work probably made during Captain Cook’s first voyage, which is reproduced in Figure 5 in Katherine Luomala’s 1960 publication ‘A history of the binomial classification of the Polynesian native dog’ in the Pacific Science journal. Dylan Gaffney assisted with digitising my original drawing.
# Table of Contents

Abstract .......................................................................................................................... ii

Acknowledgements ........................................................................................................ iv

List of Tables .................................................................................................................. x

List of Figures ................................................................................................................ xi

List of Abbreviations ...................................................................................................... xii

Chapter 1 Introduction .................................................................................................... 1
1.1 Historical biogeography of Island Southeast Asia and the Pacific ......................... 2
1.2 Pacific dogs ............................................................................................................... 6
   1.2.1 Archaeological evidence from the Pacific ....................................................... 6
   1.2.2 Molecular phylogenetic and phylogeographic studies ................................... 8
1.3 Motivation for this study ........................................................................................ 10
1.4 Research aim and questions .................................................................................. 13
1.5 Significance of this research .................................................................................. 15
1.6 Organisation of this thesis ..................................................................................... 15
1.7 Note on published works ...................................................................................... 17

Chapter 2 Dog origins and dispersal across Southeast Asia and the Pacific .................. 19
2.1 Introduction ............................................................................................................. 19
2.2 Domestication ......................................................................................................... 19
2.3 Dog domestication .................................................................................................. 21
   2.3.1 Single East Asian origins ............................................................................... 23
   2.3.2 Multiple origins ............................................................................................... 23
2.4 Dog dispersal .......................................................................................................... 24
2.5 Distribution of dogs ............................................................................................... 26
   2.5.1 East Asian, Mainland and Island Southeast Asian dogs ................................. 26
   2.5.2 Australia, New Guinea and the Torres Strait .................................................. 29
   2.5.3 Near Oceania .................................................................................................. 31
   2.5.4 Remote Oceania ............................................................................................. 31
   2.5.5 Central and East Polynesia ............................................................................. 33
   2.5.6 Micronesia ....................................................................................................... 36
2.6 Linguistics ................................................................................................................ 37
2.7 Summary .................................................................................................................. 39

Chapter 3 Ecology, behaviour and human interactions of free-ranging dogs ............... 40
3.1 Introduction ............................................................................................................ 40
3.2 Ecological studies of free-ranging dog populations........................................41
  3.2.1 Population structure..................................................................................43
  3.2.2 Life expectancy.......................................................................................44
  3.2.3 Sex ratio..................................................................................................45
  3.2.4 Diet and feeding behaviour......................................................................46
  3.2.5 Social organisation..................................................................................47
3.3 Interactions with people................................................................................49
  3.3.1 Island Southeast Asia...............................................................................49
  3.3.2 New Guinea and Australia.......................................................................51
  3.3.3 New Zealand...........................................................................................52
3.4 Characteristics of free-ranging dog populations.........................................55

Chapter 4 Models and Methods........................................................................58
  4.1 Archaeological models for human migrations in the Pacific.......................59
    4.1.1 The Lapita Cultural Complex...............................................................59
    4.1.2 The colonisation of New Zealand.......................................................63
  4.2 Molecular phylogenetics..........................................................................64
    4.2.1 DNA.................................................................................................65
    4.2.2 Phylogenetic and genealogical analyses............................................70
    4.2.3 Graphic representations and statistics..............................................73
  4.3 Ancient DNA..............................................................................................79
    4.3.1 Extraction and sequencing....................................................................79
    4.3.2 Assembly............................................................................................80
    4.3.3 Analysis...............................................................................................81
  4.4 Molecular genetic inference from archaeological assemblages..................82
    4.4.1 Natural, census and hypothetical populations....................................83
    4.4.2 Archaeological assemblages................................................................83
  4.5 Old dogs and new tricks..........................................................................86

Chapter 5 Complete Mitochondrial Genomes of New Zealand's First Dogs........87
  5.1 Introduction...............................................................................................87
  5.2 Materials and methods..............................................................................92
    5.2.1 Samples.............................................................................................92
    5.2.2 Extraction, hybridisation capture and sequencing..............................94
    5.2.3 Raw data processing and analysis.....................................................95
    5.2.4 Phylogenetic Analyses.......................................................................97
  5.3 Results.......................................................................................................98
    5.3.1 DNA preservation and sequence recovery.........................................98
    5.3.2 Sequence authenticity.........................................................................99
    5.3.3 Population structure of Wairau Bar dogs...........................................99
    5.3.4 Complete mitogenomes.................................................................101
    5.3.5 Control region haplotypes...............................................................102
  5.4 Discussion................................................................................................104
  5.5 Summary..................................................................................................106
Chapter 6 Dog introduction and dispersal during the colonisation of New Zealand..... 108
  6.1 Introduction........................................................................................................108
  6.2 Materials and methods......................................................................................110
    6.2.1 Samples and sites......................................................................................110
    6.2.2 aDNA methods.........................................................................................116
  6.3 Results...............................................................................................................117
    6.3.1 DNA preservation, sequence recovery and authenticity.......................117
    6.3.2 Genetic population structure.....................................................................117
  6.4 Discussion........................................................................................................119
    6.4.1 Composition of the founding population.............................................119
    6.4.2 Origins of New Zealand’s first dogs......................................................120
    6.4.3 Dog dispersal during the colonisation phase of New Zealand...............121
    6.4.4 Introduction and establishment of Oceanic dogs.................................122
  6.5 Summary .........................................................................................................124

Chapter 7 Ancient mitogenomes of South East Asian and Pacific dogs shows complex regional history................................................................. 125
  7.1 Introduction......................................................................................................125
  7.2 Materials and Methods..................................................................................127
    7.2.1 Ancient DNA extraction, library preparation, sequencing and data processing 127
    7.2.2 Modern dingo mitogenomes......................................................................129
    7.2.3 Phylogenetic reconstructions...................................................................129
  7.3 Results and discussion...................................................................................129
    7.3.1 East Asian ancestry for Australasian and Pacific dogs.........................130
    7.3.2 Multifaceted introduction of dingoes to Australia..................................132
    7.3.3 New lineage of Late Lapita dogs.............................................................136
    7.3.4 Later introduction of Remote Oceanic dogs..........................................137
    7.3.5 Origins of Pacific dogs............................................................................139
  7.4 Summary .......................................................................................................140

Chapter 8 Discussion .......................................................................................... 142
  8.1 Genetic variation in ancient mitogenomes of Pacific dogs and the investigation of genetic relationships and ancestry.........................................................143
    8.1.1 Sequence recovery and authenticity.......................................................143
    8.1.2 Genetic diversity......................................................................................145
    8.1.3 Phylogenetic, population genetic and phylogeographic analyses ..........147
  8.2 Phylogenetic analyses compared with archaeological models for human colonisation in the Pacific.................................................................149
    8.2.1 New Zealand.............................................................................................150
    8.2.2 Lapita Cultural Complex..........................................................................151
    8.2.3 Alternative archaeological models.........................................................153
  8.3 Contribution of ancient mtDNA for understanding dog and human mobility and interactions in Australasia and the Pacific...........................................158
    8.3.1 Subject and object....................................................................................159
    8.3.2 The place of dogs in human communities.............................................162
List of Tables

Table 1.1: Published material and material submitted or prepared for publication included in this thesis................................................................. 17

Table 1.2: Contributions of co-authors to published works and manuscripts to be submitted for publication.......................................................... 18

Table 2.1: Archaeological evidence for dogs in East Asia, MSEA and ISEA............. 27

Table 2.2: Archaeological evidence for dogs in dogs in Remote Oceania ................. 33

Table 3.1: Categories of dog populations.................................................................... 42

Table 6.1: New Zealand and Southern Cook Island archaeological specimens sequenced for this study........................................................................ 112

Table 7.1: Ancient MSEA, ISEA and Pacific samples sequenced for this study......... 128
List of Figures

Figure 1-1: Map of Island Southeast Asia and Australia, showing the Pleistocene landmasses of Sunda and Sahul, and Wallacea................................................................. 3

Figure 1-2: Map showing the islands of the Pacific, the boundary between Near and Remote Oceania, and the Andesite Line............................................................... 4

Figure 2-1: Map showing the natural distribution of wolves, the proposed centre of dog domestication and selected archaeological sites with dog remains in Mainland and Island Southeast Asia................................................................. 25

Figure 2-2: Map showing early distribution of dogs in the Pacific, from archaeological contexts and reported by eighteenth and nineteenth century European voyages........ 32

Figure 4-1: The structure of an animal cell................................................................. 65

Figure 4-2: A mitochondrial organelle....................................................................... 66

Figure 4-3: The structure of nuclear DNA................................................................. 67

Figure 5-1: Map of the Pacific region. Arrows indicate the proposed route for the introduction of dogs to Polynesia, via Island South East Asia................................. 88

Figure 5-2: Wairau Bar archaeological site................................................................. 93

Figure 5-3: Median-joining networks for Wairau Bar dog sequences....................... 100

Figure 5-4: Median-joining network of complete mitochondrial genomes of Wairau Bar dogs, and modern and ancient sequences from Haplogroup A......................... 101

Figure 5-5: Median-joining network of control region sequences (582bp) of Wairau Bar dogs, and modern sequences from Haplogroup A.................................................. 103

Figure 6-1: Map of Polynesia, showing New Zealand and the Cook Islands, inset Southern Cook Islands................................................................. 113

Figure 6-2: Locations of archaeological sites in New Zealand and the Southern Cook Islands and a median-joining haplotype network of mitogenomes of dogs.............. 118

Figure 7-1: Median joining network of mitogenome sequences from ancient Island Southeast Asian and Pacific dogs, modern dingoes sequenced for this study and dingoes, NGSD and Southeast Asian dogs from GenBank ................................................. 131

Figure 7-2: Molecular phylogenetic analysis by maximum likelihood method............ 133

Figure 7-3: Distribution of ancient dog lineages sequenced from MSEA, ISEA and Pacific archaeological sites................................................................. 140
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aDNA</td>
<td>Ancient DNA</td>
</tr>
<tr>
<td>BCE</td>
<td>Before Common Era</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BP</td>
<td>Before Present</td>
</tr>
<tr>
<td>CE</td>
<td>Common Era</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>ISEA</td>
<td>Island Southeast Asia</td>
</tr>
<tr>
<td>LCC</td>
<td>Lapita Cultural Complex</td>
</tr>
<tr>
<td>ML</td>
<td>maximum likelihood</td>
</tr>
<tr>
<td>MSEA</td>
<td>Mainland Southeast Asia</td>
</tr>
<tr>
<td>mtDNA</td>
<td>mitochondrial DNA</td>
</tr>
<tr>
<td>nucDNA</td>
<td>nuclear DNA</td>
</tr>
<tr>
<td>NGS</td>
<td>Next Generation Sequencing</td>
</tr>
<tr>
<td>NGSD</td>
<td>New Guinea Singing Dog</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>SNP</td>
<td>Single Nucleotide Polymorphism</td>
</tr>
<tr>
<td>VC-III</td>
<td>Voyaging Corridor Triple I (Green 1991b)</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

When the first groups of people spread from west to east across the islands of the Pacific they didn’t travel alone. Often dogs, pigs, chickens and rats were part of these migrations, and the bones of these animals are routinely found in archaeological sites. Their presence comes as no surprise, as the tendency to form complex relationships with other species is a notable human trait (O’Connor, 2010), and at least some of these migrants came from cultures where domesticated animals were a long-standing part of their way of life (Bellwood, 2011). Nonetheless, transporting and establishing domesticated animals across increasingly wider ocean gaps and onto often smaller, and more environmentally depauperate, islands of the Pacific, with no established agricultural subsistence base, must have posed considerable challenges. The presence of the three domesticated animals, pigs, dogs and chickens, has been incorporated into archaeological models for the human settlement of the Pacific, particularly the concept of the ‘transported landscape’. This concept describes a process where migrating groups moved across the Pacific taking with them a set of horticultural concepts, plants and animals that were then utilised on the new islands they settled (Kirch, 2000). Yet there are still many aspects of these remarkable journeys that remain little known. The introduction of domesticated animals into the region (as with people) was not a systematic process, rather the archaeological evidence for these animals is patchy and discontinuous (Anderson, 2009, Matisoo-Smith, 2007) which may reflect some of the difficulties inherent in the process, and also the choices that people may have made in relation to transportation and animal management.

My PhD research investigates the introduction and dispersal of one of these domesticated animals, the dog (*Canis familiaris*), across the Pacific, drawing on data from archaeological, archaeozoological and ancient DNA (aDNA) methods of analysis. My approach uses dog remains from archaeological sites to produce new aDNA data generated using Next Generation Sequencing (NGS) technology. This data is interpreted in relation to
archaeological models for human and animal movements, and human-dog interactions across the region. The results contribute to a greater understanding of the genetic diversity of Pacific dogs, the timing and nature of their dispersal, and the founding population of New Zealand dogs.

This chapter consists of seven sections: an overview of the historical biogeography of Island Southeast Asia (ISEA) and the Pacific, a brief background about dog domestication and the introduction of dogs to this region, the motivation for the study, the purpose and research questions, the significance of the research, and finally an overview of subsequent chapters along with a note on published material arising from this thesis.

1.1 **Historical biogeography of Island Southeast Asia and the Pacific**

The movement of dogs and people through ISEA and the Pacific took place across a region that has an incredibly diverse geology, geography, biology and human history. Any investigation of animal introductions and dispersals in the region must take this diversity, and the constraints or opportunities it provided, into account. The biodiversity of the region is patterned by the underlying geological and climatic characteristics, and all these aspects have influenced the human history of the region in varying degrees. The main aspects are highlighted below.

ISEA encompasses the island chains and islands that lie between continental Asia and the large islands of New Guinea and Australia. Fluctuations in sea levels over the past several million years resulted in episodes when two large landmasses, Sunda and Sahul, were formed, separated by 1,000–1,500 km of the Wallace Straits (Hall, 2009) (Figure 1-1). Sunda comprised a large peninsula extending from the mainland Asian continental shelf to Borneo and Java, while New Guinea, Australia, and Tasmania formed the landmass called Sahul. Even during episodes of low sea level Wallacea, being the area between the two, has never been spanned by dry land. Wallace’s Line marks a boundary between Asian fauna, and that of Australia, initially based on 19th century observations. Despite this separation, the flora and fauna across the boundary is diverse, comprising placental and marsupial
mammals, birds and reptiles. The biodiversity of the islands of the Pacific differs further from that of the continental land masses, mainly because of the differences in dispersal between biota, and the high levels of endemism arising from isolation (Kirch, 2000).

![Map of Island Southeast Asia and Australia showing the Pleistocene landmasses of Sunda and Sahul, and the area of Wallacea.](image)

Figure 1-1: Map of Island Southeast Asia and Australia, showing the Pleistocene landmasses of Sunda and Sahul, and the area of Wallacea (Matisoo-Smith, 2015, Figure 1, permission to re-use image obtained from RightsLink, Licence No. 4011141380388).

Rising sea levels in the late Pleistocene resulted in the current environment, with the separation by water of many ISEA islands, the large island of New Guinea, continental Australia, and Tasmania occurring by 8–10,000 years ago. The islands of New Guinea, the Bismarck Archipelago and the Solomon Islands extend across the area defined as Near Oceania by Green (1991a) (Figure 1-2). To the north, east and southwest east lies Remote
Oceania, a vast area of ocean interspersed with relatively small islands with a different biogeography and human history.

Figure 1-2: Map showing the islands of the Pacific, the boundary between Near and Remote Oceania (after Green 1991a), and the Andesite Line.

The human colonisation of Sahul, based on archaeological evidence, dates to between 40,000 to 50,000 years ago (Summerhayes et al., 2010). In order to reach Sahul and the Bismarck Archipelago, people needed to be able to cross several open water barriers. Although it has been suggested that these voyages may have been accidental, archaeological evidence including fish bones and fishing technology suggests that people were conducting deliberate voyages during the colonisation of Sahul (O’Connor et al., 2011). Archaeological evidence for people subsequently reaching the Solomon Islands dates to not much later around 35,000 years ago (Kirch, 2000).

Compared with the antiquity of human occupation in Near Oceania, the islands of Remote Oceania were not colonised by people until about 3,200 years ago (Kirch, 2000). The human settlement of Remote Oceania commenced during the migration of people identified in the
archaeological record by the appearance of a distinctive pottery, known as ‘Lapita’. This pottery is present in numerous archaeological sites that extend from Papua New Guinea to the western margins of Polynesia, and spans the Near/Remote Oceanic boundary. Historical linguistic reconstructions indicate that these people spoke languages from the Austronesian language family (Blust, 2013). In addition to pottery and languages, the Lapita people are associated with the introduction of agricultural practices, a suite of domesticated animals and plants, a settlement pattern of villages situated on intertidal reefs or small offshore islands, and a characteristic set of artefacts including adzes and shell ornaments (Specht et al., 2014).

Remote Oceania is further sub-divided by the Andesite Line, which runs between Fiji and West Polynesia. This line follows the Pacific Plate boundary, and the formation processes of islands on either side of the line varies. To the west are island-arc or continental islands that lie along the plate boundary, and have complex geological histories. To the east are mid-plate islands formed by volcanic eruptions. The action of erosion, subsidence, uplift and coral reef formation on these volcanic islands can result in the formation of atolls, and makatea (raised coral) islands. Beyond the Andesite Line to the east, islands are generally much smaller, further apart, and have very limited stone resources for tool manufacture. Lapita people moving out into the islands of Polynesia would have encountered a much different environment to that of the west.

The settlement of Polynesia occurred in two phases, beginning with the islands of West Polynesia around 3,000 years ago during the Lapita expansion (Kirch, 2000). Then after a substantial hiatus, East Polynesia was rapidly settled over several hundred years following the arrival of people initially to the central tropical islands about 1000 years ago (Wilmshurst et al., 2011). This second settlement phase included the islands at the margins of East Polynesia: Hawaii, Rapa Nui and New Zealand.
1.2 Pacific dogs

Dogs were the first species to emerge from the domestication process at least 15,000 years ago, yet the location, timing and nature of this process is the subject of continuing debate (Larson et al., 2012). Archaeological and genetic research has identified several possible regions for domestication: Central Asia, East Asia, the Middle East and Europe, with time frames varying between archaeological and molecular dating methods (Pang et al., 2009, Savolainen et al., 2002, Shannon et al., 2015). Two independent domestication events in East Asia and Western Eurasia (Europe and the Middle East) have recently been proposed (Frantz et al., 2016) and that research, along with analyses of nuclear markers (Pilot et al., 2015, Sacks et al., 2013), highlight the likelihood of the replacement of some early dog Eurasian lineages by those with East Asian ancestry.

1.2.1 Archaeological evidence from the Pacific

Despite their long association with people, dogs appear relatively late in the archaeological record of the islands of the Pacific, and their dispersal trajectories appear to differ from that of people. The earliest evidence of dogs in the Pacific actually comes from Australia, where dog (dingo) remains have been excavated from an archaeological context from the Madura Cave in South Australia dated to 3,500 years ago (Milham and Thompson, 1976). The dingo has subsequently adapted to surviving independently of people, and feral dingo populations are found in many areas of rural Australia. Individual dingoes may, however, be taken in by indigenous communities where they become companions, protectors, and spiritual guardians, and can assist with hunting (Smith and Litchfield, 2009). In New Guinea, the New Guinea Singing dog (NGSD) is thought to be a descendant of another relatively early introduction, although no well dated or genetically confirmed archaeological NGSD remains have been reported. They are rarely observed in the wild and the current captive population is descended from only eight individuals (Koler-Matznick et al., 2007). In Australia, New Guinea, and the smaller islands of Near Oceania, dogs were introduced to areas where people had already been living for thousands of years.
The introduction of dogs into Remote Oceania is generally associated with the first wave of human migration to these previously uninhabited islands. Dogs are thought to be one of the three domesticated animals introduced during the Lapita expansion, along with pigs and chickens. Yet dog remains in Lapita archaeological sites are relatively rare, having only been recorded in limited numbers in the Bismarck Archipelago of New Guinea (Specht et al., 2014). Dog remains have not been reported in any of the key Lapita or later archaeological sites in Vanuatu or New Caledonia (Bedford, 2006, Sand, 2000). The subsequent, post-Lapita settlement of the islands further to the east also involved the transport of dogs, pigs and chickens, although they were not uniformly distributed across the region (Anderson, 2009). The archaeological evidence for dogs in Polynesia is variable. There is no evidence of dog remains in Lapita archaeological sites in Tonga and Samoa to the west, where they do not appear in the archaeological record till about 1,000 years ago (Smith, 2002). Dogs are also absent from Rapa Nui throughout the pre-colonial period. This contrasts with many of the island groups in East Polynesia, such as Hawaii and New Zealand, where large number of dog bones are found in early settlement sites.

In East Polynesia, when European explorers arrived in the 18th century, dogs were observed as being an important part of the social and economic fabric of daily life on some islands (Luomala, 1960b). In the ranked societies of Hawaii dogs were noted as the property of chiefs and were raised in large numbers for feasts. In New Zealand, dogs were the only domesticated animal to be successfully introduced by the Polynesian ancestors of the Maori. While living, they were kept as watch-dogs, hunting dogs and general companions and were sometimes also kept for their hair; on their death they could be used as food, sometimes for ceremonial occasions or sacrifices, their bones and teeth provided industrial materials, and their pelts were used to make dog skin cloaks (Davidson, 1987). The arrival of European dogs from the late 18th century onwards, which resulted in widespread interbreeding between the Pacific dogs and these newcomers, means that Pacific dogs are no longer identifiable as a distinct breed in the modern dog population in New Zealand. Archaeological remains are now the major source of information about these Pacific dogs and their morphology, and the only source of information about their genetic history. Developments in molecular
genetics such as aDNA techniques and NGS technology are now putting aspects of this genetic history within reach.

1.2.2 Molecular phylogenetic and phylogeographic studies

Phylogenetics is the study of the relationships between organisms, based on shared ancestry. Studies may be based on morphological, behavioural or molecular characteristics. In the 1980s the discovery of the Polymerase Chain Reaction (PCR) made it possible to directly sequence DNA. This enabled the use of molecular sequences as a unit of study, and gave rise to the discipline of molecular phylogenetics. Phylogeography draws on molecular phylogenetics to investigate the relationship between the spatial distribution of species and their genetic history (Avise, 1987). Developments in DNA sequencing technology continue to increase the amount of DNA data available for analysis, which has also resulted in a concomitant growth of computational biology and bioinformatics. Ancient DNA methods developed since the 1990s have also made it possible in some circumstances to extract and sequence DNA from archaeological, museum and paleontological specimens. Until recently most aDNA studies have been based on short sections of mitochondrial DNA (mtDNA); the reasons for this are discussed in further detail in Chapter 4. Molecular phylogenetic methods have a wide application across many disciplines, including systematics, biology, medicine, ecology and anthropology.

Molecular phylogenetics has contributed new methods to the study of human-animal relationships, particularly the processes of domestication (e.g. Zeder et al., 2006a), and subsequent movements of domesticated animals. Studies using mtDNA that have investigated the origins and dispersals of the dog throughout the Pacific indicate that the dingo, NGSD and Polynesian dogs are all descended from East Asian dogs. Using a portion of the control region of the mitochondrial genome (mitogenome) from modern dogs, Savolainen and colleagues (2004) demonstrated that all dingoes sampled belonged to a lineage known as the A29 haplotype. Although this is one of a number of dog mtDNA lineages that reached ISEA it was the only one successfully established in Australia. On this basis, Savolainen and colleagues (2004) suggest that the dingo population was either founded
by a very small number of individuals, or from a group of dogs that had passed through a series of genetic bottlenecks and hence had lost substantial genetic variation. They also reported, based on an analysis of a shorter portion of the control region, that 19 ancient Polynesian dogs (sampled from Hawaii, the Cook Islands and New Zealand) belonged to two short haplotypes, Arc1 and Arc2. These are both different from the A29 control region haplotype carried by dingoes. The short Arc1 haplotype is indistinguishable from a number of widespread control region haplotypes, while Arc2 has a much more restricted distribution, and appears to belong to the A75 lineage found only in modern Indonesian dogs.

Matisoo-Smith (2007) suggested on the basis of this molecular genetic evidence, and the additional analysis of mtDNA from a dog from the archaeological site of Taurama in New Guinea, that there were at least three introductions of dogs into the Pacific. Firstly, the arrival of the dingo to Australia and New Guinea, followed by two other lineages represented by the Taurama dog and Polynesian dogs. Genetic estimates suggest the dingo was introduced to Australia between 4 to 6 thousand years ago (Savolainen et al., 2004), pre-dating the Lapita migration. The remaining two lineages may have been introduced as part of the Lapita Cultural Complex (LCC), or in two phases with the Polynesian samples representing a later introduction coinciding with an increase in the number of dog remains appearing in the archaeological record after 2000 BP.

A later study by Oskarsson and colleagues (2012) investigated the origins and routes of the introductions of dingoes, NGSDs, and Polynesian dogs in further detail, using the same control region fragments as the previous study by Savolainen and colleagues (2004). The study used samples collected from modern dogs and dingoes, previously published modern dog and dingo sequences, and the 19 ancient Polynesian short sequences. Oskarsson and colleagues (2012) compared the frequency and distribution of modern haplotypes in MSEA and ISEA, including Taiwan and the Philippines, with those from the dingo and ancient Polynesian samples. Neither the ancient Polynesian short haplotypes nor the A29 haplotype carried by dingoes, were present in modern dogs from Taiwan and the Philippines,
suggesting that dogs were not introduced into the Pacific region from or via this northeastern route. The implications of these results are discussed in more detail in the following chapter.

1.3 Motivation for this study

Although molecular phylogenetic studies have gone some way in addressing questions about dog origins and dispersals in the Pacific, the recent, short timeframe for human settlement has confounded much of this research, along with impacts of serial migration on genetic diversity. This is in addition to the already reduced level of genetic diversity in dogs resulting from the domestication process. As a result, the patterns of genetic diversity observable in mtDNA sequences from modern and ancient dogs from Southeast Asia and the Pacific, using only a portion of the mitogenome, are broad in nature (e.g. Oskarsson et al., 2012). They do not provide sufficient levels of discrimination to address questions at smaller geographical and temporal scales which are more closely attuned to archaeological research questions, such as the process of animal introduction and establishment, the relationship between human colonisation, migration and animal movements, and the incorporation of animals into social networks and interactions in the Pacific region.

Similar broad patterns are seen in results of mtDNA analyses of pigs and chickens. Analyses of a control region mitochondrial fragment from modern pigs from across MSEA, ISEA, Near and Remote Oceania demonstrated that all the pigs sampled from Pacific islands belonged to what was termed a ‘Pacific clade’ (Larson et al., 2007). Similarly only two lineages based on a control region fragment have been observed in archaeological chicken specimens from across the region. The introduction of these lineages appears to be temporally distinct however, with Clade E associated with the Lapita migrations and Clade D with a later introduction (Storey et al., 2012).

NGS technology is now making it feasible to sequence the complete mitochondrial genome, rather than just a portion. As expected, studies of mitogenomes of domesticated species, such as cattle (Horsburgh et al., 2013) and sheep (Lancioni et al., 2013), have shown that the control region sequences provide only a partial picture of genetic diversity. NGS technology
has also contributed to the viability of sequencing of mitogenomes from archaeological specimens (Paijmans et al., 2013). This is still constrained however by issues particular to aDNA, such as DNA preservation, degradation and contamination (Pääbo et al., 2004).

Investigating human and animal movements in the past using archaeological and molecular phylogenetic approaches, however, requires further consideration than simply obtaining archaeological samples for aDNA analyses. As a discipline, the principles and concepts underlying molecular phylogenetics are fundamentally evolutionary in nature. Phylogenetic trees and networks are structured by underlying assumptions about models of molecular evolution and evolutionary processes (Lemey et al., 2009). The type of data produced by aDNA analyses and its appropriateness to address archaeological questions or be combined with archaeological data needs to be taken into account. This is becoming increasingly important as the vast amount of data produced by NGS becomes subject to more complex statistical and computational analyses (McCormack et al., 2013).

Phylogenies produced by bioinformatic computational methods are constructions or hypotheses of the ‘best’ scenario to account for observed genetic variation (Matisoo-Smith and Horsburgh, 2012). They are not ‘real’ in an objective sense, and this contrasts strongly with the way that the archaeological record is viewed, where the physical remains can be seen as direct evidence of people, animals, and material culture of past societies. This evidence is, of course, also subject to interpretation and will not reflect the whole picture of past events. The concept of population size provides an insight into this difference. For an archaeologist, a population may mean the group of people who occupied a particular place, and whose activities led to the formation of the archaeological record there. The size of the population is likely to be assumed to relate to the actual numbers of people that were present in the past. In molecular genetics, population size is an important parameter in constructing evolutionary history. Different methods have been proposed to estimate the ‘effective population size’, which is related to the number of reproducing individuals compared to an ideal population operating under particular conditions (Allaby, 2009), not the actual number of people present at a given time and place. An effective population size is likely to be much
smaller than the actual number of people present, especially when using mtDNA or Y-chromosome markers. The use of terminology in different ways can have implications for the interpretation of results across disciplines.

The commensal model uses phylogeographic methods to investigate human mobility and migration, using the animals that travelled with people as a proxy for tracing human movements (Matisoo-Smith, 2009). When the term ‘commensal’ is used in this approach, it is defined as referring to animals which are transported by people, and in the Pacific this includes rats, dogs, pigs and chickens (Storey et al., 2013). This differs from its use in the biological sciences where commensalism refers to a kind of relationship between species, where the relationship is beneficial to one species and neutral for the other (Allaby, 2009). The commensal approach may draw on DNA from modern populations, and also aDNA extracted from archaeological material. The approach was first developed using a short control region fragment of mtDNA from the Pacific rat (kiore) (Matisoo-Smith, 1994). Phylogenetic analyses of modern and ancient specimens in Near and Remote Oceania did possess sufficient variation to enable multiple lineages to be detected and used to explore associated human movements. This level of diversity contrasts with that of dogs, pigs and chickens, as observed from the control region analyses described above. It is notable that the Pacific rat has a much different ecology and relationship with people, and hence also dispersal trajectory, than dogs, pigs and chickens.

Understanding the relationship between dogs and people is therefore critical when interpreting aDNA data in relation to dog population structure and mobility. How closely do circumstances conform to the assumptions of the models used to generate measures of genetic diversity and relatedness for animal populations such as dogs, which may be strongly influenced or modified by their interactions with people? How might these compare with results for natural populations or commensals, such as rats? What does this mean for the use of phylogenetic methods for domesticated animals when addressing questions about animal and human dispersals in the Pacific?
To add another layer to the complexity, archaeological samples have passed through a series of cultural, taphonomic, and analytical filters, which will have shaped the assemblage in ways which are not always able to be determined. Issues that need to be considered include assessing the representativeness of an archaeological sample in relation to the population from which it came. This issue is addressed in more detail in Chapter 4.

1.4 Research aim and questions

The purpose of this research is to investigate the introduction and dispersal of dogs throughout the Pacific. My approach uses dog remains from archaeological sites across the region to produce new aDNA data generated using NGS technology. The aDNA data is then used to investigate archaeological questions about the interactions between people and dogs, particularly in relation to current archaeological models for human and animal movements across the region. The utility of complete mitogenomes of ancient dog DNA to address such questions is also explored.

Three research questions are addressed in order to meet the aim of the thesis:

1. Do complete mitogenomes from archaeological dog specimens provide sufficient observable genetic variation to investigate genetic relationships and ancestry at an archaeological site, island or archipelago level?

2. How do the outcomes of phylogenetic analyses of dog genetic diversity and relatedness compare to archaeological models for colonisation era human mobility and interactions in Oceania and New Zealand?

3. Are phylogenetic patterns of dog diversity and relatedness based on mitogenomes informative about dog-human mobility and interactions in the Pacific?

This study focuses on two episodes in the history of dog dispersal in the region, firstly the introduction of dogs into Oceania, generally thought to be associated with the Lapita expansion, and secondly, the introduction and dispersal of dogs in New Zealand. Both these episodes are associated with a body of archaeological research which includes well developed
models for human migrations, social networks and interactions, and also the movement of raw materials within these networks. These two episodes also provide different contexts for archaeological material, and enable an evaluation of aDNA data in two quite different spatial, temporal and environmental contexts.

The introduction of dogs to Oceania in association with the arrival of the Lapita peoples took place around 3,500 years ago (Kirch, 2000). The Lapita settlement of New Guinea, the Bismarck and Solomon island arcs, and Remote Oceania is marked by a rapid dispersal across a large number of islands spread across over 6,000 kilometres. People had been living on some of these islands for tens of thousands of years. The introduction of dogs to New Zealand represents the arrival of dogs and people to a previously uninhabited area. This introduction occurred at the end of this last major pre-industrial human migration, only about 700 years ago (Jacomb et al., 2014). This episode enables exploration of the utility of aDNA to investigate recent population movements and human-animal interactions at the terminus of a major set of migrations.

The distribution of populations across geographical space has formed the basis of phylogenetic and phylogeographic studies of wild populations, commensal species such as rats, and domesticated species, including dogs. Each of these groups of animals have different genetic, behavioural and ecological characteristics, that influence their relationship with people. Populations in the wild have little to no interactions with people, commensal species may be transported by people but once introduced have little or no human intervention in their reproduction and ecology apart from inadvertently providing a domestic niche, while people may play an active and ongoing role in the ecology of domestic species such as dogs.

A more detailed understanding of the genetic variation and relatedness of dogs at a particular place and time, for example, colonisation era sites in New Zealand, may provide insights into aspects of dog mobility and human-dog interactions, beyond reconstructions of phylogenies over large spatial and temporal scales. Mobility is associated with the movements of animals within, around and between settlements, and this may, or may not, involve deliberate human action. These kinds of interactions are often fundamental to archaeological interpretations
of the place of animals in human societies, including domestication, animal management practices, and the functional roles of dogs. Conversely, dogs are animate and aspects of dog behaviour may also be implicated in phylogenetic and phylogeographic patterns.

1.5 Significance of this research

This research presents new data generated from aDNA analyses of complete mitogenomes of dogs from archaeological samples from ISEA and the Pacific. These include the first complete mitogenomes for Pacific dogs, and the first archaeological mitogenomes from ISEA. These data include dog sequences from the furthermost southern extent of dog distribution across the Pacific ocean.

The data presented here are considered with reference to archaeological models for animal and human introductions and dispersals in the Pacific region. This research demonstrates the value of complete mitogenomes for studies where control region analyses are insufficient to discriminate between different lineages in a meaningful manner. The limitations of this phylogenetic approach are also discussed.

The interpretation of this molecular data set in relation to archaeological problems is carefully considered. The different disciplinary underpinnings of molecular phylogenetics and archaeology are addressed and the appropriateness of the mitogenomic dataset is demonstrated for the dogs in this Pacific context. The influence of cultural and social variables may be also applicable to other taxa in different regional contexts.

1.6 Organisation of this thesis

This thesis is organised as follows. After this introduction, Chapter 2 explores the current views about the origins of dogs in the Pacific in further detail, based on archaeological evidence for dog domestication, their introduction and distribution across the Pacific, and the data from phylogenies constructed from short sections of the mitogenome. Chapter 3 then examines ecological information about free-ranging dogs in contemporary settings, with attention given to the ways in which dogs may behave which may influence genetic structure. This includes behaviours that occur independently of people. Interactions between
people and dogs documented in villages in the Asia-Pacific region from contemporary and historical observations are also explored.

Chapter 4 then turns towards an examination of the archaeological models for the migrations associated with the Lapita expansion, and for the later colonisation of New Zealand, and the expectations of these models for genetic data. Phylogenetic and phylogeographic approaches to investigating relationships between populations are considered next, including the underlying parameters and assumptions of different methods. The use of aDNA obtained from archaeological samples is also addressed.

The results of the aDNA extraction, sequencing and phylogenetic analysis of dog remains from Wairau Bar, a colonisation era archaeological site in New Zealand, are presented in Chapter 5. This chapter also provides an account of the methods for archaeological sample selection, and aDNA extraction, library preparation, sequencing and phylogenetic analyses adopted for this thesis. This small preliminary study, undertaken to establish the appropriateness of complete mitogenome sequencing for phylogenetic analysis in New Zealand and the Pacific, prepares the way for the two larger studies that follow in Chapters 6 and 7.

Chapter 6 presents the results of phylogenetic analyses of dog samples from Wairau Bar alongside with samples from four additional colonisation era archaeological sites in New Zealand and a similar aged archaeological site from the Southern Cook Islands, following the method set out in Chapter 5. The results are then compared with expectations derived from archaeological models of human mobility and interaction networks during the colonisation of New Zealand. Chapter 7 likewise presents the results of phylogenetic analyses of dog samples from sites across the Pacific, and considers these results in relation to current models for Lapita colonisation and interactions.

Chapter 8 moves from the data to a consideration of the ways in which phylogenetic studies may intersect with archaeological frameworks that attempt to explain how people and dogs move and interact during the colonisation process. The thesis concludes with Chapter 9,
which summarises the outcomes of this thesis, considers its limitations and implications for future work.

### 1.7 Note on published works

This thesis has been produced in accordance with the University of Otago policy ‘Guidelines for the Inclusion of Material from a Research Candidate’s Publications in their Thesis’ (5 June 2014), which encourages publishing during candidature for a doctorate. This thesis is a hybrid thesis, where published material and material to be submitted for publication has been included in the thesis (Table 1.1). This work however has been modified where appropriate to enable the thesis to stand alone as a coherent document.

**Table 1.1: Published material and material submitted or prepared for publication included in this thesis.**


As expected by the University, I am the first author on all publications. Co-authors include my supervisors who assisted with research design, the interpretation of results and editing, colleagues who carried out bioinformatic and laboratory analyses, and archaeologists who provided samples for sequencing and details about archaeological site context and dates. The contributions of the co-authors for each publication or manuscript are detailed in Table 1.2.
Table 1.2: Contributions of co-authors to published works and manuscripts to be submitted for publication. Bracketed numbers refer to works listed in Table 1 above.

<table>
<thead>
<tr>
<th></th>
<th>PlöS one article (1)</th>
<th>Routledge book chapter (2)</th>
<th>Submitted to JPA (3)</th>
<th>Manuscript (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept and design</td>
<td>KG, CJ, RW, LMS</td>
<td>KG, RW, LMS</td>
<td>KG, RW, LMS</td>
<td>KG, RW, LMS</td>
</tr>
<tr>
<td>Laboratory work</td>
<td>KG, KAH</td>
<td>n/a</td>
<td>KG</td>
<td>KGaDNA, KMcd dingoes</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>KG, SP, JB</td>
<td>n/a</td>
<td>KG, JB</td>
<td>KG, JB</td>
</tr>
<tr>
<td>Archaeological context</td>
<td>KG, CJ, RW</td>
<td>n/a</td>
<td>KG, RW</td>
<td>KG, DA, MA, BD, MG, CH, IM SO, CT, FL, RW</td>
</tr>
<tr>
<td>Linguistics</td>
<td></td>
<td>R W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpretation of results</td>
<td>KG, JB, CJ, RW, LMS</td>
<td>KG, JB,RW, LMS</td>
<td>KG, JB, RW, LMS</td>
<td></td>
</tr>
<tr>
<td>Drafting of manuscript</td>
<td>KG</td>
<td>KG</td>
<td>KG</td>
<td>KG</td>
</tr>
<tr>
<td>Critical revision</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
</tbody>
</table>

Each chapter that includes published or work to be submitted for publication also begins with a short preface that outlines how the work has been incorporated in the chapter.
Chapter 2
Dog origins and dispersal across Southeast Asia and the Pacific

Preface

This chapter contains published work from: Greig, K., Walter, R. & Matisoo-Smith, E. (2016). Dogs and people in Southeast Asia and the Pacific. In: Oxenham, M. & Buckley, H. (eds.) *The Routledge Handbook of Bioarchaeology in Southeast Asia and the Pacific Islands*. Oxon, Routledge. pp. 462-482. The book chapter is given in its entirety, with the exception of: the full introduction; the concluding comments (aspects of which are incorporated into Chapter 9 of this thesis); and a section titled 'Living with dogs in Southeast Asia and the Pacific’ which is included in the following chapter.

2.1 Introduction

This chapter examines the current state of knowledge about dog origins, dispersal and distribution throughout Southeast Asia and into the Pacific, and explores the relevance of these for understanding the place of dogs in the settlement history of the Asia-Pacific region. The current molecular genetic models for dog domestication, and subsequent dispersal, have implications for understanding the mechanisms and timing of dog movements in the Asia region and out into the Pacific. Unlike the expansion of modern forms of domesticated animals (such as dogs, sheep, cows), the prehistoric dispersal of dogs into the Pacific is linked to the domestication process itself. We therefore begin with an overview of the domestication process and how it relates to dogs, before examining the current molecular genetic models with specific reference to the Asia/Pacific region. We then address dog distribution across the region based upon archaeological and linguistic data, and consider how it relates to the genetic models of dog dispersal based on modern and aDNA studies.

2.2 Domestication

Domestication has been described as the most profound transformation that has occurred in human-animal relations (Russell, 2002: 285). Domestication has proved difficult to define
but involves behavioural changes in both animals and people, and physiological changes in animals (Zeder et al., 2006b). Domestication marks a major turning point in human history; accordingly animal (and plant) domestication has been of interest to a wide range of disciplines, including anthropology, archaeology, zoology, biology and genetics. Much of the conceptual framework for considering the domestication of animals and people’s relationships with domesticated animals has arisen around the transition from hunting and gathering to sedentism that began in the early Holocene in continental Eurasia. Research over the last few decades, however, has demonstrated that rather than arising from a single point of origin, domestication actually occurred in multiple locations around the world, and that the processes driving domestication may have varied through time and across space (Bellwood and Diamond, 2005, Zeder, 2006). Domestication is a much more complex process than initially envisioned and is in no way linear and unidirectional—from a state of wild to domesticated—nor were the processes the same for each animal species or in every culture.

Archaeological research into the process of animal domestication has generally focused on identifying the presence of morphological markers of the physiological changes associated with domestication. Morphological changes that are genetically driven, passed from one generation to another and seen in most domestic animals, include a decrease in body size, changes to facial structure and tooth size and placement, and for ruminants, changes in horn size and structure (Zeder et al., 2006b). Other morphological changes can arise in response to changes in the environmental conditions experienced by individual animals. For example, the development of pathologies from methods of controlling animal movements, such as tethering or penning, may leave an imprint on bones, although marks arising from domestication can be difficult to discriminate from those that occur in wild populations (Zeder et al., 2006b: 174). Joint pathologies possibly arising from the use of animals for traction have also been reported in water buffalo remains from the site of Ban Chiang in Thailand (Higham et al., 1981). Non-morphological markers that may provide evidence for domestication include changes in demographic profiles, changes in the abundance or movements of animal populations, the presence of animals outside their natural distribution
zones, and other indirect archaeological evidence such as the presence of structures for animal control (Zeder et al., 2006b). Dietary differences detectable by stable isotope analyses between domesticated and wild animals may also indicate animals that are more closely associated with human groups. For example, isotopic analysis of the diet of people and animals in sites from China dating from 2,000 to 4,000 BCE distinguished between wild boar, and pigs in the early stages of domestication (Hu et al., 2008).

Molecular genetic studies of domestication have focused on differences detectable in the genomes of domesticated animals, e.g. observing overall genetic differences between domesticates and their wild ancestors, and from these, estimating the time frame for the divergence (Zeder et al., 2006b). Initial studies utilised variation in the control region of the mitogenome in order to address issues of divergence and dispersal. Advances in genetic techniques and analysis, particularly in the field of aDNA, are enabling a greater range of genetic data, including nuclear markers, to be targeted. Recent studies have investigated variation in functional genes such as those controlling coat colour, and other characteristics not identifiable in archaeological remains (reviewed in Linderholm and Larson, 2013).

### 2.3 Dog domestication

The complexity of the process of domestication is perhaps most apparent in the case of the dog. Both archaeological and genetic evidence indicate that dog domestication occurred much earlier than that of other animals. While archaeological and morphological evidence for the domestication of cattle, sheep, goat and pig all appear in the early Holocene, morphological changes thought to be associated with putative aspects of dog domestication may be observable in archaeological remains from Central Europe during the Pleistocene, as early as 31,000 to 34,000 BCE (Germonpré et al., 2009, Ovodov et al., 2011), although this is debated (e.g. Crockford and Kuzmin, 2012). By the end of the Pleistocene, between 10,000 and 12,000 BCE, there is clear evidence for domesticated dogs present in sites across Europe and the Middle East, and by 8,000 BCE in East Asia and North America (reviewed in Morey, 2010). Differences in the timing of the appearance of dogs across the world raise questions about the original relationships between dogs and humans, and our ability to recognise these
in the archaeological record. While the circumstances surrounding dog domestication remain unclear, current research suggests a variety of selective, co-evolutionary and environmental forces were involved (Miklósi, 2007). Co-habitation would have resulted in selection for behavioural characteristics associated with tolerance for close proximity to people, and also perhaps for the relaxation of group hierarchical and territorial behaviours.

Recent developments in whole genome sequencing are now providing different kinds of information about dog ancestry and the selection of physiological and behavioural traits associated with domestication. Whole genome analyses of wolves from the three hypothesised centres of dog domestication (Asia, Europe and the Middle East), and the golden jackal, combined with basal dog linages (Basenji and Dingo) indicate a complex history of dog domestication (Freedman et al., 2014). Freedman and colleagues (2014) argue this occurred at a single place of origin between 9,000 to 14,000 BCE, before the transition to agriculture. Building on an earlier study showing changes in dog genes related to digestion and metabolism (Axelsson et al., 2013), they found the amylase gene, related to starch digestion, was present in varying copy numbers in both wolves and dogs. They suggest that the increase in copy numbers in modern dogs is a recent development post-dating domestication and that it related to eating a more carbohydrate-rich diet as a result of the agricultural transition in Asia, Europe and the Middle East. In a study of genes that are expressed in brain function, Li and colleagues (2013) found significant differences between the genes of wolves and Chinese native dogs, which they attribute to the rapid evolution of dog-specific behaviours during domestication. These studies indicate the potential of genetic research to provide further insights into the changes brought about by the domestication process, and the importance of understanding the wider cultural context in which these processes took place.

In addition to questions about the processes and driving factors behind dog domestication, the location, timing and subsequent dispersal of dogs remains unclear. Dogs are thought to be a descendant of the grey wolf or a recent shared common ancestor (Freedman et al., 2014). The current distribution of grey wolves spans North America, Eurasia and northern Africa.
Given this distribution of wolves, the obvious question is: did domestication happen once and spread, or were there multiple domestication events? There are currently two contrasting models.

### 2.3.1 Single East Asian origins

Analyses of mitochondrial and Y-chromosome DNA obtained from modern dogs from around the world indicate a single East Asian origin for dogs (Savolainen et al., 2002, Ding et al., 2012). Mitochondrial DNA analysis suggests a particular area in the extreme southeastern part of Asia, from south of the Yangtze River in China to Southeast Asia (Pang et al., 2009). This is known as the ASY (Asia South of the Yangtze) Model, and genetic evidence suggests domestication could have occurred here as early as 14,000 BCE (Pang et al., 2009). Archaeological evidence for morphologically domestic dogs in China, however, currently dates no earlier than around 8,000 BCE (Liu and Chen, 2012).

### 2.3.2 Multiple origins

The argument for multiple origins is based on archaeological and genetic evidence that suggests additional locations for domestication in the Middle East and Eurasia (vonHoldt et al., 2010, Ovodov et al., 2011, Thalmann et al., 2013, Frantz et al., 2016). Canid remains with some dog-like features have been identified in several sites in Europe, however, the classification of these canids as early dogs on the basis of morphology is debated (Crockford and Kuzmin, 2012, Boudadi-Maligne and Escarguel, 2014). It has been suggested that some of these very early European dog-like canids were ‘proto’ dogs from lineages that did not survive climatic, social and settlement pattern changes at the end of the Pleistocene (Ovodov et al., 2011, Thalmann et al., 2013). By around 10,000 to 12,000 BCE there is compelling archaeological evidence for truly domesticated dogs across Europe and the Middle East and, from this time onwards, dogs are found in deliberate burials, sometimes in association with people (Morey, 2006).
2.4 Dog dispersal

These conflicting domestication models do have slightly different implications for the arrival of dogs into MSEA and their subsequent dispersal across ISEA and the Pacific. The ASY Model sees the origin and expansion of dogs as part of a regional evolutionary phenomenon, whereas the Multiple Origins Model sees it as a global process. This latter model could include a local domestication event in East Asia but would also allow for the movement into and out of the region of multiple dog lineages at different times. Whatever the actual history of domestication, dog populations were present in the archaeological record of East Asia by 8,000 BCE and lineages ultimately deriving from East Asia are found in many modern dog populations.

Sacks and colleagues (2013) have proposed a ‘Neolithic replacement’ hypothesis in an attempt to reconcile the genetic evidence for the dominance of East Asian lineages in dogs today with the lack of archaeological evidence for dogs in the region prior to around 8,000 BCE. Using Y-chromosome markers in modern village dogs from ISEA, dingoes, NGSDs and a range of modern dogs, the hypothesis suggests that although dogs were not originally domesticated in East Asia, a major expansion of dogs from the region of the Yangtze and Yellow River basins arising 6,000 to 3,000 BCE resulted in the partial replacement of the earliest dogs possessing western continental lineages. This replacement coincides with a major shift in food production in East Asia around 6,000 BCE and the rapid spread of food producing communities across the region, including the expansion of Austronesian language speakers through Southeast Asia, ISEA, and eventually into the Pacific.

While domesticated or semi-domesticated dog populations from across northern Eurasia could have regularly back-crossed with wild wolf populations, the movement of people out of East Asia and into MSEA would have moved dogs beyond the natural distribution of wolf populations and may have been an important point in the process of domestication. From this point on dog populations would be reproductively isolated from wolves and dog behavioural and morphological characteristics would have become fixed (Figure 2-1). Southeast Asia is therefore a significant region for understanding the history of dog
domestication, the relationship between dogs and people, and the subsequent dispersal of dogs around the world, and particularly into the Pacific.

Figure 2-1: Map showing the natural distribution of wolves, the proposed centre of dog domestication based on DNA analysis south of the Yangtze River and selected archaeological sites with dog remains in Mainland and Island Southeast Asia mentioned in the text (wolf distribution based on Larson et. al 2012).

The expansion of dog populations east of MSEA involved an increasingly difficult exercise, technologically and ecologically. The further one moves to the east out into the islands of the Pacific the smaller and more dispersed the landmasses become and there is an accompanying drop in biodiversity (Green, 1991a). While some natural expansion of dog populations is possible in Mainland and parts of ISEA and Australia, further east all dog dispersal would
have required deliberate human intervention. To understand the expansion of dogs east beyond MSEA it is important to understand as fully as possible what the nature of the dog-human relationship is likely to have been. One way of doing this is to look at ecological and ethnographic information about interactions between dogs and people in the region—this will be considered in more detail in next chapter—but first we will continue to track dog dispersal from East Asia across ISEA and out into the Pacific.

2.5 Distribution of dogs

Drawing on archaeological and genetic evidence we can trace the distribution of dogs across the Asia-Pacific region and, to some extent, the timing of that distribution.

2.5.1 East Asian, Mainland and Island Southeast Asian dogs

Archaeological evidence

Dog remains are documented in archaeological sites relating to hunting and gathering communities in China from the end of the Pleistocene and are present in mixed economy and agricultural sites from the earliest periods of the Neolithic onwards (Chi and Hung, 2012, Underhill, 1997). The earliest record of dogs in China is from the northern site of Nanzhuangtou which is associated with millet cultivation, dating to 8,000 BCE (Table 2.1). Other early Neolithic sites dated between 7,000 and 6,000 BCE, where food obtained by hunting was an important component of the diet, also contain deliberate dog burials (Liu and Chen, 2012). Liu and Chen (2012) suggest this is indicative of a close relationship between dogs and people, possibly related to their contribution to hunting. Dog remains are present in Neolithic sites mixed with domestic refuse, which indicates that dogs were also being eaten, and in some places dogs were also used as sacrificial offerings (Liu and Chen, 2012). In southern coastal China, Chi & Hung (2012) document the presence of dogs in sites with mixed hunting, gathering and early agricultural economies from 5,000 to 3,000 BCE. It is notable that at this time other domestic animal remains were not present. For example, the sites of the Dingshishan indigenous forager culture, spanning southeastern coastal China and northern Vietnam, have dogs around 5,000 BCE, but no domestic pig or water buffalo (Lu cited by Liu and Chen, 2012).
Table 2.1: Archaeological evidence for dogs in East Asia, MSEA and SEA mentioned in the text and shown in Figure 2-1 (reported in English literature sources)

<table>
<thead>
<tr>
<th>Site or context</th>
<th>Location</th>
<th>Date</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanzhuangtou</td>
<td>Northern China</td>
<td>8,000 BCE</td>
<td>Chi &amp; Hung, 2012; Underhill, 1997</td>
</tr>
<tr>
<td>Dingsishan culture</td>
<td>Coastal southeastern China and northern Vietnam</td>
<td>5,000 BCE</td>
<td>Chi &amp; Hung, 2012</td>
</tr>
<tr>
<td>An Son</td>
<td>Vietnam</td>
<td>2,485–1,055 BCE</td>
<td>Bellwood et al., 2011</td>
</tr>
<tr>
<td>Ban Chiang</td>
<td>Northeastern Thailand</td>
<td>2,016–1,538 BCE</td>
<td>Higham et al., 2011</td>
</tr>
<tr>
<td>Khok Phanom Di</td>
<td>Central Thailand</td>
<td>1,850 BCE</td>
<td>Higham, 2014</td>
</tr>
<tr>
<td>Nong Nor</td>
<td>Central Thailand</td>
<td>1,950 BCE</td>
<td>Higham and Thosarat, 1998</td>
</tr>
<tr>
<td>Matja Kuru 2</td>
<td>East Timor, Indonesia</td>
<td>1,050 BCE</td>
<td>Gonzalez et al., 2013</td>
</tr>
<tr>
<td>Aru Islands</td>
<td>Indonesia</td>
<td>mid-Holocene</td>
<td></td>
</tr>
<tr>
<td>Nagsabaran</td>
<td>Batanes Islands, Philippines</td>
<td>500 BCE</td>
<td>Amano et al., 2013, Piper et al., 2014</td>
</tr>
<tr>
<td>Uattamdi, Gua Siti Nafisah</td>
<td>Northern Maluku Islands, Indonesia</td>
<td>150 BCE</td>
<td>Flannery et al., 1995</td>
</tr>
</tbody>
</table>

Some of the earliest reported evidence for dogs in MSEA comes from the site of An Son, a Neolithic settlement within the Mekong River system in Vietnam, which dates from 2,458 to 1,055 cal BCE (Bellwood et al., 2011). An Son is one of a group of food-producing Neolithic settlements and cemetery sites in the region which appear in the archaeological record around or soon after 2,000 BCE. The subsistence strategies at these sites were varied and involved differing combinations of foraging and food production (Piper et al., 2014b). The high proportion of dog remains recovered from An Son middens indicates the use of dogs as a food item, particularly as a number of dog bones have evidence of butchery marks (Piper et al., 2014b). Dog remains are also documented from Man Bac in northern Vietnam where pigs were the dominant domesticate, dated to between 1,745–1,538 and 2,016–1,775 cal BCE (95.4%) (Sawada et al., 2011).

Dog remains first appear in archaeological sites in Thailand slightly later. In the northeast, at Ban Chiang dogs are present alongside other domesticated animals and rice cultivation
from the Neolithic occupation levels dated to 1608 and 1442 cal. BCE (Higham et al., 2011). There is evidence consistent with their use as food, such as burning and butchery marks, and a dog crania was also found interred with a human burial (Higham et al., 1980). At the site of Khok Phanom Di, dogs appear part way through the sequence at c. 1,850 BCE, before other domesticated animals, at around the same time as the adoption of rice cultivation (Higham, 2014). Nearby, the earlier hunter gatherer occupation at the site of Nong Nor dated to the late third millenium BCE (Higham et al., 2011), is notable for its lack of dogs (Thosarat, 2001).

The appearance of dogs in Mainland Southeast Asia (MSEA) archaeological sites is consistent with the scenario of a movement of people into the region associated with the development and spread of food producing economies (Bellwood, 2011, Sacks et al., 2013). The region south of the Yangtze River is a critical area for understanding the domestication and dispersal history of the dog, and further archaeological research is required to explore the complexity of this process and the ways in which dogs were maintained and incorporated into communities involved with foraging and food production. The shifting emphasis from hunting to animal husbandry may have involved significant changes to people’s relationships with dogs.

In ISEA, the temporal concept of a transition to a Neolithic food-producing economy is less useful. Increasingly, researchers are highlighting the complexity of the interactions and movements of people, ideas and plants and animals across the islands of Southeast Asia, which do not fit into a simple model of a transition from foraging to farming economies (Barker and Richards, 2013, Spriggs, 2011). Dogs continued to disperse across ISEA in the mid to late Holocene, however, the mechanisms by which this happened are not clear. Dog bones are present in middens in the northern Philippines by around 500 BCE, in the Savidug Dune site in the Batanes Islands (Piper et al., 2014a) and at Nagsabaran in northern Luzon (Amano et al., 2013). Cut marks on the bones suggest butchery and skinning, and the use of dogs as a food item. By the mid-late Holocene, dogs are present on Timor and the Aru Islands (O’Connor and Aplin, 2007, Gonzalez et al., 2013). A tooth recovered from the
Uattamdi site on Kayoa Island and some cranial fragments from Gua Siti Nafisah indicate that dogs had reached the Maluku Islands by 150 CE (Flannery et al., 1995).

**Genetic evidence for dog dispersal**

Analysis of mitochondrial DNA from dingoes, modern dogs from South China, MSEA, ISEA and archaeological Polynesian dogs, identified distinct geographical patterning, indicating that the Polynesian dogs (and dingoes and NGSDs) trace their ancestry back to South China through MSEA and Indonesia (Oskarsson et al., 2012). At present there is no mtDNA evidence for a connection with Taiwan or the Philippines. This route is also consistent with results from molecular studies of pig dispersal in the region (Larson et al., 2007) and with the linguistic evidence for dog introduction into Oceania (see below).

A recent study of modern Y-chromosome markers from dogs in Southeast Asia did, however, suggest a possible link between Taiwanese dogs and ISEA dogs (Sacks et al., 2013). The study, however, was based on samples from modern dogs, which may not be representative of past dog populations. Additional studies incorporating aDNA are necessary to investigate any possible links between Taiwan and ISEA dogs through time.

Very little aDNA research has been undertaken on East and Southeast Asian dogs, despite the importance of these regions for understanding dog domestication and dispersal. Short mitochondrial DNA sequences have been obtained from samples from Ban Chiang (Matisoo-Smith, 2007), and from Timor (dated to 1,050 BCE) (Gonzalez et al., 2013). These sequences are consistent with the mitochondrial lineages of modern village dogs in the region (Oskarsson et al., 2012).

### 2.5.2 Australia, New Guinea and the Torres Strait

The earliest archaeological evidence for the dog (dingo) in Australia is from the Madura Cave in South Australia dated to 3,500 years ago (Milham and Thompson, 1976). This predates archaeological evidence for dogs in New Guinea and raises the possibility of an earlier dog dispersal prior to the Neolithic expansion of food producing groups into ISEA.

At present the earliest dates for dogs in the islands of the Torres Strait, which lies between
Australia and New Guinea, are around 150 CE (McNiven, 2008). Subsequent introductions are also indicated in historical accounts, which suggest that maintaining dingo populations on the islands may have required periodic restocking. The dates for dogs in the Torres Strait are consistent with the dispersal of dogs to the area via coastal New Guinea, yet the question of how dingoes were introduced to Australia several thousand years earlier remains unanswered.

In New Guinea the earliest recorded dog bone comes from late Lapita sites in Caution Bay, on the south coast of Papua New Guinea, near Port Moresby, in levels dating to 500–300 BCE (McNiven et al., 2012). The majority of other coastal sites in New Guinea do not contain dog remains until around 50 CE or later (Bulmer, 2001). Bulmer (2001) has argued for the early, pre-Lapita introduction of dogs to the New Guinea Highlands based on dog bones recovered from the sites of Yuku and Kiowa, but the dates for these sites, and indeed the identification of the dog and pig remains, are problematic. At this point, dogs are not securely dated in the Highlands before around 950 CE (Sutton et al., 2009).

NGSDs possess mtDNA that is either identical or very closely related (one mutation away) to those haplotypes identified in dingoes (Savolainen et al., 2004). That, and a similarly low, but shared, Y-chromosome diversity suggests a very restricted introduction of the first dingoes into Australia, probably from New Guinea (Ardalan et al., 2012). In addition, another genetic marker, the dog leucocyte antigen, is similar across both populations (Runstadler et al., 2006). Behavioural and morphological characteristics also suggest a close relationship with dingoes, but no ancient NGSD skeletal remains have been recovered. As a result it is difficult to confirm whether or not there were two dog introductions to the island of New Guinea. A study of a short control region fragment from archaeological dog bones recovered from a puppy buried under a house floor from the coastal site of Taurama in New Guinea, dating to 550 - 50 BCE, showed that this dog does not, however, carry the dingo/NGSD mtDNA markers, but is similar to lineages found in archaeological specimens from East Polynesia (Matisoo-Smith, 2007).
2.5.3 Near Oceania

Close scrutiny of the archaeological record of dog remains in early Lapita sites (i.e. c. 1,400–550 BCE) in Near Oceania indicate that dog does not appear to have been an important part of the Lapita economy or culture. While one drilled dog tooth has been identified from the site of Kamgot, in the Anir group in the Bismarck Archipelago, no other confirmed early Lapita associated dog remains have been identified (Table 10.1 in Matisoo-Smith, 2007). Interestingly, the date of 500 BCE for the late Lapita Caution Bay sites in New Guinea (McNiven et al., 2012) is also the time when evidence of pig bone appears in New Guinea sites such as Watom, New Britain, and may indicate a date of domestic mammal introductions to New Guinea and the Bismarck Archipelago. Dog bone is present in sites in Near Oceania from 2,000 years ago (see Specht et al., 2014 for a review) though not in large quantities, which suggests that much lower densities of dogs may have been present in these village communities.

2.5.4 Remote Oceania

The distribution of archaeological dog bone in Remote Oceania is most intriguing (Table 2.2, Figure 2-2). No prehistoric dog bone has ever been reported from New Caledonia or Vanuatu. While early Lapita sites in Vanuatu do contain pig and chicken bones (Bedford, 2006), neither of these animals have been recorded in New Caledonia (Sand, 2000). In Fiji,
the earliest secure dates for dogs are around 950 CE (Anderson, 2009), with the identification of early dogs in Lapita contexts at the site of Naigani now seen as questionable (Clark and Anderson, 2009). Dog is also absent from Lapita sites in Tonga and Samoa (Smith, 2002). Therefore, it appears that dogs were not introduced to Remote Oceania with initial Lapita expansion in the region.

Figure 2-2: Map showing early distribution of dogs in the Pacific, from archaeological contexts and reported by eighteenth and nineteenth century European voyages in the region.
Table 2.2: Archaeological evidence for dogs in dogs in Remote Oceania, as discussed in the text and shown in Figure 2-2.

<table>
<thead>
<tr>
<th>Island/group</th>
<th>Present prior to European contact</th>
<th>Present at European contact</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Caledonia</td>
<td>No</td>
<td>No</td>
<td>Sand, 2000; Luomala 1960</td>
</tr>
<tr>
<td>Vanuatu</td>
<td>No</td>
<td>No</td>
<td>Bedford, 2006; Luomala, 1960</td>
</tr>
<tr>
<td>Fiji</td>
<td>Yes (950 CE)</td>
<td>?</td>
<td>Anderson, 2009</td>
</tr>
<tr>
<td>Tonga</td>
<td>Yes (post Lapita)</td>
<td>No</td>
<td>Smith, 2002; Luomala, 1960</td>
</tr>
<tr>
<td>Samoa</td>
<td>Yes (post Lapita)</td>
<td>No?</td>
<td>Smith, 2002; Luomala, 1960</td>
</tr>
<tr>
<td>Tokelau</td>
<td>Yes</td>
<td>No</td>
<td>Ono and Addison, 2013</td>
</tr>
<tr>
<td>Niue</td>
<td>No</td>
<td>No</td>
<td>Walter and Anderson, 2002</td>
</tr>
<tr>
<td>Cook Islands</td>
<td>Yes</td>
<td>?</td>
<td>Walter, 1998</td>
</tr>
<tr>
<td>Caroline Islands, Micronesia</td>
<td>Yes (5–450 CE)</td>
<td>?</td>
<td>Rainbird, 2004</td>
</tr>
<tr>
<td>Fais, Micronesia</td>
<td>Yes (400 CE)</td>
<td>?</td>
<td>Intoh, 2008</td>
</tr>
<tr>
<td>Marshall Islands</td>
<td>Yes</td>
<td>?</td>
<td>Weisler et al., 2012</td>
</tr>
<tr>
<td>Nukuoro</td>
<td>Yes</td>
<td>No</td>
<td>Davidson, 1971</td>
</tr>
<tr>
<td>Kapingamairangi</td>
<td>No</td>
<td>No</td>
<td>Leach and Ward, 1981</td>
</tr>
<tr>
<td>Hawaii</td>
<td>Yes (from initial settlement)</td>
<td>Yes</td>
<td>Pearson et al., 1971, Luomala, 1960</td>
</tr>
<tr>
<td>Marquesas</td>
<td>Yes (from initial settlement)</td>
<td>No</td>
<td>Rollet, 1998; Suggs, 1961; Sinoto, 1966; Luomala, 1960</td>
</tr>
<tr>
<td>Society Islands</td>
<td>Yes (from initial settlement)</td>
<td>Yes</td>
<td>McCoy and Sinoto, 1979; Luomala, 1960</td>
</tr>
<tr>
<td>Mangareva</td>
<td>Yes</td>
<td>No</td>
<td>Green and Weisler, 2004</td>
</tr>
<tr>
<td>Rapa Nui</td>
<td>No</td>
<td>No</td>
<td>Steadman et al., 1994</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Yes (from initial settlement)</td>
<td>Yes</td>
<td>Brooks et al., 2009; Higham et al., 1999</td>
</tr>
</tbody>
</table>

2.5.5 Central and East Polynesia

In contrast with the Lapita sites to the west, dogs were clearly a major component of the colonising voyages into Central and East Polynesia, although the prehistoric distribution of dogs remains patchy. Dog is present in the earliest sites in the Cook Islands but is not found at all in sites in Rapa Nui (Steadman et al., 1994) or Niue (Walter and Anderson, 2002).
some islands too, dogs were present in early periods, but absent by the time of European contact. Examples of dogs becoming absent from the archaeological sequence after having been present previously include Tokelau (Ono and Addison, 2013) and Mangareva (Green and Weisler, 2004). Dog bone appears as a significant component of faunal assemblages in sites in the Marquesas Islands, Society Islands, Hawaii and New Zealand.

Recent reappraisal of radiocarbon dates has identified two phases in the colonisation of East Polynesia, an initial migration to the Society Islands around 1,025–1,120 CE, followed by a short hiatus than one major pulse to all remaining islands c. 1,190–1,290 CE (Wilmshurst et al., 2011). Dogs are recorded from archaeological contexts associated with human arrival in the Society Islands, such as Vaito’otia (Sinoto and McCoy, 1975), and in Hawaii, from the Bellows Dune site (Pearson et al., 1971). In the Marquesas dogs are found throughout the archaeological sequence, demonstrating that they were present from the initial settlement onwards (Rolett, 1998). At the early Hanamiai Dune site there are cut marks on some of the bones, suggesting they were being eaten (Rolett, 1998: 192). Dogs are also recorded from the initial settlement layers at Ha’atuatua (Suggs, 1961) and Hane, where two dog burials were excavated (Sinoto, 1966). It seems, however, that dogs had almost disappeared by the time of European contact, although numbers appear to have subsequently increased during the historic period, which is possibly related the introduction of European breeds (Rolett, 1998: 37). In New Zealand, the dog, or kuri, was the only domesticated animal successfully introduced by Polynesian migrants, with dog bone present at Wairau Bar, the earliest dated archaeological site in the country, occupied around 1,200 CE (Brooks et al., 2009, Higham et al., 1999).

Although dogs, pigs and chickens all appear to have been successfully introduced on many island groups in Polynesia, there are islands where some, or all, of these animals are not present in archaeological contexts. As a result of these absences, the role of those domesticates that did become established appears to have become elevated in importance, perhaps also due to the scarcity of indigenous terrestrial mammals in these remote islands of the Pacific. For example, Rapa Nui and Niue only had chickens and in both places their
economic significance was much higher than in any other Polynesian islands, as reflected in the historic accounts from Rapa Nui and the archaeological record of both places. In New Zealand, dog bones as food remains are often recovered from middens. Dog bone was also used to make fish hooks, tools and items for personal adornment. At the time of Cook’s voyages in the late 18th century CE, dog skin cloaks were associated with people of chiefly status. There are also recorded instances of deliberate dog burials in New Zealand (Clark, 1996).

Although New Zealand has a large landmass, there is no evidence that feral dog populations ever became established in prehistory. Feral dog populations are, however, recorded following the introduction of European dog breeds in the mid-nineteenth century CE (Clark, 1995). This lack of feral dog establishment contrasts with the situation in both Australia and New Guinea and suggests that the native biota of New Zealand may not have been sufficient to enable dogs to survive on their own. If the early high density biomass of flightless birds ever did support feral populations, there is no evidence of this in either the archaeological or the natural palaeontological record.

The colonisation of Central and East Polynesia is now considered likely to have occurred no earlier than about 950 CE (Wilmshurst et al., 2011) and it is also about this time, or just prior, that dog appears commonly in the Polynesian archaeological record. Given the lack of prehistoric dogs in New Caledonia and Vanuatu it is thus possible that dogs were not introduced into Polynesia as part of the first Lapita colonisation. Instead, their arrival and distribution through Polynesia may have been associated with the burst of movement that resulted in the expansion into East Polynesia from around 1,000 CE (Addison and Matisoo-Smith, 2010). This implies the possibility of a later (post-Lapita) introduction of dogs into Remote Oceania.

Although dog is currently not well documented in Lapita sites its presence is widely assumed (e.g., Kirch and Green, 2001) and is supported by linguistics (discussed below). The presence of a well attested Proto-Polynesian term for dog suggests its introduction well before the colonisation of East Polynesia. If dog was introduced by Lapita colonists, the linguistic and
archaeological record both strongly suggest that it, and the other two domestic animals (pig and fowl), played a very different role in Polynesia than in earlier Lapita society.

**Genetic evidence for dogs in Polynesia**

Ancient DNA analyses of a short (263 bp) fragment of mtDNA obtained from 19 archaeological dog bones from East Polynesia (Hawaii, New Zealand and the Cook Islands) have identified two haplotypes which have been designated as Arc 1 and Arc 2 (Savolainen et al., 2004, Oskarsson et al., 2012). These haplotypes are clearly distinct from those identified in dingo and NGSD. Both Arc 1 and Arc 2, along with the A29 dingo and NGSD haplotype, have been found at low frequency in South China and modern MSEA dogs but not in modern dogs sampled in Taiwan or the Philippines. The distribution and frequencies of the two ancient haplotypes in modern ISEA dogs vary. Arc 2, but not Arc 1, has been identified in Kalimantan, and both have been found in Bali. The Taurama puppy from New Guinea carried a Pacific dog haplotype, as opposed to that of a dingo or NGSD (Matisoo-Smith, 2007), and at present this sample provides the only archaeological link between Polynesian dogs and those in Near Oceania.

**2.5.6 Micronesia**

Dog remains have been found in some of the earliest layers of sites in the Caroline Islands, including Kosrae and Pohnpei, dating to between 50 BCE to 450 CE (Rainbird, 2004). Excavations on Fais, a raised coral reef island, have shown that dogs and pigs were established by 400 CE, following initial settlement up to 200 years earlier (Intó, 2008). In the Marshall Islands, the easterly most atoll chain in Micronesia, dog bones are present in archaeological sites on Utrik and Maloelap, which date to the initial occupation of the atolls around the first few centuries CE (Weisler et al., 2012). The two Polynesian outliers in Micronesia show different patterns. No dogs are known from archaeological excavations on Kapingamarang (Leach and Ward, 1981), but they are present from early horizons on Nukuoro (Davidson, 1968). The distribution of dog remains in Micronesia and the slightly earlier dates there may indicate a source population and direction of movement into
Polynesia, though further dating and aDNA analyses of Micronesian dog remains are necessary to investigate these possible connections.

2.6 Linguistics

Linguistics provides another line of historical evidence concerning the prehistoric dispersal of dogs in the Asia Pacific region (e.g. Kirch et al., 1987, Kirch and Green, 2001). In modelling patterns of dog movement into Oceania we make the simple assumption that in areas of highly efficient expansion of dogs with people, there is likely to be an equally efficient transfer of terms for dog. Conversely, in areas where the introduction of dogs was patchy, discontinuous and punctuated, there is a higher likelihood of variability in terms. Archaeological evidence shows that dogs were present in the early Austronesian expansion into the Philippines and were present in numerous Austronesian communities through ISEA. Given the strong association of dogs and Austronesian speakers we might expect there to be a well attested Proto-Austronesian (*AN) term for dog in the Pacific, but this is not the case. There is a *AN reconstruction (*asu) but it is restricted to languages of Taiwan, the Philippines, western Indonesia and some of the lesser Sundas and south-central Moluccas. It has no cognates in any language further to the east (i.e., in the Eastern Malayo-Polynesian family) (Osmond and Pawley, 2011: 240). One interpretation of this pattern is that there was an interruption in the transport of dogs out of ISEA and into Oceania. A possibility is that the eastern end of ISEA marked the biogeographic boundary beyond which dogs could no longer simply accompany human migrations without direct and sustained human intervention. Thus dogs may not have survived the first wave of human expansion into Oceania and arrived in a series of later movements. Various terms for dog in Western Oceanic languages, and in both Austronesian and non-Austronesian languages of eastern Indonesia, suggest a proto-term resembling *kapuna which spread east into Oceania (and not the other way around) (Osmond and Pawley, 2011: 240, Donohue and Denham, 2010: 226).

There are also hints in the linguistic data that once established in Oceania, there were still limitations in the ability of dogs to freely move with expanding populations. Osmond and
Pawley (2011: 240) (see also Hudson, 1989, Lynch, 1991, Donohue, 1995) point out that in the Oceanic languages the term for ‘dog’ is ‘notoriously’ variable. Indeed there is no clear Proto-Oceanic (*OC) term for dog, although Blust (2002) suggests *guan as the most likely contender. Proto-Oceanic is a good candidate for the language of the Lapita colonists who were the first people to expand into Remote Oceania. It is interesting to contrast the variability in the terms for ‘dog’ in Oceanic languages with the relative homogeneity in cognate forms for the other mammal domesticate – ‘pig’ for which there is a clear *OC term (*beRek). This might suggest that dogs followed a more complex set of pathways through the Pacific than did pig, including periods of loss, followed by reintroduction from different Oceanic sources. On the other hand, the archaeology suggests within archipelagos, pig and dog had a similar history of discontinuous distribution, loss and reintroduction to each other.

In Polynesia, however, the situation appears to have changed for dog and the other two domesticates pig and fowl. The Proto-Polynesian (*PN) terms for all three domesticates are unique innovations with regular reflexes in most daughter languages but not reflected in higher order subgroups (Osmond and Pawley, 2011: 242). This suggests to us that in Polynesia, behavioural adaptations in humans and animals, and perhaps even biological adaptations in the domesticates, occurred that allowed the extremely efficient transmission of these animals with people. Thus we are suggesting that it was the Polynesians, not their Lapita ancestors, who developed the biological and behavioural innovations for the rapid and efficient transport and establishment of domestic animals. Wherever Polynesians travelled they introduced the animals, and the terms they used to describe them. The *PN term for dog for example (*kuli) is well attested throughout triangle Polynesia and has spread westward, with the movements associated with the dispersal of the Samoic Outlier languages to appear in Samoic languages in the Solomons, Vanuatu and Micronesia. A similar pattern can be observed with fowl (*moa) and pig (*puaka).
2.7 Summary

The arrival of dogs in Polynesia marks the end point of process that began with their dispersal beyond East Asia, through mainland and ISEA, and ultimately across long stretches of the Pacific Ocean. Once beyond the mainland, this becomes an increasingly difficult exercise requiring deliberate human intervention. Current understandings based on mtDNA and Y chromosome markers from modern ISEA village dogs suggests a southern route for this dispersal, following that proposed for pigs (Larson et al., 2007). MtDNA data from ancient dogs sampled from archaeological sites, however, suggests there may be more complexity involved in the process, resulting in the introduction of more than one dog lineage into the Pacific. Archaeological and linguistic evidence supports this, demonstrating a patchy and discontinuous process of dog introductions. In the next chapter, I explore additional sources of information that may assist with developing a better understanding of the dispersal process.
Chapter 3

Ecology, behaviour and human interactions of free-ranging dogs

Preface


3.1 Introduction

Beyond MSEA, dog dispersal across the islands of the Pacific becomes reliant on deliberate human intervention. To investigate how this process may have occurred, it is important to consider what kind of conditions might have influenced the establishment and maintenance of dogs in new settlements, and to understand as fully as possible what the nature of the dog-human relationship is likely to have been. One way of doing this is to look at ecological, ethnographic and historical information about dog populations and their interactions with people in the region.

This chapter begins by developing an understanding of free-ranging dog ecology and behaviour. Dog ecology is a relatively recent area of research, and the vast majority of studies have been directed at understanding dog ecology as part of zoonotic disease control amongst human and animal populations (Slater, 2001). In addition, ecological studies of dingoes and other canids and their interactions with wildlife have been undertaken as part of conservation efforts (Hughes and Macdonald, 2013). The African and Asian regions are notable for their populations of ‘free-ranging’ village dogs, where dogs live in settlements, in association with people, but free from direct human control. This way of keeping dogs contrasts with contemporary Western notions of peoples’ relationships with dogs, where they are generally considered as pets or working dogs. For this thesis, I am particularly interested in these studies of free-ranging dog populations in village or rural locations. These
village and rural settings are more likely to be comparable to the environment present during the dispersal of dogs across ISEA and the islands of the Pacific, rather than urban settings with high human population densities and cramped living conditions.

The chapter then moves to a consideration of what information can be gleaned from ecological, ethnographic and historical observations about dog-human interactions in the Asia Pacific region. These descriptions are informative about the nature of free-ranging dog populations, as well as the various kinds of interactions people may have had with dogs in similar settings in the past. It is possible that some characteristics of dog populations, and their interactions with people, also have the potential to influence the contexts of archaeological samples and hence the interpretation of molecular genetic data.

3.2 Ecological studies of free-ranging dog populations

The wide range of different circumstances in which dogs live alongside human communities becomes quickly apparent in a review of ecological studies about dog populations. Ecological studies use two main factors to categorise the nature of the dog population being studied: the relationship between the dog population and people, and the environment that the dogs inhabit. Key criteria may include whether or not the dogs are owned, restrained or able to roam freely, their attitudes towards people, and whether the population is situated in an urban, village or rural location. Vanak and Gompper (2009) reviewed common classificatory categories used in studies of dog populations. These categories are summarised in Table 3.1, and demonstrate the variety of contexts in which dogs may be associated (or not) with human communities.
Table 3.1: Categories of dog populations (after Vanak and Gompper, 2009)

<table>
<thead>
<tr>
<th>Category</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owned dogs</td>
<td>Dogs that are owned by people, and have their movements restricted to a particular outdoor or indoor area.</td>
</tr>
<tr>
<td>Urban free-ranging dogs</td>
<td>Dogs that are not owned by people, but are commensal and live primarily on food derived from people’s activities</td>
</tr>
<tr>
<td>Rural free-ranging dogs</td>
<td>Dogs that are owned or associated with people, but not confined to a particular outdoor area.</td>
</tr>
<tr>
<td>Village dogs</td>
<td>Dogs that are associated with villages in rural environments, and are unconfined, but do not leave the immediate vicinity of the village.</td>
</tr>
<tr>
<td>Feral dogs</td>
<td>Dogs that are completely disassociated with people, and are independent of food derived from human activities.</td>
</tr>
<tr>
<td>Wild dogs</td>
<td>Dingoes, feral dogs and their hybrids in Southeast Asia and Australia, that were originally associated with people, but have a subsequent history of independence and are no longer considered domesticated.</td>
</tr>
</tbody>
</table>

Vanak and Gompper (2009: 268) also noted that most researchers agree that these categories are flexible and dogs may fall into more than one category or shift from one to another. Dog populations may also be made up of sub-populations, such as dogs that share peoples’ homes, those that are unrestrained but rely on people for food and shelter, and those that are free-roaming with less reliance on people for food and shelter. These sub-populations may have differing demographic and behavioural traits (Hughes and Macdonald, 2013). There are other issues inherent in defining dog populations which are relevant when considering dog-human relationships, such cultural perceptions about ownership and dog-keeping. There may be differing perceptions of dogs across cultures, and ambivalent attitudes to dogs within cultures (Serpell, 1995a).

The concept of “ownership” is unlikely to mean exactly the same things across the various cultures where studies have been carried out. Wandeler and colleagues (1988: 686) contrast the rights and responsibilities of dog ownership in industrialised countries with those of other parts of the world, where the obligations may appear on the surface to be less
restrictive, but actually dog ownership and attendant responsibilities may still be proscribed by social rules.

Pet-keeping, similar to ideas about ownership, has a particular meaning in industrialised society, where the pet population has been expanding at a rapid rate since the 1960s (Serpell and Paul, 2011). In contrast to ecological studies addressing epidemiological issues and wildlife protection in developing countries, many dog behavioural studies focus on individual dogs living in modern households as pets. Pet-keeping, however, varies widely across human societies, and in some cultures pets, including the linguistic term for them, do not exist (Herzog, 2014).

In western society, conceptions of domestic animal-human relationships tend to be polarised between those with individual animals living with people as companion animals and pets, and aggregations of animals which are farmed and are removed from day to day contact with people outside the farm environment. Yet there are many other possible ways in which animals and people interact today and in the past. These interactions are addressed in more detail in section 3.2, but first I describe the main characteristics of free-ranging dog populations from ecological studies.

3.2.1 Population structure

The structure of a given population can be described in terms of reproduction, mortality and age distribution. Modern dogs average two breeding cycles per year, this contrasts with wild canids, including dingoes, which have a single annual breeding cycle. However, some dog populations, including Basenjis, Indian pariah dogs, and West Bengalese dogs have an annual breeding cycle (Pal, 2001). Some studies of reproduction in free-ranging dog populations further indicate the possibility of seasonal or synchronised breeding (e.g. Pal, 2001), however this has not been observed in all studies. Litter size is variable, depending strongly on maternal size and age, but an analysis from 15 locations gave a mean litter size of 4.9 (± 1.2) (Gompper, 2014).
Dogs have a high reproductive potential. Wandeler and colleagues (1988: 684) state that under optimal conditions a given dog population could nearly triple every year, but is limited, as with other species, by the carrying capacity of the environment. The carrying capacity depends on the pattern of availability of resources such as shelter, food and water, and their quality. They also note that any reduction in the dog population because of increased mortality, may be compensated for by improved rates of reproduction and survival in the remaining dogs during the period when competition for resources is reduced.

Wandeler and colleagues (1988) also comment that the primary difference between wild and domestic animals is that humans control resources for domestic animals, and make the observation that the carrying capacities for dog populations are closely related to the human groups with which they are associated. Vanak and Gompper (2009) also point out that because of human subsidies, free-ranging dogs may be insulated from food scarcity and can achieve high-density populations. For example, in an area of the Gokwe Communal Lands in Zimbabwe all dogs are owned but free-ranging, and breed freely whilst also depending on people for their basic needs (Butler and du Toit, 2002). The exponential growth of the human population in the past century has led to an equally rapid increase in dog numbers. In 1954 there were 250,000 dogs nationally, but by 1994 dog numbers in the communal lands alone had reached 1.36 million, with an annual growth rate of 6.5% (Butler and Bingham, 2000).

The carrying capacity of the environment appears to be a significant factor in the size of free-ranging dog populations. The extent to which the patterning of dog remains in archaeological sites may reflect the carrying capacity of the human population, and the possible implications of population size on genomic analyses, will be considered in more detail in following chapters.

3.2.2 Life expectancy

Counter to their high reproductive potential, high early mortality is also common feature of free-ranging dog populations. Causes of death included human actions (either directly through deliberate culling, or indirectly through car accidents), poisoning, disease, and old age (Acosta-Jamett et al., 2010, Gompper, 2014). The mean average age of dogs in 37
populations studied around the world was 2.8 (± 1.2) (Gompper, 2014). Most populations studied comprised young dogs, with only a small percentage of old dogs. This contrasts with the greater ages reached by confined pets (Beck, 1975). Daniels and Beckhoff (1989) observe that a free-ranging existence, in modern dogs at least, may therefore reduce the chances of survival markedly. They also note that breeding females required much greater energy requirements during gestation and lactation, which may be difficult to maintain if resources are scarce. Dogs do not regurgitate food for their offspring (Newsome et al., 2001), making pups highly vulnerable to the availability of food when young.

It is notable that feral dogs in Italy and India, with no direct human contact, also suffer high rates of mortality, to the extent that groups appear unable to maintain their population levels without immigration from outside the group (Pal, 2001, Boitani and Ciucci, 1995). This is also the case for free-ranging community dogs in the central Philippines, where community-owned dogs in villages are usually juveniles or young adult dogs, not associated with households and forced to become self-sufficient in terms of food (Beran, 1982). Beran (1982) observed that the populations of these dogs sat at just at, or slightly above, the carrying capacity of the environment in terms of available food. Their fecundity, survival of puppies and life expectancy was too low to sustain the population without continual additions of surplus household dogs. The removal of dogs for butchery for food had no lasting effect on population size, however, as only some dogs were selected, and their place was then taken by other dogs from the communally-owned group.

3.2.3 Sex ratio

Studies of several free-ranging dog populations observed a skewed sex ratio towards males. In North America, a study of dogs in two settlements on a Navajo reservation, found ratios of 4:1 and 3:1 (Daniels and Bekoff, 1989). This also corresponds with sex ratios in urban areas in Baltimore, Maryland (Beck, 1975) and in New Jersey (Beck, 1973). In Chile, the sex ratio was also skewed towards males (Acosta-Jamett et al., 2010). As there is not a bias towards either sex at birth, the reasons for these differing ratios are most likely to relate to human influences. Daniels and Bekoff (1989) suggest that males may be selected as pets hence
increasing male numbers in urban dog populations, and that females are deliberately removed from the population by people. For example, in the Coquimbo region of Chile, people used population control measures such as selectively killing female puppies (Acosta-Jamett et al., 2010). In India, however, feral dog populations also show a male skewed sex ratio, but are free from human influences (Pal, 2001). Pal (2001) points out that further research is required to explain the male bias in feral dog populations, which may relate to male based immigration into study areas at particular times of the year.

3.2.4 Diet and feeding behaviour

The feeding behaviour of dogs is essentially opportunistic scavenging. Dogs are generalists, capable of consuming and surviving on a wide range of food types, ranging from human-derived garbage to animals that may be several times their body mass (Hughes and Macdonald, 2013, Vanak and Gompper, 2009). Dogs may also consume amphibians, reptiles, birds, rodents, insects and non-food items (Campos et al., 2007, Hughes and Macdonald, 2013, Newsome et al., 2001).

Studies of free-ranging dog populations from around the world have identified the importance of food derived from human activities in the dogs' diet (Vanak and Gompper, 2009). This includes deliberate feeding, and scavenged food refuse (plants and animals), remains of livestock and other carcasses, and faeces. Although some studies have documented dogs killing and feeding on wildlife, Vanak and Gompper (2009) point out that such studies are generally focused on the effects of predation on the prey species, rather than the overall picture of dog foraging ecology. They argue that reliance on human derived materials is typical, even when wildlife is also killed and consumed.

In Zimbabwe, for example, a large proportion of the country’s dog population is located in communal lands (rural areas of traditional agro-pastoralism), many of which share borders with wildlife reserves. The impact of free-ranging dog populations on wild carnivores, as carriers of canine diseases and as potential competitors, is a matter of concern. Butler and du Toit’s (2002) investigation of the diet and scavenging behaviour of dogs on the communal lands demonstrates the dogs’ reliance on human waste and animal carcasses. Homesteads in
the communal lands surveyed for the study were typical of those found throughout in the communal lands, comprising thatched buildings with un-fenced yards. Household waste such as left-over food was disposed of in pits at the edge of the yard. While 8% of households had pit latrines, the majority of residents defecated in the open.

The diet of the Zimbabwe communal lands dogs, recorded on the basis of observations made in the field, was wide-ranging. It included plant matter, mammals, insects, birds, reptiles and fish, and faeces (Butler and du Toit, 2002). The three most important items by mass were sadza (a maize/sorghum/millet porridge which is the local staple food), human faeces and cow carrion. Of the food eaten, only 13% was fed to the dogs by people, the remaining 87% was independently obtained. Over half of this food was scavenged from mammalian carcasses, with human faeces contributing a further 25%. Although dogs obtained a major proportion of their diet independently of people, human-derived food made up 88% of the diet. The availability of carcasses of domestic animals appeared to be an important aspect of the diet. Over the course of Butler and du Toit’s study 59% of domestic animal fatalities were left where they died, rather than being burned, buried or eaten. Dogs proved to be efficient scavengers, finding over two thirds of the animal carcasses available to them over the study period.

Studies of two ISEA dog populations also illustrate the dogs’ reliance on human derived foods. In the Philippines rural barangays, dogs associated with a household were fed regularly, and dogs also scavenged garbage, unused portions of plant and animal food, and frequently faeces (Beran, 1982 :268). Community owned dogs were not fed, and had to be self-reliant in terms of obtaining food from garbage piles and elsewhere in the village. In Flores, most dogs are not deliberately fed, but may be given left over food scraps by their owners (Newsome et al., 2001). They would also hunt and scavenge for food, including eating unattended food, which was seen as a negative aspect of having dogs.

### 3.2.5 Social organisation

Social organisation as applied to the study of dog populations refers to spatial relationships, group composition, and social interaction patterns (Bekoff and Wells, 1986). Although
popular culture assumes that dog behaviour resembles that of wolves, ethological research has demonstrated that domestication has had a profound impact on how dogs relate to each other, and to people (Miklósi, 2007). The notion that dogs are pack animals with strict hierarchical leadership structures in groups has been shown to be an over-simplification. Overall, studies show that there is a stronger tendency for dogs to avoid each other, rather to form groups (or packs). Dog density influences the social system, with the tendency for positive social interactions decreasing in relation to increased dog population (Daniels and Bekoff, 1989). Other factors affecting social organisation include distribution of shelter and food resources, and changes in these in relation to changes in population size.

Daniels and Bekoff (1989) suggest that for dogs in urban and rural communities, group living is likely to confer little advantage. As Beck (1975) observed, food resources beyond those provided by people were able to be exploited more efficiently by individuals rather than groups. He also noticed that dog ownership practices influenced social behaviour, with dogs which shared a home site with other dogs because of multiple dog-ownership, socialising more frequently with each other. In Flores, dogs did spend time together, often forming groups with dogs from neighbouring households (Newsome et al., 2001).

In a North American study of free-ranging and feral dog populations on a Navajo reservation, the free-ranging dogs exhibited territorial behaviour in relation to their home site only (Daniels and Bekoff, 1989). As food was also provided by the dogs’ owners at the home site, these locations represented small, easily defended areas of resource abundance. The lack of social aggregations by dogs in the reservation villages contrasted with the social behaviours exhibited by two feral dog packs that were also part of the study, where a large proportion of the dogs were observed living in packs either year round, or seasonally.

Limited studies to date have shown that in relation to range or spatial areas, free-ranging dogs in rural areas spend most of their time close to their associated households, within infrequent forays away from home (Sepúlveda et al., 2015). When forays do occur, the route taken primarily follows human-dominated landscapes.
3.3 Interactions with people

Only a small number of ecological studies have specifically considered the interactions of dog populations with people, beyond determining ownership status and the level of control exerted over the dogs’ movements. However, two studies in ISEA that addressed the spread of rabies both included human-dog interactions. In Flores, Indonesia, Newsome and colleagues (Newsome et al., 2001) argued that the ecology of dogs on the islands would be incomplete without consideration of the strong bonds between dogs and people, and the effect that these have on dog population size. In the Philippines, understanding the relationship between dogs and human society was also identified as a critical aspect of successful rabies control, as often the failure of control programmes in developing countries was due to inappropriateness in terms of people’s relationships with dogs, rather than technical issues (Beran, 1982). In addition to Flores and Philippines, this section also draws on ethnographic research from New Guinea and Australia, and looks specifically at historical observations about dogs in New Zealand at the time of the first European arrivals and over the early European colonisation period.

3.3.1 Island Southeast Asia

In the central Philippines, Beran (1982) studied dog populations in urban areas, and also rural, coastal and mountain barangays (villages). The barangays were described as highly homogeneous social units, with residences close together and connected with paths, rather than being separated by fences. Two categories of dogs were identified in the rural barangays, those dogs owned by a particular household, and also other dogs which were community-owned. People still knew however which household the community dogs may have originated from, or when they had arrived in the village. Most households in the barangays owned one or more dogs, which were guards, scavengers, and were sometimes used as a source of meat. Dogs, however, were not treated as pets to be handled or played with, they were considered part of the village but were mainly left to themselves. The dog populations in the barangays were highly independent, and wary of unusual actions that might threaten their freedom. Generally, dogs and people paid little attention to each other, and the dogs
appeared docile and detached from what the people were doing, unless they perceived approaching people as threats.

In Flores, an island in Indonesia, dogs were also observed as being an integral part of the environment and had a wide variety of interactions with people (Newsome et al., 2001). They guarded people’s homes and gardens, and people enjoyed their company and considered them pets or friends. The dogs were not allowed in houses, instead they lived around the outside and slept under the eaves, sheltered by sheets of bamboo or wood. Adults and children were often observed cuddling and playing with dogs; which contrasts with the attitudes of people in the Philippines. In addition, dogs were eaten and played a role in certain ceremonial activities. The long distance movement of dogs in Flores was strongly associated with people; dogs either followed people into strange areas, or were deliberately taken long distances on trading trips or other visits. In addition to following their human companions over land, dogs may also travel with people in boats and thus could be transported across water barriers.

Dogs interact with other animal species in their environment, often as part of their wider interactions with people. These are mostly other domesticated species, and commensal species which also inhabit human environments. In Flores, dogs not only guard against human interlopers, but also against the intrusion of other animals (Newsome et al., 2001). Dogs guard people’s crops against wild and free-ranging domestic pigs, monkeys and rodents, and chase away any that come near. The dogs do not, however, engage with chickens, which are also common on the island (Newsome et al. 2001). Although it has been suggested that dogs and pigs may have been competitors for food in the Pacific in the past (Bay-Petersen, 1979), it appears that the animal husbandry methods of some human communities may mediate against direct competition. In Flores, domestic pigs were usually kept in pens and fed by people, while the dogs were free-ranging and obtained much of their own food without human assistance (Newsome et al., 2001).
3.3.2 New Guinea and Australia

Ethnographic research in New Guinea has documented various aspects of peoples’ relationships with dogs, including their use for hunting, as companions, and as valuables in social exchange. Hunting dogs are used to discover, flush out, drive and run down large game such as pigs, wallaby and terrestrial cu cu, and hold arboreal mammals up trees and prevent them from fleeing across the ground (Bulmer, 1968). Chowning’s (1991) ethnographic work with the Molima on Fergusson Island, in the D’Entrecasteaux group off the eastern tip of New Guinea, demonstrates the complexity of people’s relationships with dogs there. In contrast to pigs, dogs are not only valued for food for ceremonial occasions, some are used for hunting and most are pets. Dogs are not eaten by their owners, rather they are passed to others for food, in accordance with social obligations and exchanges. Dog’s teeth are also minor valuables and are removed after the death of the dog. Although dogs are recognised as non-human, a dog’s owner may be referred to as its parent, and their children its siblings. Keeping dogs can, however, cause problems for people as they steal food, cause property damage, kill and eat young pigs and bite people. Dwyer and Minnegal (2006) document the problems the Kubo people had introducing chickens to their village due to predation by their dogs. Following a distemper outbreak that killed most the village dogs, chickens were reintroduced. Any subsequent dogs were able to be trained to avoid the chickens.

In Australia and the New Guinea Highlands, dogs descended from Southeast Asian lineages are known respectively as dingoes and New Guinea Singing Dogs (Savolainen et al., 2004). Dingoes and NGSDs are highly unusual as they are descended from Southeast Asian domesticated or village dogs but they have successfully adapted to living independently of people. This has resulted in some debate about whether these groups of dogs should be considered a separate species, sub-species, or a distinct breed (Bulmer, 2001). Despite their ability to live without human aid, animals drawn from these feral populations may still live in human settlements and they are part of various cultures in these areas. Although they were introduced to Australia around 3,500 years ago (substantially later than initial human arrival) (Gollan, 1984), dingoes have been absorbed into indigenous culture and are part of the creation stories of the Dreaming. Dingoes (and more recently European dogs) are an
important part of indigenous communities, and interact with people in a wide range of ways. Although they are not generally eaten or considered useful hunting aids, they are highly valued as protectors and guardians (both physical and spiritual) and as companions (Meehan et al., 1999, 2009). Contemporary attitudes towards dingoes and camp dogs also demonstrate aspects of the ambivalence noted towards some pets and animals within cultures (Serpell, 1995a). Meehan and colleagues (1999) observed a lack of care for camp dogs in Anbarra society, including people beating dogs, allowing them to fight, and breed freely.

It is notable that in Australia, the Torres Strait and parts of New Guinea where dogs were incorporated into existing societies that did not previously have dogs, these groups do not usually consider them to be an acceptable food source. In Australia, dingoes are not eaten, unless as a famine food, and even then with great reluctance (Smith and Litchfield, 2009). In the Torres Strait, dogs are also not eaten; although their remains are often found in ritualized midden deposits reflecting their status as spirit totems (McNiven and Feldman, 2003). Most Highland groups in New Guinea also did not eat dogs (Bulmer, 1975, 2001). Attitudes to dogs further east in Remote Oceania provide a strong contrast. In East Polynesia, European historical accounts record that dogs were valued as a food and were often eaten during ceremonial events (Luomala, 1960b), and dog remains are often found in archaeological contexts associated with the disposal of cooking waste (e.g. New Zealand Davidson, 1987).

3.3.3 New Zealand

Observations about dogs in the Pacific are found throughout European historical literature dating from the initial European exploratory voyages to the subsequent settlement of New Zealand. This literature comprises ship’s journals, including those kept by Captain James Cook and his scientific officers during Cook’s three voyages to the Pacific from 1768–1779, followed by published accounts of experiences and travels in New Zealand, and papers of scientific interest. In addition to written accounts, there are also paintings and drawing of dogs and other animals from this period (Luomala, 1960a). Although it is likely that aspects of dog behaviour may be referred to in Maori oral traditions, drawing on such source material is beyond the scope of this thesis.
There are numerous references to dogs in the historical literature, although the majority of journal accounts address the description of physical characteristics, sometimes with a brief comment about the uses of dogs by Maori (for a review see Colenso, 1877). Later authors, making observations about life in New Zealand, also mention dogs in passing. It appears, however, that as dogs were familiar to these Europeans, they received cursory attention in comparison to aspects of Maori culture that were new or unusual to European eyes. Nonetheless, from these accounts it is possible to gain some information about the first dogs introduced to New Zealand, their biology, behaviour and relationships with people.

It is difficult to obtain a clear picture of the numbers of dogs associated with settlements when contact with Europeans first occurred. During Cook’s first voyage to New Zealand he made the observation that there were no four footed animals, with the exception of dogs and rats, and that these were very scarce. During his second voyage while anchored in Queen Charlotte Sound, both Cook and Forster comment on there being plenty of dogs, including those travelling with people in canoes (Colenso, 1877: 143). Although the numbers may have varied, dogs were seen in association with people and settlements throughout the country.

The majority of dogs seen by early European travellers were un-restrained, with the freedom to roam throughout settlements and scavenge for food. Dieffenbach (1843) and Earle (1832) both describe villages with fenced enclosures or houses to prevent entry by dogs and pigs (introduced by Europeans). Dieffenbach (1843: 111) noted that favourite dogs and pigs were allowed access to the interior of some enclosures. Some people slept with their dogs. Many travellers described ‘wata’ or ‘fata’, which were storehouses on poles or stilts several feet above the ground, used to store and protect food such as seed potatoes, or dried fish from dogs, pigs and rats. Earle (1832: 36) also noted that left over food was kept between meals in baskets on poles, to protect it from dogs and pigs. Stages or platforms above the ground on poles were also recorded, where items associated with deceased people were kept. It is likely that these were also designed to keep important objects away from the dogs and other animals in the villages.
Observations made by European travellers describe dogs being fed food scraps or left to fend for themselves. Dogs were fed 'the skin of potatoes and well picked fish-bones' (Polack, 1838: 437). One author, Crawford (1880: 162), in his recollections of earlier travels made special mention of seeing a dog being fed a potato, as it was the only instance in all his travels that he had seen such a thing. During Cook’s second voyage observations were also made of dogs eating human remains which were left lying on the beach after an altercation with Cook’s crew (Buick, 1900). Dieffenbach (1843) and Ward (1840) in the 1830s both commented on how lean the dogs were, and Polack (1838: 220) noted that he had seen a few older dogs that had survived being eaten.

Several uses of dogs were noted—dogs were watch dogs, used for food, and for clothing. Although dogs were not observed to bark, they would howl and cry when strangers entered a village, and were known to prevent attacks by stealth. Thomson (1859) and Colenso (1877) both mention briefly the castration of dogs to improve their size and flavour, and Colenso makes reference to an unusual way of doing so that is not described. The use of dogs to hunt birds is also recorded by several authors, however, these accounts date from the second half of the 19th century, and may represent a relatively recent activity not necessarily involving non-European dogs. Several commentators recorded the affection that people had for their dogs. During Cook’s second voyage, Forster thought people seemed fond of their dogs, and Colenso (1877) records Forster’s observation that people ‘gave names to them, prized and petted them’.

By the 1830s, with the introduction of European breeds, dog numbers appear to have increased to the point of being considered a nuisance. Dogs had been brought to New Zealand by Europeans as early as Cook’s voyages, and sealers and whalers also brought dogs with them (Clark, 1995). Cook also transported European and Polynesian dogs around the region (Colenso, 1877). Hunting dogs in particular were introduced relatively quickly by sealers and whalers. These dogs may have formed the basis of packs of wild dogs later recorded in many regions. Wakefield (1845) noted that pig dogs were readily available and whalers also bred dogs for fighting. Earle (1832: 36) considered dogs to be the worst
introduction by Europeans, due to their rapid increase and the injuries they caused to other animals. By this time, pigs and cats were also often noted as being present along with dogs. In 1856, Fitton (1856: 371) in his advice for European settlers states that the passage for a dog is 5 pounds, but that it was not necessary to bring a pig dog as there were plenty in New Zealand already.

By the mid-nineteenth century, Polynesian dogs had become indistinguishable as a distinct breed as they had been completely absorbed into the burgeoning European dog population. Bright, in his handbook of 1841, advised that the native dog had disappeared, but that every [Maori] family had a pack of dogs, but they were mongrels (Bright, 1842: 92). Shortland (1851) recorded that the native dog was present, but seldom in its original purity. Thomson (1859) made the observation that dogs were not protected as a breed once pigs arrived, suggesting that pigs may have taken over at least some of the functions of dogs.

Packs of wild dogs present in the mid-19th century appear to be feral European dogs or Polynesian dog hybrids, which in the South Island coincide with the arrival of flocks of sheep in central regions. These animals would have provided a source of food for feral dog packs. Thomson (1859) notes that in the 'Middle Island' (the South Island) the so-called wild dog packs were not native dogs, rather 'shepherd’s curs'. Dieffenbach (1843) also observed that a native dog could not bring down a sheep (presumably because of its small size), but that cross-breeds and introduced dogs would do so.

3.4 Characteristics of free-ranging dog populations

This chapter has drawn on modern dog ecological studies to gain information about the possible nature of free-ranging dog demographics, diet, social organisation, home range, mobility and interactions with people. While historical observations about dogs recorded in the eighteenth and nineteenth centuries are less systematic and often less detailed, they still provide information about diet, behaviour and dog-human interactions. Both types of information demonstrate the ways in which dogs may co-habit spaces with people, often not as pets, but as a group sharing the living spaces and resources with people in a village setting.
Even within this free-ranging lifestyle, some dogs may be treated differently to the group as a whole, being kept as companion animals or for a particular use, and having a different relationship with people.

There are many characteristics of free-ranging dog populations that are relevant to the interpretation of archaeological and molecular evidence for dog introductions and dispersals in the Pacific:

- Dogs are opportunistic scavengers, with access to food refuse or other human derived material and do not necessarily need to be fed deliberately, or have a large protein component in their diet.
- Although the dogs may exhibit some aspects of hunting behaviour, they are unlikely to be efficient hunters of wild prey, and are thus dependant on people for food.
- Dog population levels are highly associated with human populations, resource availability and the carrying capacity of the domestic niche, and can fluctuate rapidly. Disease outbreaks, and deliberate culling can also have a significant impact on population numbers. Despite high puppy mortality, free-ranging dog population numbers can triple in a year.
- In New Zealand, no evidence has been published to date that supports the presence of feral dogs until the arrival of European dogs in the early 1800s.
- Free-ranging dog populations are characterised by the lack of spatial control by people over their movements. This means that breeding is likely to be uncontrolled. Evidence from free-ranging dog studies also suggests that in some instances there may be a skewed sex-ratio, towards males.
- Dog distribution, and also mobility, is likely to be closely related to associated human settlement patterns, social and trade networks. Some dogs are highly mobile, accompanying people on travel by foot and boat. Dogs may move between settlements with their owners, but also potentially as items of exchange. Localised movements of small numbers of dogs may be a part of sustained contact between human communities.
The accounts from across Asia and the Pacific demonstrate the complexity of the interactions between dogs and people in village settings, and also highlight possible characteristics of free-ranging dog populations that may influence the contexts of archaeological samples and interpretation of molecular genetic data. Taking these issues into account, the following chapter examines archaeological models for human migration in the Pacific, and the use of phylogenetic analysis, with a specific focus on the relationship between these models and methods and understanding the early history of dogs in the Pacific.
Chapter 4  
Models and Methods

This chapter addresses the two components that, when combined, make up the conceptual and methodological framework for this thesis: the archaeological models for two periods of Pacific colonisation, and the molecular genetic methods which are used to generate data to investigate these processes. The implications arising from the use of this kind of framework are also considered.

This thesis focuses on two periods of colonisation in the Pacific that involved the transportation of dogs: the emergence and distribution of the Lapita Cultural Complex (LCC), and the later human colonisation of New Zealand. The LCC is essentially an archaeological construct which unites a particular set of archaeological artefacts (pottery, stone, shell and bone) and other aspects inferred from the archaeological record, such as settlement pattern, subsistence economy and exchange networks (Kirch, 1997: 13). These archaeological markers are argued to be distinguishable from archaeological evidence of the earlier inhabitants of Near Oceania (but see Specht et al., 2014 for an alternative view). The LCC emerges in the archaeological record of the Pacific (from the Bismarck Archipelago to West Polynesia) almost simultaneously around 3,500 years ago, and is associated with the spread of an ethno-linguistic group of people from ISEA who probably spoke Austronesian languages (Kirch, 2000). The colonisation of New Zealand which took place much more recently, around 700 years ago, marks the final stages of the human colonisation of the Pacific. Unlike the Lapita expansion in Near Oceania, New Zealand was previously uninhabited by people.

The development of aDNA methods and NGS technologies offer opportunities to broaden the kinds of information we can use to investigate past patterns of animal population history and interactions, and the way these may articulate with colonisation processes. The analytical methods used may however have been developed for other circumstances than the
investigation of material recovered from archaeological sites, and may therefore influence possible interpretations.

The first section of this chapter sets out the archaeological models, and explores the possible implications and expectations in relation to the origins and dispersal of Pacific dogs. In the second section, the methods used to generate molecular genetic data are outlined, particularly the main methods for the construction of phylogenetic trees and networks, with particular reference to underlying parameters and assumptions. The third section addresses the use of aDNA obtained from archaeological samples, while the specific methods used to generate and analyse the ancient mitogenomes for this thesis are documented in the following chapter. This chapter concludes by discussing some of the implications of drawing on molecular phylogenies, inferred from archaeological samples, to investigate past patterns of animal population history and interactions.

4.1 Archaeological models for human migrations in the Pacific

There are two main types of models in archaeology; firstly, explanatory models developed to assist with understanding the phenomena under study, and secondly, formal models, developed using mathematical or statistical approaches from which hypotheses are derived and evaluated using archaeological datasets. In this thesis, I am concerned with the group of explanatory models developed to explain the archaeological phenomena of the LCC, and the colonisation of New Zealand. These explanatory models conceptualise and simplify a complex set of processes, and focus attention on what are believed to be the most important factors and the relationships between them (Clarke, 1972).

4.1.1 The Lapita Cultural Complex

Models for the emergence and spread of the LCC through Near Oceania and out into the islands of Remote Oceania usually rely in varying degrees on archaeological data. The outcomes of linguistic studies of language patterns, and molecular studies of phylogenies of humans and animals have also been used to explore the implications of these models, to expand or challenge them, and to develop alternate models. Archaeological models are based
primarily on the spatial and temporal distribution of Lapita-associated archaeological markers, including material culture, structural features, and settlement pattern, and attempt to extrapolate a broad explanatory framework in which further evidence can be accommodated.

Green’s ‘Triple I’ model (1991b), and later revision (2003, see below) represent the most recent attempt to provide a descriptive and explanatory archaeological model for the LCC. This model takes into account the substantial body of archaeological work that was carried out in the 1970s and 1980s that addressed initial questions about the pottery of the LCC, and distribution and chronology of associated archaeological structural, faunal and other material culture evidence. Green’s (2003) elaboration of the Triple III model acknowledged the concept of a ‘voyaging corridor’ (VC-III), advocated by researchers such as Irwin (1992). This voyaging corridor describes the area of islands and sea in Wallacea, where movements across water could be undertaken in short hops, never out of sight of land. One of the fundamental aspects of the VC-III model is the identification of three mechanisms (intrusion, innovation and integration) by which elements of the LCC are brought together to form the complex. Intrusions are aspects of the complex brought into the area by LCC migrants, innovations arise within the LCC during the colonisation process, and finally integration describes the incorporation of some aspects into the LCC sourced from local Near Oceanic cultures.

Since Green’s VC-III model there have been no substantial revisions. Ongoing archaeological work and human and animal genetic studies continue to provide data that enable the evaluation of the VC-III model, resulting in some researchers questioning the appropriateness of the LCC in its entirety, or pointing out archaeological and other evidence, or the lack thereof, that appears contrary to the model (e.g. Addison and Matisoo-Smith, 2010, Donohue and Denham, 2010, Specht et al., 2014). Questions been raised about the appropriateness of conceptualising the LCC as a unified ‘package’ of characteristics that came into being at some presently unidentified point and was exported as a single unit across the region. Specht and colleagues (2014) ‘deconstructed’ the archaeological evidence for the
LCC in the Bismarck archipelago and argue that rather than appearing as intrusive elements introduced rapidly by new migrants, the evidence suggests a gradual introduction over several centuries. Some engagements with explanations for the LCC highlight those phenomena that do not ‘fit’ with the existing model. For example, Sheppard (2011) points out that the islands of the Solomon Island chain lack evidence of early Lapita settlement, and proposes that the pattern of Lapita colonisation is better described as a ‘leap frog’ rather than a standard ‘wave of advance’ movement.

Concerns about the place of domesticated animals in the LCC ‘package’ have been noted for some time. Interestingly, in the VC-III model, Green (2003 Table 6) includes dogs within the group of the commensal animals that were commonly included within the LCG, but had yet to be confirmed as part of the archaeological core elements. Archaeological work and faunal analysis has continued to demonstrate the patchy nature of the archaeological evidence for domesticated animals in Lapita sites. It seems that dogs, pigs and chickens did not arrive in Near Oceania as a group with the first appearance of the LCC, despite the attraction of the transported landscape concept. In addition, the mechanism by which dogs were being moved into the region is not necessarily made explicit. It is possible that dog dispersal does not always equate with the movement of particular groups of people.

Dogs have been documented from archaeological sites in ISEA, and across Remote Oceania, but are scarce in Near Oceanic LCC era archaeological sites (see review in Chapter 2). This hiatus in the presence of archaeological remains needs to be taken into account when considering the potential of aDNA to explore aspects of human-dog interactions during the spread of the LCC, and when interpreting the role of dogs in the Pacific through time. Gaps in sampling coverage may not always relate to the availability of samples or taphonomic processes, rather to an underlying absence of dogs altogether. Given this patchiness, the limited archaeological work in some areas, and also the challenges of the survival of DNA in tropical areas, what are the expectations of VC-III that might be able to be investigated using ancient mitogenomes of dogs?
One of the main aims of models for the appearance and distribution of the LCC is to account for the varying components as either intrusions bought by incoming groups of people, as items already present within in the local cultures of Near Oceania and integrated into the LCC, or as LCC innovations. Green (1991b) notes that innovation may relate to material culture and technology. As well as questions relating to the origins of the domesticated animals found in LCC archaeological sites, it is possible that the LCC may involve changes in the ways that people view and use animals. Changes in the place of animals in the LCC may be complex, and incorporate aspects of both integration and innovation.

To date, only very small fragments of ancient mtDNA have been sequenced from ancient dogs in ISEA and Near Oceania, from a dog from East Timor (Gonzalez et al., 2013) and a puppy from the Taurama archaeological site in Papua New Guinea (Matisoo-Smith, 2007). Interpretation of these results was limited by the small fragment size and the lack of comparative data, both archaeological and modern. The Taurama puppy did appear to carry the Arc2 lineage also found in ancient Polynesian dogs. Comparison of aDNA from archaeological samples from Polynesia with modern dogs and dingoes does suggest an Indonesian link (Savolainen et al., 2004, 2012) and possibly also links with Taiwan (Sacks et al., 2013), as reviewed in Chapter 2. The contribution of dogs from either place is not however mutually exclusive. In addition, given the absence of dogs in Lapita archaeological sites immediately prior to the settlement of East Polynesia, genetic links between East Polynesian dogs may be more informative about events occurring sometime after the Lapita colonisation of West Polynesia.

There are several possible sources for the dogs found in Lapita archaeological sites in terms of the parameters of the VC-III model. Were dogs 1) an intrusive element brought by people moving south from Taiwan and the Philippines, 2) obtained from local populations present on the New Guinea mainland, as modern evidence from genetics suggested that NGSDs were present in New Guinea prior to the Lapita migrations, or 3) accessed via the voyaging corridor linking Near Oceania with southern ISEA, or some combination of these possible source populations? And, at the end of the process, are East Polynesian dogs descendants of
these Lapita dogs? These scenarios will be investigated using ancient mitogenomes and the results presented in Chapter 7.

4.1.2 The colonisation of New Zealand

The colonisation of New Zealand took place at the terminus of the movement of people across Remote Oceania, which began with the appearance of the LCC in the uninhabited islands of the Southeast Solomon Islands some 3,000 years earlier. The eastern most extent of the LCC is marked by archaeological sites in Tonga and Samoa (Kirch, 2000). After a pause of some fifteen hundred years, people moved out from West Polynesia to rapidly settle the islands of East Polynesia. This colonisation event is again marked by a highly visible archaeological horizon appearing in the previously uninhabited islands of Eastern Polynesia. Archaeological affinities with material culture and linguistics relationships between Maori and other Polynesian languages indicate that the people who arrived in New Zealand came from the islands of Central Eastern Polynesia (Walter, 1996).

In New Zealand archaeology, the orthodox model of human colonisation has faced serial revisions as the radiocarbon chronology has been refined since the 1960s. The orthodox model proposed the initial settlement of New Zealand occurring around 800–1000 AD, with the loss of large game such as moa marking a transition between early and later occupation periods (Davidson, 1987). The revision of the radiocarbon chronology for New Zealand, spearheaded by Anderson (1991) resulted in a contraction of the period of human occupation as the time of first human settlement was pulled forward by centuries. Recently, Wilmshurst and colleagues (Wilmshurst et al., 2011) proposed that colonisation of East Polynesia took place when people arrived in the Society Islands around A.D. \( \sim 1025–1120 \), followed by a second pulse to all remaining islands, including New Zealand, between A.D. \( \sim 1190–1290 \). The impact of these revisions raises a need for a new way of looking at New Zealand colonisation, as it now seems that archaeological evidence of early settlement appears almost simultaneously across the landscape. The East Polynesian migrants are increasingly conceptualised as highly mobile, and the implications of the rapidity of the colonisation process are still being explored (Anderson and Smith, 1996, Jacomb et al., 2014).
Dogs accompanied at least some of these first groups of people to arrive in New Zealand, as dog bones have been excavated from early archaeological sites across New Zealand. Patterns of dog mtDNA diversity and relatedness therefore may be informative about the nature of the founding dog population, possible source populations, and the nature of the colonisation process and interactions between human communities.

Having established the current archaeological models for the emergence and spread of the LCC and the colonisation of New Zealand, I now consider how aDNA and molecular genetic analyses may be used to investigate animal introductions and dispersals during these processes. The focus of this section is on non-human mitochondrial DNA, as this thesis is concerned with the production and analysis of ancient mitogenomes from dogs, rather than people.

4.2 Molecular phylogenetics

Phylogenies, or models of genealogical history, are used in many branches of biology, and sub-disciplines are developing in areas such as ecology, anthropology, and linguistics which apply phylogenetics to their areas of interest (Yang and Rannala, 2012). Systematics and taxonomy use molecular phylogenetic analysis to establish the relationships between species by inferring their common history from selected DNA sequences, while population genetic analyses use similar methods to investigate the demographic histories of closely related species, populations and individuals within populations. Molecular phylogeography examines the spatial and temporal dimensions of genetic variation between and within populations. Molecular phylogenetic approaches have also been applied at a cellular and genetic level to investigate the development of cancers and the evolutionary and epidemiological dynamics of pathogens (Yang and Rannala, 2012). The basis of molecular phylogenetics is the production and analysis of DNA and RNA sequences. The following section provides a brief overview of basic concepts surrounding DNA, followed by a discussion of the main methods of analysis, and the use of aDNA. Finally, the application of aDNA methods to samples from archaeological sites is considered.
4.2.1 DNA

DNA holds genetic information that enables the construction and function of cells in every organism. It has two main roles, protein synthesis and self-replication, and it also regulates these processes (Matisoo-Smith and Horsburgh, 2012). Protein synthesis is the process of building proteins, which are necessary for cellular function. The majority of DNA is found in the nucleus of the cell, densely coiled together in pairs of chromosomes (Figure 4-1). In humans and other mammals one pair of chromosomes contains the sex chromosomes (X and Y), that determine the sex of the organism; females have two Xs, and males an X and a Y. One chromosome from each pair, including the sex chromosomes, is provided by each parent. DNA is also present outside the nucleus of the cell, in the mitochondria. These are organelles that have a range of functions including cellular energy production and the immune response. Each somatic cell contains the total set of DNA for that organism, called its genome. Where specially stated, the nuclear genome refers to the DNA held in the nucleus of the cell, while the mitochondrial genome describes the DNA in the mitochondria only.

Figure 4-1: The structure of an animal cell. DNA is found in the nucleus where it is bundled up in the chromosomes and in the mitochondria, which are present outside the nucleus (Courtesy: National Human Genome Research Institute).
While each cell contains only one set of nuclear DNA (nucDNA), outside the nucleus there are on average 2,000 mitochondria, all containing several copies of its genome (Matisoo-Smith and Horsburgh, 2012: 33) (Figure 4-2). The replication of mtDNA does not involve recombination, as occurs during reproduction involving the nuclear genome when genetic material from each parent contributes to their offspring. Mitochondrial DNA is maternally inherited, passed from a mother to her children, and in turn her daughters will pass their mtDNA to their children, however, their sons will not. This maternal line of descent makes mtDNA an important molecular marker for investigating maternal ancestry and genetic relationships. The high copy number of mtDNA in each cell has also meant it is an important source of DNA for aDNA studies (discussed in more detail below).

Figure 4-2: A mitochondrial organelle, which contains circular mtDNA molecules (Courtesy: National Human Genome Research Institute).
Nuclear and mtDNA have a different structure. The chromosomes, linked in pairs in the cell nucleus, have a linear form, joined in the middle (Figure 4-3). The chromosomes are made up of coiled DNA molecules. The nucDNA molecule looks like a twisted ladder, or a double helix, complete with stringers and rungs. The mtDNA molecule is circular, made up of a heavier outer strand and lighter inner strand (comparable to the stringers of the ladder) with rungs between them. The rungs of the DNA ‘ladders’ comprise four nucleotides or bases, usually abbreviated to the first letter of each name: adenine (A), cytosine (C), guanine (G), and thymine (T). Each rung in the ladder is made up of a pair of chemically bonded nucleotides (base pairs). The nature of these bonds means that A can only bind with T, and C with G. The stringers or sides of the ladder are known as DNA strands, so each DNA molecule has two strands. The chemical structure means that each strand is complementary, as the nucleotide on each strand will be complementary to the other in accordance with the nature of the bond. A DNA sequence is the series of letters (nucleotides) in the order it occurs along the DNA molecule.

Figure 4-3: The structure of nuclear DNA (Courtesy: National Human Genome Research Institute).
Genes are sections of DNA that hold the instructions for a particular protein or function. Genes are present in the chromosomes and the mitochondria. In addition to genes, there are sections of DNA which do not code for proteins. These stretches of DNA, known as non-coding regions, are not well understood (Shabalina and Spiridonov, 2004). Some parts may be involved in the maintenance of chromosome structure, or the regulation of protein production, but in some cases there is not an obvious function, or the function may no longer be necessary. In the mitochondria, the non-coding control region is also referred to as the displacement or D-loop, which as has two hypervariable regions (HVR1 and HVR2) (Matisoo-Smith and Horsburgh, 2012: 34).

Mutations or changes to the DNA sequence provide a means of identifying and investigating genetic variation between organisms, and through space and time. Mutations arise randomly in the genetic makeup of organisms. They can occur during DNA replication, when errors arise that are not corrected by cellular repair mechanisms. Mutation rates vary across the genome—the mutation rate of mammalian mitochondria is approximately 10 times higher than that of nucDNA (Brown et al., 1979). Mutations tend to accumulate in non-coding regions, as changes to parts of the molecule associated with the production of proteins have the potential to have major impacts on protein production and hence function.

The process of selection acts on these random changes that occur in genes. If a mutation is beneficial and increases the reproductive success of those individuals carrying it then it will become more frequent in a population, conversely a mutation with a negative effect will become less frequent. If changes severely affect protein production the organism may not survive, and hence the mutation would not be passed on to its offspring. A population may contain several versions, or polymorphisms, of a gene at the same time; the original unmodified gene, and any mutations. If a mutation occurs in a non-coding region of the genome, however, it may be passed on but may not affect the organism at all. Not all mutations however may be informative at a population or species level. Unique mutations, known as private or de novo mutations, arise in individuals during reproduction every generation, but may only be carried by that individual or their immediate family. In addition,
there are some parts of the genome, known as ‘hotspots’ that show evidence of rapid mutation, which are generally not used for phylogenetic analysis (Matisoo-Smith and Horsburgh, 2012: 34).

Mutations can take several different forms. Changes may occur in relation to individual nucleotides, where one base changes to another, e.g. C to T, an additional nucleotide pair made be added (an insertion) or a nucleotide pair may be lost altogether (a deletion). Changes to a single base are known as single nucleotide polymorphisms (SNP, called a ‘snip’), while the addition (insertion) or loss (deletion) of a nucleotide is often abbreviated as an ‘indel’. In population genetics, SNPs are identified where they occur in over one percent of the population, while private mutations (single nucleotide variations), occur in less than one percent. A single base is more like to mutate to its complementary pair, i.e. A<->G, C<->T, because of structural similarities. This is called a transition, whereas changes across these nucleotide pairs are called transversions (e.g. A<->C, G<->T). The variations in the frequency of nucleotide change is an important consideration when constructing phylogenies, and most of methods discussed below incorporate a specific model of evolution in the production of trees and networks. In addition, insertion mutations may affect multiple nucleotides (Matisoo-Smith and Horsburgh, 2012). These tend to accumulate rapidly in non-coding regions of the nuclear genome. Short tandem repeats (STRs) occur when small sequences of DNA (1–8 bp) are repeated during DNA replication, these repeats are sometimes called microsatellites. Longer stretches may also be repeated (9–100 bp), which are called minisatellites. Mutation rates of STRs tend to be higher than SNPs, and their high level of variability may be used for identifying individuals for forensic purposes. Sacks et al. (2013) utilised STRs and non-coding Y markers in their study of dingo and NGSD ancestry.

The process of recombination also has a major effect on the genetic composition of organisms (Lemey et al., 2009). In sexual reproduction, during the process of meiosis, chromosomal DNA inherited from both parents is shuffled prior to the formation of the egg or sperm. As result, individuals inherit a mix of genetic material from both parents. The mixed inheritance that results as an outcome of recombination contrasts with the uni-
parental inheritance of mitochondrial DNA and the non-coding region of the Y-chromosome. The differing patterns of inheritance are an important aspect when considering the results of phylogenetic analysis.

Processes such as mutation and recombination may result in genetic variation at a genomic level. In addition, genetic variation between organisms can arise from processes that operate on a population level, and these are considered next.

### 4.2.2 Phylogenetic and genealogical analyses

Evolutionary processes can be studied at different scales. Macro-evolutionary studies tend to be concerned with systematics, establishing the relationships between different species across sometimes extremely long time periods. Phylogenetic analyses can be carried out to reconstruct speciation events, and may be concerned with dating the divergence between species. Population genetic analyses on the other hand tend to focus on micro-evolutionary processes within species, investigating the outcomes of selection, migration, population size changes, and recombination on population genetic structure (Bloomquist et al., 2010). In keeping with this distinction, (Lemey et al., 2009: 16) points out that the term 'phylogeny' was originally used to describe evolutionary relationships between species, while 'genealogy' describes the patterns of shared ancestry within a population. When phylogeography was first described in the late 1980s by Avise and colleagues (1987), it was seen as a way of bridging the gap between these macro- and micro-evolutionary approaches, advocating the phylogenetic analysis of an organism within the context of its geographical distribution.

Improvements in sequencing technology and computational power since the 1990s have engendered increasingly complex approaches to the investigation of evolutionary and demographic processes. In addition, conceptual developments have contributed to the availability of a 'bewildering array' of statistical phylogeographic methods which enable statistical testing of hypotheses and estimation of demographic parameters such as population size, divergence time and migration rates (Knowles, 2004: 7). Traditionally, population genetic and phylogeographic analyses have been based on samples from a single point in time, or at most across several generations. The ability to sequence aDNA from
fossil, archaeological, and historical samples now enables studies of demographic history to incorporate data from different time periods. The development and application of statistical frameworks that incorporate multiple time frames are now being developed (Leonardi et al., 2016).

1.1.1.1 Generating DNA sequence data
Phylogenetic analyses were able to be commonly carried out using molecular genetic data following the development of the Polymerase Chain Reaction (PCR) and its application in Sanger sequencing in the 1980s. PCR involves the replication of a particular segment of DNA, using a DNA enzyme, two primers which are short synthesised DNA fragments that flank the area of interest on the strand of DNA, and additional nucleotides as building materials for new strands replicated from the template (Erlich et al., 1991). A cycle of heating and cooling causes the two strand to separate, the primers to bind to the DNA, and new complementary strands to be replicated by DNA polymerase using the additional nucleotides, resulting in the formation of another double strand. Under appropriate conditions, cycling can result in the exponential replication of the original segment (Erlich et al., 1991). Sanger sequencing incorporates fluorescently labelled nucleotides during the PCR process, which can be detected by a laser and the base identified. The output can be viewed as a chromatogram, with different coloured peaks for each base, which corresponds to the DNA sequence (Matisoo-Smith and Horsburgh, 2012). A Sanger sequence represents the composite picture of all the bases identified during sequencing. Sanger sequencing is a ‘first generation’ technology, and newer technologies are referred to as next generation sequencing (Metzker, 2010).

NGS has revolutionised the use of molecular genetic data in the biological sciences with the ability to produce extraordinary amounts of sequencing data relatively quickly and inexpensively. NGS technologies comprise a number of proprietary systems (e.g. Roche 454, Illumina, Life Technologies), with differing methods of sample preparation, and sequencing and imaging (Metzker, 2010). NGS technologies sequence a DNA ‘library’ made up of small segments of the DNA template or area of interest. Individual libraries can be tagged with a
‘barcode’ and sequenced together, making it possible to sequence large numbers of samples at the same time. Post-sequencing computational processing can then be used to segregate sequences relating to each library based on their barcode. For modern DNA, library preparation involves the shearing or fragmentation of the DNA into small segments, and the ligation of sequencing adaptors. During sequencing NGS technologies amplify (replicate) these individual library segments and often immobilise them to solid supports, enabling thousands to billions of sequencing reactions to take place simultaneously. NGS can also be used to target specific parts of the genome, using various techniques to enrich the library with DNA from the area of interest such as custom-designed microarrays, chips, and solution-based hybridisation methods (Metzker, 2010). The output from the sequencing platform contains raw data made up of all the ‘reads’ produced by the sequencing process, comprising the series of bases identified by the sequencing platform for each fragment of DNA. This differs from the output from Sanger sequencing, as NGS provides data on all the DNA fragments sequenced, while Sanger sequencing provides one file per sample that is a ‘majority rules’ consensus based on the data for each position. The NGS output file also includes quality scores describing the confidence with which each base has been determined.

A range of bioinformatic tools, usually complied in a bioinformatic computational protocol, known as a ‘pipeline’, are used to process the raw NGS reads and produce a final sequence, that can then be used for phylogenetic analyses. The basic steps carried out via the pipeline are: quality control, mapping of short reads to a reference genome or de novo assembly, post-alignment processing, and variant and SNP calling (Altmann et al., 2012). Additional processes are required for aDNA, and these are addressed in more detail in section 4.3.1.

Molecular markers
There is an increasing variety of molecular markers used in phylogenetic and phylogeographic studies. For example, a quick review of the literature on dog domestication, evolution and phylogeography show the use of fragments of the hypervariable portion of the mtDNA control region, microsatellites, single or multi-loci mitochondrial or nuclear genes, concatenated series of genes, or even complete genomes (Brito and Edwards, 2009,
McCormack et al., 2013). Marker selection depends on the question at hand, and the technology and resources available. MtDNA has been used to study evolutionary relationships between and within species, because of its simple genetic structure, maternal inheritance, lack of recombination, and reduced effective population size (Rubinoff and Holland, 2005). Because of mtDNA’s faster rate of evolution than nucDNA (Brown et al., 1979) the number of variable sites will generally be greater for mtDNA than nucDNA, making mtDNA particularly useful for population level analyses of between and within population variation. In addition, because of its smaller effective population size, mtDNA completes lineage sorting faster than nucDNA, making it useful for the more recent past (Fahey et al., 2014). However, because mtDNA is non-recombining and inherited as a single unit, selection acting on any part of the mitogenome, or nucDNA that interacts with it, will have an impact on demographic history, which will show only the most recent selective sweep obscuring any previous demographic history (Fahey et al., 2014).

As mtDNA is maternally inherited, it does not represent the entire bi-parental genetic inheritance contained in nucDNA, while Y-chromosome markers also show a male line of descent only. Autosomal genetic markers provide data on genetic diversity from both parents, although substitution rates for nucDNA are more variable than mtDNA so divergence estimates may have much wider confidence intervals (Fahey et al., 2014). Increasingly both sets of markers are preferred of demographic studies, and disagreement between markers can enhance the inferences possible about demographic history. Incongruence between different markers may highlight the outcomes of introgression, complex population structure and sex-based gene flow (Rubinoff and Holland, 2005). Along with the advent of NGS sequencing technologies, the use of multiple rather than single data sets require increasingly sophisticated statistical treatment to evaluate resulting inferences.

4.2.3 Graphic representations and statistics

Inferences about evolutionary relationships between taxa are often shown graphically as trees or networks. In addition, aspects of demographic history may be estimated using
statistical methods that compare observed sequence variation with what would be expected under a neutral model of evolution (Rubinoff and Holland, 2005).

Trees
Traditionally, evolutionary relationships have been shown as phylogenetic ‘trees’, with terms for the different parts of the tree, such as root, branch and leaf corresponding to this metaphor. The branching pattern of the tree is called its topology. The leaves or tips are external nodes and represent characteristics of the samples, such as DNA sequences or other data. Depending on the nature of the study, branch tips may represent species, or haplogroups within species (inherited patterns of SNPs that discriminate groups within a population). In intraspecific phylogenies, haplogroups comprise groups of haplotypes that share a common ancestor. The internal nodes represent putative ancestral sequences. In an unrooted tree, the individual sequences are shown relative to each other without the direction of evolution being shown (Lemey et al., 2009: 20). A rooted tree includes distantly related taxa which enables the most common recent ancestor of the taxa under study to be inferred, hence also showing the direction of evolution (usually from left to right). The branch length (horizontal) measures the amount of change, usually expressed in units relating to genetic distance, such as the numbers of substitutions per site (Yang and Rannala, 2012). Estimations of divergence time draws on the ‘molecular clock’ hypothesis that assumes that mutations occur at a constant rate over time (Lemey et al., 2009: 23). However, violations to this assumption may occur due to differing metabolic rates, generation times, population sizes and selective pressure. This has resulted in the development of new models to allow for variation in mutation rates. Numbers adjacent to branches can show the level of statistical support for the node, from methods such as bootstrapping or Bayesian posterior probabilities (see below).

Some phylogenetic trees may be more accurately called ‘gene trees’, as they represent the inferred history of the gene used for the analysis, rather than the species. Further, individual gene trees and species tree may not be congruent (Yang and Rannala, 2012). This incongruence may arise as the divergence time between genes may not correspond to the
divergence time of species, due to a range of possible factors, including effective population size and differences in copy number (Lemey et al., 2009: 23).

Phylogenetic trees can be inferred from the data—they are not directly observed—and there are many methods available to construct phylogenetic trees, with associated software (see Yang and Rannala, 2012 for a review). Regardless of which marker is used for analysis, in order to trace changes to DNA sequences and hence evolutionary history, sequences need to be aligned or matched against each other, so that the same positions are being compared with each other. The alignment of sequences can be relatively easily done where sequences are closely related and there are few changes in the overall structure, but where sequences have diverged over a long period of time, alignment can be more difficult. There is an increasing range of computer programs that produce sequence alignments, particularly for NGS outputs (see Li and Homer, 2010 for an overview).

Tree building methods can be grouped by the kind of data they use, and the algorithmic approach of the method (Lemey et al, 2009, Yang and Rannala, 2012). Distance matrix methods, such as neighbour-joining, calculate the distance between pairs of sequences to produce a pair-wise distance matrix, and phylogenetic relationships are then inferred from this matrix. This method works from the data up, grouping clusters together to ultimately generate one tree. In character-state methods, such as maximum parsimony, maximum likelihood and Bayesian methods, the position in an aligned sequence is the character, and the nucleotide or base (A,C,G,T) is the state. The analysis takes into account all changes to character positions independently and calculates a score for each tree. Distance matrix, maximum likelihood and Bayesian methods all incorporate a substitution model that describes the substitution rates for each nucleotide. Different models can be used that may assume an equal rate of evolution, or allow for different rates between transversions and transitions (Yang and Rannala, 2012). Models can also allow for different frequencies of nucleotides, and variable rates between sites or sequences. Software such as ModelTest (Posada and Crandall, 1998) can be used to determine the most appropriate substitution model of analysis.
With character-state methods, a comparison of all possible trees should show the ‘best’ tree, being the tree with the highest score. The value of the score varies depending on the method used. Maximum parsimony parallels the concept of Occam’s Razor and assumes that the tree that requires the least number of changes or mutations is the best (Matisoo-Smith and Horsburgh, 2012). The score for maximum likelihood is the log-likelihood value, and for Bayesian inference it is the posterior probability (Yang and Rannala, 2012). For large datasets it is not computationally possible to compare all the possible trees generated by character-state methods, so heuristic tree search algorithms are used to find the tree with the best score (Yang and Rannala, 2012). Parsimony and likelihood methods may also utilise statistical measures of support to assess the reliability of a particular tree (Matisoo-Smith and Horsburgh, 2012: 54). Bootstrapping uses resampling to test the reliability of the topology, and bootstrap values of over 95% are often viewed as providing strong support. Bayesian methods are based on the Bayesian method of statistical inference, and enable the incorporation of a priori information about the tree or its parameters, in addition to the sequence data, the posterior probability or tree score is based on the outcome of the analysis using both prior parameters and data. Bayesian methods differ from maximum-likelihood methods, as the parameters in the Bayesian model are viewed as random variables with statistical distributions, while in maximum likelihood they are unknown fixed constants (Yang and Rannala, 2012). Bayesian methods offer the opportunity to include parameters such as dates in the generation of trees, and for a relaxation of the molecular clock (Drummond et al., 2006).

The various methods have different strengths and weakness in comparison with each other, relating to, for example, computational processing speed, accuracy, impact of gaps where nucleotide data is unavailable in sequence alignments, and model selection (for a review see Yang and Rannala, 2012 Table 2). Som (2013) assessed the performance of five different methods (Bayesian, maximum likelihood (ML), neighbour joining, maximum composite likelihood, and maximum parsimony) on the basis of accuracy, consistency and computational efficiency, using computer simulated datasets which included mitochondrial genes. Results showed that Bayesian and ML methods were far more accurate than the rest,
and that ML was the most consistent. Although the rest were quick to perform, their accuracy and consistent was poor compared to Bayesian and ML methods. A comparison of the typologies generated using different methods such as likelihood and Bayesian approaches can provide a qualitative measure of the robustness of the topologies produced for a particular dataset. A greater confidence in the results can be assumed where similar topologies are generated by different methods (Rubinoff and Holland, 2005).

Networks

The production of phylogenetic trees assumes that the evolutionary process is bifurcating, that is, branches on a tree will only split in two and that these branches or lineages will not interact (Lemey et al., 2009: 631). There are a range of circumstances where a bifurcating tree may not best represent evolutionary processes, for example, where explosive radiation results in multifurcation of a branch, if ancestral and descendant genes co-exist, where there has been recombination, hybridisation or horizontal gene transfer, or, perhaps most commonly, where there is insufficient data to discriminate between sequences. A network provides an alternative representation of evolutionary relationships, where vertices correspond to nodes, and the edges to the branches on a tree. Networks do not force the data into a tree-like shape and also allow for cycles, where paths can begin and end at the same node (Lemey et al., 2009: 632, Morrison, 2005). For intraspecific phylogenies, networks may better represent population-level processes and enable a more detailed display of information than is possible on a bifurcating tree (Posada and Crandall, 2001).

There are a number of methods for network construction, with associated software packages, including statistical parsimony, median-joining, and minimum spanning networks. Woolley and colleagues (2008) used simulations to evaluate the performance of eight network methods in comparison to standard tree building approaches. They found that all methods, with the exception of minimum spanning, produced the correct topology and branch lengths most of the time when the substitution rate was low and there was no recombination. Maximum parsimony and union of maximum parsimony trees were the most accurate where substitution rates were higher. Where recombination was present, no method could
accurately estimate branch lengths and the successful rate of inference of the correct topology was halved.

**Summary statistics**

The size and growth rate of populations over time may be investigated by statistical tests on population genetic data. The observed sequence variation is often compared with expectations under a neutral model. A neutral model describes the equilibrium between the loss of genetic variation through random drift and the accumulation of diversity through mutation, expected to be achieved in a stable population in the absence of natural selection and migration (Grant, 2015).

The main parameter calculated to estimate population size from genetic markers, as it relates to the genetic variability of a given population, is the effective population size \((N_e)\). This parameter differs from the census population size. Wang (2005: 1395) observes that it is ‘notoriously difficult to estimate, mainly because of the highly stochastic nature of the processes of inbreeding and genetic drift for which \(N_e\) is usually defined and measured, and because of the many factors (such as time and spatial scales, systematic forces) confounding such processes’. These systematic forces include selection, mutation, and migration.

Fahey and colleagues (2014) identify three kinds of tests that may reveal changes in population size over time. 1) Tests such as Tajima’s \(D\), \(D^*\) and \(F^*\), that use mutation frequencies and the excess of rare mutations; 2) Statistics such as \(F_S\) that use composite haplotypes; and 3) Those that use the distribution of pairwise sequence differences to test for deviations from the null model. In addition, mismatch distribution and the Bayesian skyline plot are methods that can detect population expansion and decline, and also provide information about the timing of these events, and historic population sizes (Fahey et al., 2014). Tests of neutrality may be performed to identify statistically significant population shifts, which can then be investigated using mismatch and Bayesian analyses.

The next section of this chapter examines the use of aDNA in molecular genetic analysis.
4.3 Ancient DNA

The term ‘ancient DNA’ is applied to DNA from any degraded sample or where the DNA sample was not recently collected from a living organism and preserved (Matisoo-Smith and Horsburgh, 2012: 60). Ancient DNA may be obtained from numerous sources, including bones, teeth, skin and hair, from archaeological specimens, museum collections and natural fossil deposits. Ancient DNA differs from DNA found in living organisms, as once the organism dies the DNA begins to decay, becoming fragmented and subject to physical and chemical changes. As a result, aDNA fragments are often only 100 bp long (Heintzman et al., 2015). Decay can result in breaks to the DNA molecule, either single or double stranded, and the formation of miscoding or blocking lesions where nucleotides are missing or damaged. These lesions either prevent amplification for sequencing or result in an incorrect sequence (Heintzman et al., 2015).

The production and analysis of aDNA sequences generally follows that for modern DNA but must deal with a range of challenges not present in modern samples, arising from the degraded nature of aDNA. These challenges include the extraction, sequencing and assembly of very short fragments of aDNA, the avoidance of contamination from exogenous DNA and modern DNA present in laboratory environments, and possible sequencing or processing errors resulting from the degraded state of the ancient fragments (Marciniak et al., 2015).

4.3.1 Extraction and sequencing

Experimental work involving aDNA is subject to a range of stringent protocols and standards to ensure the authenticity of results (Willerslev and Cooper, 2005). The risk of contamination is present from the time a sample is excavated to when it is prepared for sequencing (Linderholm, 2016: 153). Physical measures are taken during the extraction and pre-PCR steps of library preparation include working in a dedicated aDNA laboratory, wearing protective clothing and following protocols to keep levels of environmental DNA low (Knapp et al., 2012a).
Initially aDNA research used a combination of PCR and Sanger sequencing to generate ancient sequences, but as with molecular genetics generally, the development of NGS has revolutionised the sequencing and analysis of aDNA (Knapp and Hofreiter, 2010, Marciniak et al., 2015). If a NGS platform is to be used for sequencing, then aDNA libraries need to be prepared. These libraries contain the target DNA fragments from the extract, ligated to oligonucleotides that may include a barcode or index to identify each individual library, and adaptors that have a functional role during the sequencing process. Ancient DNA libraries may be enriched by methods such in-solution hybridisation capture to increase the amount of target DNA in the library (Der Sarkissian et al., 2015, Linderholm, 2016). In addition to the reduced costs, speed and increased amount of data generated, NGS is also amenable to the short fragment length of aDNA, and enables better identification of possible contamination.

4.3.2 Assembly

As with modern DNA, the NGS aDNA reads are assembled using a series of bioinformatic tools compiled in a pipeline. The reads need to be prepared for mapping, mapped to a reference sequence and pass quality control checks. Generating sequences from aDNA libraries involves several additional bioinformatic steps that are not usually carried out for modern samples, in order to maximise the information able to be obtained, whilst ensuring the validity of the results (Der Sarkissian et al., 2015). Heintzman and colleagues (2015: 251) outline a typical aDNA pipeline for an ancient genome as follows: merging paired-end reads, trimming adaptor sequences, removing low complexity and short reads, and separating individual samples based on the barcode or index. These filtering steps are following by mapping reads to the reference sequence and removing duplicates. Finally, quality control steps are carried out calculate the average coverage across the genome and estimating rates of aDNA damage and contamination. Studies have also identified where the parameters of bioinformatic tools can be optimised in relation to the characteristics of aDNA (Leonardi et al., 2016).
As NGS data provides information for each fragment of DNA, descriptive statistics can be generated and compared with the expected characteristics of aDNA to evaluate authenticity of reads. Criteria include fragment length and damage patterns that are characteristic of aDNA. A better understanding of the processes of degradation, such as the deamination of cytosine bases at the end of the aDNA molecule, enables distinctions to be made between the aDNA target, exogenous and contaminant sequences (Hofreiter et al., 2015, Linderholm, 2016) Some of the original criteria used to assess the authenticity of aDNA, such as replication, have become obsolete with the use of NGS, particularly given the large number of reads produced and the ability to examine damage patterns (Linderholm, 2016).

4.3.3 Analysis

Ancient DNA data have been generated for phylogenetic and phylogeographic studies of extinct and extant animal species, as well as functional studies of genes and the process of domestication. Population genetic studies have also been undertaken but to a lesser extent (for reviews see Hofreiter et al., 2015, Leonardi et al., 2016). Rather than inferring changes in genetic diversity and population size indirectly from modern populations, aDNA offers the opportunity to sample populations before, during and after significant events in a species’ history thus measuring genetic shifts directly (Hofreiter et al., 2015: 290)

Until recently, the time frame feasible for aDNA studies was the last 40,000 years. Studies often addressed questions about major changes in biodiversity, for example the extinction of megafauna at the end of the Pleistocene, or the impact of key human transitions such the shift from foraging to farming in the Neolithic (Wang et al., 2014, Boivin et al., 2016). These studies usually sequenced portions of the mtDNA control region, because of the substantially greater relative proportion of mtDNA in each cell and its non-recombining nature (Matisoo-Smith and Horsburgh, 2012). The use of mtDNA for modern phylogeographic studies also meant that comparative data sets were available for human populations and many domesticated animal species. Small fragments of mtDNA formed the basis of most studies, however, in some cases where DNA preservation was excellent, extraction of nucDNA was possible (Hofreiter et al., 2001).
The development of NGS has transformed aDNA studies, enabling the sequencing of mitogenomes, portions of nucDNA and even whole genomes for several species, over greater time depths than previously possible (for a review see Leonardi et al., 2016). Within the last few years the time range possible has been significantly pushed back with the publication of a complete mitochondrial genome of a ~400,000 cave bear (Dabney et al., 2013) and a draft genome from a 560,000–780,000 year old horse bone (Orlando et al., 2011). Alongside developments in aDNA methods, the expanding range of genetic markers and genomes being sequenced has made the application of increasingly sophisticated statistical methods for investigating past population demography feasible. Leonardi and colleagues (2016) recently reviewed the application of coalescent theory to aDNA datasets, as well as methods for detecting and quantifying population structure and admixture using whole genomes.

Ancient DNA from archaeological specimens has also been incorporated into studies based on the commensal model, which uses phylogeographies of domestic and commensal animal species as proxies for tracing human migrations. This approach was pioneered in the Pacific region, where analysis of modern and ancient specimens of the Pacific rat were used to evaluate hypotheses derived from a range of scenarios for the human settlement history of the region (reviewed in Storey et al., 2013).

4.4 Molecular genetic inference from archaeological assemblages

Samples from archaeological and fossil assemblages often form the basis of phylogenetic studies over long time frames and large geographical areas (e.g. Knapp et al., 2009, Allentoft et al., 2014, Librado et al., 2016). In such cases the specimens are sufficiently separated though time and space to be considered to be representative of different species, sub-species or populations, for the purpose of modelling and statistical analysis. But when archaeological samples are used for phylogeography and population genetic analyses of taxa more closely related in terms of space and time, assumptions about the population may not necessarily align with the nature of archaeological assemblages from which the sample was obtained. In the following section, I address the concept of a ‘population’ in biology, compare this with
what constitutes an archaeological assemblage, and then highlight possible inconsistencies between the two that may need to be taken into account when they are combined.

4.4.1 Natural, census and hypothetical populations

The term ‘population’ is broadly defined as a group of individuals of the same species that occupies a particular area (Allaby, 2009: 503). In this context, a population is an objective entity, observable and measurable. Although researchers may not be able to take a total census of a natural population, strategies are designed to estimate census population size using methods such as aerial counts and mark-recapture. Whilst there is general agreement about population referring to a group of individuals from the same species, Camus and Lima (2002) highlight the variability that can occur when describing the spatial extent of a given population.

In addition to the objective description of a group of individuals, the term is also used in a hypothetical context, as part of statistical modelling and hypothesis testing (Martinez-Abrain, 2010). Statistical modelling of populations may assume various characteristics such that the members of an ideal population comprise equal numbers of males and females, which mate randomly, and are not intergenerational. Discussions of the meaning and use of the concept of a population within the biological and life sciences have therefore stressed the importance of researchers being explicit and precise in their use of the term (Camus and Lima, 2002).

4.4.2 Archaeological assemblages

An archaeological assemblage is defined in relation to an archaeological context, that is, an assemblage is the material that has been excavated from a particular location. In the absence of this contextual information, an assemblage has little or no analytical archaeological value. The context will vary depending on the nature of the excavation, the research questions being addressed, and the resulting methodology. Archaeological contexts are usually defined in relation to a spatial co-ordinate system and a vertical scale. In post-excavation analysis this vertical scale is usually correlated to a site-based chronology, and may also be dated using
methods such as radiocarbon dating. An assemblage may comprise all the archaeological material that has been excavated, or may be broken down into its constituent parts, such as artefacts manufactured from different materials and fauna. In the Pacific this is often stone, pottery, shell and bone artefacts, and different classes of fauna—mammals, fish, bird, shellfish etc. The composition is shaped by the practices at the place during deposition, subsequent taphonomic processes, and the archaeological research process.

Archaeozoology is concerned with the interactions between people and animals in the past, so a substantial amount of research relates to investigating how these interactions can be reconstructed from animal remains, including the cultural practices that led to the deposition of the bones in archaeological sites. It is beyond the scope of this thesis to consider these methods, but it is important to recognise that a large component of faunal assemblages represent the outcomes of choices made by people as to the animals that they caught and processed to consume as food, to use for raw materials, or for symbolic and ritual purposes. Faunal assemblages may therefore be patterned by these different practices, and so are unlikely to be representative of a natural population (Grayson, 1984). Indeed, some archaeozoologists have used comparisons between natural modern populations and archaeological faunal assemblages to identify cultural practices such as hunting and herding (Lyman, 1987). Kewsani (1994) has also drawn a distinction between societies where livestock production is geared towards state or market economies, or where group subsistence relies on the maintenance of a particular herd; with those that do not rely on particular domesticated animals as a staple (for either meat or secondary products). She argues that specialised management strategies are likely to be found in the former (e.g. culling of particular sex/age classes) while in the latter patterns of animal exploitation and consumption are more strongly influenced by social considerations and ritual practices that utilitarian ‘optimising’ models do not take into account.

The behavioural ecology of the animal species may also play a part in structuring the assemblage. A population may not necessarily comprise equal numbers of sexes or age classes, whether in a natural or domesticated population. As discussed in Chapter 3,
ecological studies of free-ranging dog populations have shown in some instances there are differences in sex ratios in some groups. In addition, high juvenile mortality rates may also affect population structure, and have a flow-on effect for archaeological assemblages.

A further factor affecting the composition of archaeological faunal assemblages are the processes that take place between the death of the animal and its excavation; these may be cultural or natural, or a combination of both. The focus of taphonomic studies in archaeology is to understand the impact of such processes on archaeological assemblages (Lyman, 1994). Taphonomic processes include natural processes such as weathering, etc., while cultural processes relate to food and artefact processing and deposition. Taphonomic processes that include alternation to chemical and structural properties of bone, such as weathering and burning, are likely to have an impact on the survival of DNA (Ottoni et al., 2009).

The assemblage that is curated at the conclusion of an archaeological research project represents the end point of a process that involves a series of sub-sampling exercises (Grayson, 1984). Archaeological sites are rarely excavated in their entirety, so any assemblage is highly likely to be a subset of the fauna deposited in the site, throughout its occupation. In some sites, faunal material may have been deposited in concentrated areas such as midden dumps, or rubbish heaps, so sampling strategies may target these areas depending on the research questions. In these midden contexts, animal bones are usually disarticulated and jumbled together rather than being in discrete burials. Sampling different individuals from these contexts may not be possible. While faunal analysis is routinely included in reporting of archaeological excavations in the Pacific, this may be subsidiary to the main purpose of the excavation, so large amounts of material may not be excavated. Conversely taphonomic processes may mean that there is very little faunal material present, despite a research interest in this aspect of the past.
4.5 Old dogs and new tricks

The application of molecular genetic techniques to animal bones from archaeological sites has the potential to greatly enhance our reconstructions of the past, by offering an additional method of analysis to complement archaeozoological methods, and providing a time depth previously unavailable to molecular studies of contemporary populations. Samples from archaeological assemblages, however, are not equivalent to samples from a natural population. Archaeological assemblages may differ from a natural population because:

- The assemblage is highly likely to contain a subset of the animal bones deposited in the archaeological site
- The assemblage may be the result of cultural practices that bias particular age or sex classes
- Individuals may not be able to be differentiated in an assemblage, and may be represented by multiple bones that could be sampled more than once
- The assemblage may span a time period of hundreds of years.

As long as models, statistical analyses and interpretations can take into account these possible deviations from natural or hypothetical populations and their impact on the analysis, molecular genetic analysis can offer new sources of information about human and animal history. Possible avenues of research include selection, migration, contact between groups, loss and reintroduction of species, continuity of species, and processes of introduction and dispersal.

In the next three chapters, I investigate this potential by generating and analysing dog mitogenomes in relation to two periods in the history of the colonisation and settlement of the Pacific. Firstly, the colonisation of New Zealand that took place at the end of the last major pre-industrial human migration, and secondly, the appearance of the LCC in the Pacific which begins around 3,500 years ago.
Chapter 5
Complete Mitochondrial Genomes of New Zealand's First Dogs

Preface

This chapter (with some modifications) has been published in PLoS one as follows: Greig, K., Boocock, J., Prost, S., Horsburgh, K. A., Jacomb, C., Walter, R. & Matisoo-Smith, E. 2015. Complete Mitochondrial Genomes of New Zealand’s First Dogs. PLoS one, 10 (10), e0138536. Supplementary information is provided in Appendix A.

5.1 Introduction

Dogs are found in human communities throughout the world where they may be companions, working dogs or simply co-inhabitants of villages. They have a special place in human history as the first domesticated animal, having appeared in the archaeological record around 15,000 years ago, well in advance of other domesticates (Larson et al., 2012). Since this time they have successfully moved with people across the globe, including through the islands of Oceania. Recent trends in dog keeping, particularly the rise in intensive breeding for particular physical characteristics, lie as a veneer over a much longer and more complex relationship between dogs and people.

Despite their long association with people, dogs appear relatively late in the archaeological record of Oceania (Figure 5-1). Near Oceania includes the large island of New Guinea, the Bismarck Archipelago, and the Solomon Islands, and has evidence of human occupation that dates to as far back as the late Pleistocene (ca. 40,000 years ago). The antiquity of human occupation in Near Oceania contrasts with the relatively recent human colonisation of the islands of Remote Oceania to the north, east and southwest, which were not settled until about 3,000 years ago (Kirch, 2000). In addition to demarking two phases in the settlement of the region, the distinction also relates to differences in biogeography. The continental islands of Near Oceania possess the greatest biodiversity within the region. Further east
across the Pacific the islands become smaller and more dispersed, and there is an accompanying drop in biodiversity (Green, 1991).

Figure 5-1: Map of the Pacific region. The boundary between Near and Remote Oceania is shown by the dotted line. Arrows indicate the proposed route for the introduction of dogs to Polynesia, via Island South East Asia. Pie charts and the summary table show the relative proportions of Arc1, Arc2 and A29 haplotypes observed in samples from across the region (after Oskarsson et al., 2012).

The oldest evidence of dogs in Oceania comes from Australia, where the earliest dog (dingo) remains date from 3,500 years ago (Gollan, 1984). The dingo has subsequently adapted to surviving independently of people and is found in many areas of rural Australia. Although the dingo is considered a feral dog, individuals may also be taken in by indigenous communities where they become companions, protectors, and spiritual guardians, and can assist with hunting (Smith and Litchfield, 2009). In New Guinea, the NGSD is thought to be a descendant of another relatively early introduction, although no well dated or genetically confirmed archaeological NGSD remains have been reported. They are rarely observed in the wild and the current captive population is descended from only eight individuals (Koler-
Matznick et al., 2007). In Australia, New Guinea, and the islands of Near Oceania, dogs were introduced to areas where people had already been living for thousands of years.

The introduction of dogs into Remote Oceania is associated with the migration of people to these previously uninhabited islands, which began about 3,200 years ago (Kirch, 2000). This process is marked by the appearance of a distinctive pottery, termed ‘Lapita’, which is present in numerous archaeological sites that extend from Papua New Guinea to the western margins of Polynesia. The presence of Lapita pottery has been linked to the arrival of people originating from Island South East Asia, who spoke Austronesian languages. In addition to pottery and languages, these people introduced agricultural practices, a suite of domesticated animals and plants, a settlement pattern of villages situated on intertidal reefs or small offshore islands, and a characteristic set of artefacts including adzes and shell ornaments (Specht et al., 2014). Dogs are generally considered to be one of the Lapita domesticated animals, along with pigs and chickens. To date though, Lapita dog remains are relatively rare, having only been recorded in limited numbers in archaeological sites in the Bismarck Archipelago of New Guinea (Table 10.1 in Matisoo-Smith, 2007) and have not been reported in any of the key archaeological sites in Vanuatu or New Caledonia.

The subsequent, post-Lapita settlement of the islands further to the east also involved dogs, pigs and chickens, although they were not uniformly distributed across the region. The settlement of Polynesia occurred in two phases, beginning with the islands of West Polynesia around 3,000 years ago during the Lapita expansion. Then after a substantial hiatus, East Polynesia was rapidly settled over several hundred years following the arrival of people initially to the central tropical islands about 1,000 years ago (Wilmshurst et al., 2011). This second settlement phase included the islands at the margins of East Polynesia: Hawaii, Rapa Nui and New Zealand. In some of these island groups, such as Hawaii and New Zealand, the large number of dog bones found in early archaeological sites contrast strongly with the limited evidence of dog remains in the Lapita archaeological sites further west.

By the time European explorers arrived in Polynesia in the eighteenth century, dogs were present on some but not all island groups. In East Polynesia, dogs were often an important
part of the social and economic fabric of daily life (Luomala, 1960b). In the ranked societies of Hawaii dogs were observed as the property of chiefs and were raised in large numbers for feasts. In New Zealand, dogs were the only domesticated animal to be successfully introduced by the Polynesian ancestors of the Maori. While living, they were kept as watchdogs, hunting dogs and general companions and were sometimes also kept for their hair; on their death they could be used as food for ceremonial occasions, their bones and teeth as industrial materials, and their pelts to make dog skin cloaks (Davidson, 1987).

Previous studies using mtDNA to investigate the origins and dispersals of the dog throughout the Pacific indicate that the dingo, NGSD and Polynesian dogs are all descended from East Asian dogs (Figure 5-1). Using a 582 base pair (bp) fragment of the control region of the mitogenome from modern dogs, Savolainen and colleagues (2004) demonstrated that all dingoes sampled belonged to the A29 haplotype. Although this is one of a number of dog mtDNA lineages that reached ISEA it was the only one successfully established in Australia. On this basis, they suggest that the dingo population was either founded by a very small number of individuals, or from a group of dogs that had passed through a series of genetic bottlenecks and hence had lost substantial genetic variation. Savolainen and colleagues (2004) also reported that the 19 ancient Polynesian dog sequences, based on analysis of a shorter 263 bp control region fragment, belonged to two short haplotypes, Arc1 and Arc2, both different from the A29 control region haplotype carried by dingoes. The short Arc1 haplotype is indistinguishable from a number of widespread control region haplotypes, while Arc2 appears to belong to the A75 lineage found in modern Indonesian dogs.

A later study by Oskarsson and colleagues (2012) investigated the origins and routes of the introductions of dingoes, NGSDs, and Polynesian dogs in further detail, using the same 582 bp and 263 bp Arc1 and Arc2 control region fragments. The study used samples collected from modern dogs and dingoes, previously published modern dog and dingo sequences, and the ancient Polynesian short sequences. Oskarsson and colleagues (2012) compared the frequency and distribution of modern haplotypes in Mainland and ISEA, including Taiwan and the Philippines, with those from the dingo and ancient Polynesian samples (Figure 5-1).
Arc1 and Arc2 short haplotypes were found in 10% of modern dogs samples in South China, 16% in Mainland South East Asian dogs, and 42% in Indonesian dogs. The dingo haplotype A29, was found in 2% of modern dogs sampled in South China, 1% of dogs in Mainland South East Asia, and 8% of Indonesian dogs. The short Arc1 haplotype appears to have a predominantly mainland and western Island South East Asia distribution, while Arc2 is found in increasing frequency across Island South East Asia. Additionally, they found that the short Arc1 haplotype corresponded to a possible 13 control region haplotypes, while Arc2 was limited to a possible two, A75 and A120. Neither the ancient Polynesian short haplotypes nor the A29 haplotype carried by dingoes, were present in dogs from Taiwan and the Philippines, suggesting that dogs were not introduced into the Pacific region from or via this north-eastern route.

Following recent advances in sequencing technology we can now use NGS to obtain complete ancient mitogenomes. As would be expected, recently published studies of complete mitogenomes of cattle (Horsburgh et al., 2013) and sheep (Lancioni et al., 2013) have shown that the control region sequences provide only a partial picture of genetic diversity. This has also been the case in a recent study of mitogenomes from four human burials at Wairau Bar, which demonstrated that the people buried there possessed greater than expected mtDNA diversity (Knapp et al., 2012b). The ability to observe a more fine-grained level of diversity can be particularly significant when addressing founding population structure, particularly in a geographic area with a settlement history such as that in the Pacific, which most likely involved population bottlenecks and a relatively short chronology of human occupation.

Here we describe complete mitogenome sequences for fourteen dogs from the site of Wairau Bar, one of the earliest archaeological sites in New Zealand and probably a central place during the colonisation phase (Jacomb et al., 2014). Wairau Bar comprises a settlement with middens, living floors, specialist work zones, food preparation areas, and a burial ground. The site also contains numerous bones of moa, other extinct birds, and a suite of artefact types that link New Zealand with the tropical East Polynesian homeland zone. Because
Wairau Bar is among the earliest settlement sites in New Zealand, it is a critical place for understanding the colonisation processes of the archipelago including the successful introduction of dogs to the southern-most point of their Oceanic distribution.

5.2 Materials and methods

5.2.1 Samples

Dog teeth were selected from faunal material recovered from a large cooking feature (Oven Pit 1) during archaeological excavations in 2009 (Jacomb et al., 2014) (Figure 5-2, Appendix A). These excavations were fully authorised by the appropriate government agencies (NZ Historic Places Trust Authority No. 2009/121), and the excavated faunal collections are held at the Otago Archaeological Laboratories, University of Otago, Dunedin, New Zealand (Jacomb et al., 2014). The oven fill contained bone and shell midden, charcoal, and fire-cracked rocks. In addition to dog, fauna identified from the oven included the extinct moa (bone and eggshell) and at least 60 species of other birds, marine mammals, fish and shellfish. Moa eggshell from Oven Pit 1 was used to develop a detailed chronology drawing on aDNA, and Bayesian analysis of radiocarbon dates, which showed the oven had been filled with midden during a single event that occurred sometime after the early 1320s and no later than about 1350 (Jacomb et al., 2014). The rapid deposition of the oven fill indicates that the dogs were all likely to have been a part of the same dog population living at Wairau Bar immediately prior to their death and incorporation in the midden. Over 400 dog bones, teeth and bone fragments were excavated from the oven, which represents at least 21 dogs. The large number of dog bones excavated from the site allowed 16 specimens of the same element (left maxillary fourth premolar) to be sampled for inclusion in this study. This means that each of the 16 specimens represents a different individual, and ensures that no dog was sampled more than once.
Figure 5-2: Wairau Bar archaeological site. Location of the archaeological site (top left), location of Oven Pit 1 (top right), and section view of Oven Pit 1 (horizontal scale 2 m, vertical scale 0.5 m).
5.2.2 Extraction, hybridisation capture and sequencing

All DNA extraction and sequencing library preparation before PCR amplification was carried out at the Ancient DNA Laboratory at the University of Otago, where stringent procedures are in place to avoid contamination (Knapp et al., 2012a). DNA extraction was performed on sixteen teeth. To reduce surface contamination, the teeth were submerged in 6% bleach for ten minutes and rinsed in ultra-pure water prior to grinding for DNA extraction. Teeth were ground using a mortar and pestle and 150 to 250 mg was used for extraction following a silica based extraction protocol (Rohland and Hofreiter, 2007a). At least one extraction blank was processed in parallel with each set of five samples. Extraction blanks were then treated as samples throughout processing and sequencing.

Barcoded sequencing libraries were prepared directly from the DNA extract as described by Knapp and colleagues (Knapp et al., 2012c) for Illumina sequencing, with the following modifications. Quantitative PCR (Stratagene MxPro 3000P) was used to determine the number of cycles necessary to reach amplification plateau. Libraries were immortalised by PCR amplification to plateau using ABI’s AmpliTaq Gold with the following reagent concentrations: 1x AmpliTaq PCR Buffer, 2.5 mM MgCl2, 1 mM dNTPs, 0.2 μM of each extension primer and 3.75 units of AmpliTaq Gold. Amplifications were stopped when PCR plateau was reached to avoid the manufacture of chimeric molecules (Meyerhans et al., 1990, Odelberg et al., 1995, Judo et al., 1998, Thompson et al., 2002). Immortalised libraries were purified using MinElute columns (Qiagen) following the manufacturer’s protocol with the addition of a second PE wash and a 5 minute incubation with 0.1x TE buffer instead of the provided elution buffer.

Dog mtDNA was enriched in the libraries by in-solution hybridisation capture (Maricic et al., 2010), with the following modifications. One PCR of 1 μl barcoded library in a 50μl reaction volume were performed for each sample prior to capture to obtain 2μg of sequencing library per sample, using KAPA’s HiFi DNA Polymerase with the following reagent concentrations: 1x KAPA HiFi Buffer, 0.3mM of each dNTP, 0.3μM of each primer, and one unit of KAPA HiFi DNA Polymerase. Each sequencing library was enriched
independently, and pooled in equimolar ratios after a further 10 cycles of PCR amplification. Additionally, the libraries were eluted from the MyOne Streptavidin C1 Dynabeads (Invitrogen) by heating for 3 minutes to 95°C instead of treatment with sodium hydroxide. Libraries were quantified using the Qubit 3.0 fluorometer (Life Technologies), pooled in equimolar concentrations and sequenced using 2 x 75 base pair-end runs on the Illumina MiSeq sequencing platform. Libraries prepared from extraction blanks were subsequently sequenced on a separate run. In addition, the bait used for capture (tissue obtained from a local veterinary clinic) was separately sequenced and checked against sample sequences at variable positions.

5.2.3 Raw data processing and analysis

As aDNA fragments are often short enough to contain the sequencing adaptor, we preprocessed the raw reads to remove adaptors and merge paired end fragments, which overlapped by at least 11 base-pairs, using AdapterRemoval (v1.5.4) (Lindgreen, 2012). Reads were also processed to remove stretches of Ns, bases that had a low quality score (<30), and short reads (<25).

Contamination from modern DNA represents a challenge for any aDNA study, particularly when archaeological samples were not collected specifically for aDNA analyses, as sequenced reads from modern organisms may map to the reference genome of interest. To identify the impact of this effect, all read mapping was done onto a composite reference genome consisting of the Cambridge Reference Sequence for humans (Andrews et al., 1999), in addition to mitochondrial reference genomes from cow (Bos taurus, Gen Bank NC_006853), pig (Sus scrofa, NC_0012095.1), and chicken (Gallus gallus, Gen Bank NC_001323.1). The contamination ratio was determined by calculating the ratio of all reads with mapping quality >= 20 for each reference, compared with those for the dog mitochondrial reference sequence (NC_002008) (Kim et al., 1998).

Ancient DNA is usually characterised by C→T transitions at the 5’ ends of the molecule, and following double stranded library preparation G→A transitions at the 3’ ends of the molecules (Green et al., 2010). Reads were aligned to the composite reference genome using BWA
(0.7.10) (Li and Durbin, 2009) with recommended aDNA settings (Schubert et al., 2012); specifically, seeding was disabled (-l 1024), the number of gap opens was set to 2 (-o 2), and the maximum edit distance was set to 0.03 (-n 0.03). PCR duplicates were removed from the merged reads using a python script originally developed for the Neanderthal genome project (Fu et al., 2014). These PCR duplicates were also removed from the unmerged reads using Picard’s MarkDuplicates tool (http://broadinstitute.github.io/picard/). To confirm the authenticity of our aDNA, the program mapDamage (v2.0.2-9) (Jónsson et al., 2013), which identifies characteristic aDNA patterns, was used with the '-rescale' option to lower the quality score of likely damaged sites. The plots of characteristic damage patterns were individually created for merged and unmerged reads. Finally, Ts (thymines) found at the 5’ end of a read and Gs (guanines) at the 3’ end of the read, within the first two bases, had their quality scores rescaled to zero. Samples with less than 95% coverage of the dog reference genome were removed, as retaining samples where the DNA only partially covers the dog reference genome would have significantly impacted downstream analyses (Joly et al., 2007).

Reads that mapped to the dog reference genome were extracted from the composite alignment and variant call files (VCFs) were generated using the GATK Haplotype Caller (v3.3) (McKenna et al., 2010) with settings specific for haploid genomes, such as the mitochondrial genome. This file was filtered for mapping quality (<20), and bases supported by fewer than three reads (Thalmann et al., 2013, Knapp et al., 2012b). This approach was justified under the assumption that the above quality control (QC) procedure preferentially reduces the quality of the damaged positions in the sequencing reads, which would lead to a reduction in the probability that the damaged position would be called downstream. Finally, to maximise the usability of all the sequences obtained in our analyses, we carried out imputation using Beagle (v4.0 r1399) with default settings and the 'gt' argument for the VCF input, imputed sites with a probability greater than 0.8 were retained (Browning and Browning, 2007). To confirm the results of the QC, and imputation, visual inspection of all sites in the alignment files was undertaken using IGV (Robinson et al., 2011). The imputed VCFs were then compared with the positions in the bait sequence that varied from the reference. The bait sequence was processed separately using the same computational
protocol as the Wairau Bar samples. In addition, libraries prepared from negative controls were processed and aligned to the composite reference genome.

Coverage plots were created for all samples, and plots of fragment length were created for the merged reads using the ggplot2 package for the R programming language (Wickham, 2009, RDC Team, 2008).

Consensus sequences containing indels (insertions and/or deletions) were created for each sample that passed QC, and these have been deposited in GenBank (Accession Numbers KT168369-KT168382). In addition to the aforementioned QC measures, for these GenBank sequences non-variant sites that were supported by fewer than three reads were changed to ‘N’s.

5.2.4 Phylogenetic Analyses

The population structure of the Wairau Bar dogs was characterised using network analysis. Networks usefully visualise the relationships between haplotypes, and can help to understand how haplotypes are related to each other (Bandelt et al., 1999). The imputed VCF files were converted to FASTA and then to NEXUS format, using a custom python script and Biopython (Cock et al., 2009) respectively. This dataset was then used to create a median-joining network using Popart (v1.7.1) (http://popart.otago.ac.nz/) with default settings. This method works by resampling from clusters of trees that minimise the distance between haplotypes to generate a parsimonious network.

To investigate the place of the ancient New Zealand dogs in wider regional and global dog populations, we combined the aDNA sequences from the Wairau Bar dogs with sequences reported from previous studies of ancient and modern dogs. To date, the only reported ancient dog sequences from the Pacific comprise very short fragments of the control region from Polynesian and New Zealand dogs (Savolainen et al., 2004). Thalmann and colleagues (2013) have recently documented complete mitogenomes for ancient dogs. Although these samples came from archaeological contexts in Europe and North America we have included samples from the major dog haplogroup A in our analyses as they provide the only other
published ancient mitogenomes for comparison with the Wairau Bar dogs. Mitochondrial genome sequences and control region haplotypes were also obtained from studies of modern dogs from the Asia-Pacific region (Pang et al., 2009, Savolainen et al., 2004, Oskarsson et al., 2012) to investigate the geographical context of the samples.

We used the multiple sequence aligner MUSCLE (v3.8.31) (Edgar, 2004) to align the Wairau Bar sequences to the ancient and modern mitogenomes, and the modern control region haplotypes obtained from the literature. In their study of mitogenomes from dogs and wolves worldwide, Pang and colleagues (Pang et al., 2009) analysed only 16,195 bp of the mitochondrial genome to allow them to exclude repetitive and difficult-to-align regions. The control region (15458–16039) was extracted from Wairau Bar sequences using bioawk (https://github.com/lh3/bioawk) for comparison with the modern control region haplotypes.

Median joining networks were then created with these datasets. In addition, the ancient Polynesian short control region fragment (15458–15720) was extracted from the Wairau Bar sequences and compared to the Arc1 and Arc2 haplotypes to investigate if they are present within our dataset (Savolainen et al., 2004, Supplementary material).

5.3 Results

5.3.1 DNA preservation and sequence recovery

Acceptable sequence data was obtained from fourteen of the sixteen samples from which libraries were prepared (Appendix A). Two samples failed to generate acceptable consensus sequences (greater than 95% of the reference sequence covered by a read depth of greater than two) so were discarded from further analyses. Complete mitogenomes were obtained from three specimens, with the remaining eleven specimens having nearly complete coverage (Appendix A). The control region had proportionally lower read depth than elsewhere in the mitogenome in all the samples (Appendix A), which may be related in part to the presence of a repetitive region confounding capture and sequencing in these regions (Pang et al., 2009). The median average read depth for the fourteen samples was 86x (varying
across samples from between 6x to 424x (Appendix A). The two discarded samples, MS10064 and MS10067, had much reduced coverage in comparison with other samples and also substantially lower read depth.

These results are significantly better than those reported for the human remains from Wairau Bar, where only four out of nineteen individuals provided sufficient mtDNA for downstream analyses (Knapp et al., 2012). This may be related to two factors; firstly, improved methods (Illumina vs. Roche 454 sequencing platforms), and secondly, that the dog remains were excavated very recently. The human remains were excavated over forty years ago, and had been housed in a museum collection prior to aDNA extraction.

5.3.2 Sequence authenticity

Throughout DNA extraction and library preparation, blank extractions were processed alongside samples to provide negative controls. None of the controls contained any reads mapping to the dog reference mitogenome. The possibility of contamination due to laboratory reagents has also been raised in aDNA studies, based on the results presented by Leonard and colleagues (2007). Our analyses show very low levels of human and cow contamination in the dog samples (Appendix A). No chicken DNA was found in any of the samples. Additionally, our analyses showed that only six of the 479,410 reads from the blank extraction libraries mapped confidently (MAPQ > 20) to the composite reference genome containing mitochondrial DNA from human, dog, pig, and chicken. No reads supporting variable positions in the bait sequence were detected. Damage patterns such as short fragment lengths (Appendix A) and deamination patterns (Appendix A) are consistent with those expected from aDNA (Sawyer et al., 2012).

5.3.3 Population structure of Wairau Bar dogs

The median-joining network constructed for the Wairau Bar dogs (Figure 5-3) has a single main node that comprises eight samples, with four nodes radiating outwards. Two of those nodes contain two samples each, separated by one mutation. Another single sample differs from the central node by only one mutation. There is also a further single sample that is
separated from the central node by two mutations. The network illustrates the low level of genetic diversity among the dogs we have sampled from Wairau Bar, with five haplotypes spread across the fourteen specimens. These five haplotypes are closely related to one another, differing by a maximum of three mutations across 16.7kb of the mitogenome and the entire network. The central node contains over half of all the dogs sampled.

Figure 5-3: Median-joining networks for Wairau Bar dog sequences. a) Complete mitochondrial genomes. Sample numbers: (I) MS10062, MS10068, MS10070, MS10130, MS10131, MS10132, MS10135, MS10137, (II) MS10065, MS10136, (III) MS10069, (IV) MS10129, MS10133, (V) MS1006. b) Comparison with 263 bp Arc1 and Arc2 short control region fragment.

Since the archaeological site at Wairau Bar is one of the earliest sites in New Zealand, sufficient time is highly unlikely to have elapsed since settlement for new lineages to arise within the dog population. It is possible therefore that the haplotypes observed in the dogs we have sampled represent the founding lineages for the initial Polynesian dog population in New Zealand. Determining the number of founding haplotypes based on aDNA can, however, be a challenging task as observed variation can be the result of private mutations or false positives resulting from DNA damaged variants being called (Witt et al., 2015). In addition, given the tight temporal and spatial context of the sample (all individuals came from the same oven feature, deposited in one event) samples may represent several generations carrying the same lineage, possibly bitches and their offspring. This tight context may constrain our ability to identify founding lineages from within the wider Wairau Bar dog population.
5.3.4 Complete mitogenomes

The network in (Figure 5-4) shows the Wairau Bar dogs along with published ancient mitogenomes from the Americas, ancient breeds the Dingo and Basenji (Thalmann et al., 2013) and modern East Asian dogs (Pang et al., 2009). The ancient Wairau Bar dogs form a cluster that is located on a major branch of the network, while the Basenji and other ancient dogs are on other branches. The major Wairau Bar dog branch, which has two other branches one of which includes the Dingo sample, corresponds to Pang and colleagues’ (2009) sub-group a2 for Haplogroup A, which has a predominantly East Asian distribution. The modern individual on the same branch as the Wairau Bar dogs comes from a dog sampled in Thailand that carries the A75 control region haplotype.

Figure 5-4: Median-joining network of complete mitochondrial genomes of Wairau Bar dogs, and modern and ancient sequences from Haplogroup A. Wairau Bar = red, Dingo = purple, ancient American dogs = blue, Basenji = pink.
5.3.5 Control region haplotypes

In the absence of complete mtDNA sequences from ancient dogs from the Asia-Pacific region, we compared sequences from the Wairau Bar dogs with published control region haplotypes (582 bp, positions 15,458–16,039) for modern dogs within the dog haplogroup A (Pang et al., 2009, Oskarsson et al., 2012). These control region haplotypes have been used extensively for studies of dog domestication and dispersal throughout the Old and New Worlds. The Wairau Bar dogs all carry the control region haplotype A192. This haplotype has only been reported in modern village dogs from Bali in Indonesia (Oskarsson et al., 2012).

The network shown in Figure 5-5 was constructed from control region sequences from the Wairau Bar dogs and modern Haplogroup A dog sequences, ignoring insertion or deletion mutations (indels) which are not taken into account in network construction. The Wairau Bar dogs form a node with the A75 and A192 haplotypes. The A192 haplotype contains several indels. If the indels are discounted, the A192 haplotype corresponds to A75 (Oskarsson et al., 2012, Supplementary material). This A75 haplotype has been found in modern dogs in small numbers in China and Thailand, and in much higher frequencies in Indonesia (40%) (Oskarsson et al., 2012). Also expanding off this node are two haplotypes (A194, A195) also obtained from samples of modern dogs from Bali, and one (A145) reported from China and Bali (Oskarsson et al., 2012).

When compared with the Arc1 and Arc2 short fragment reported for 19 ancient Polynesian dog samples, generated by the Matisoo-Smith laboratory (Savolainen et al., 2004), all the Wairau Bar dogs produced sequences that would be assigned to the short haplotype Arc2 (Fig. 2b, S3 Table). This haplotype occurs in modern dogs throughout Mainland and Island South East Asia (Savolainen et al., 2004, Oskarsson et al., 2012). To date, the Arc1 haplotype has not been identified in any dogs from Wairau Bar, either as part of the short fragment study (Savolainen et al., 2004), or during this study. The Arc1 haplotype observed in the short fragment study was present in dogs from three New Zealand archaeological sites, at
Otehei Bay, Urupukapuka Island, Bay of Islands, Harataonga, Great Barrier Island, and Redcliffs, Christchurch, mainland South Island (Matisoo-Smith, unpublished data).

Figure 5-5: Median-joining network of control region sequences (582bp) of Wairau Bar dogs, and modern sequences from Haplogroup A. Modern = red, Wairau Bar = purple, A75 = yellow, A29 = blue.
5.4 Discussion

Dogs from Wairau Bar likely represent part of the initial population of dogs introduced to New Zealand, having arrived with people around the beginning of the fourteenth century AD (Jacomb et al., 2014). We sequenced fourteen complete or nearly complete mitogenomes, from a sample of dog teeth obtained from one oven feature at this important site. We observed five different mitochondrial haplotypes, which may have contributed to the establishment of the dog population of New Zealand. While it is possible that these haplotypes represent private mutations that are not carried further in the dog population, they may also be founding lineages. Further analyses of dogs from early archaeological sites in New Zealand would assist with determining the status of these lineages, if any additional haplotypes were introduced elsewhere, or if the Wairau Bar dog population can be considered representative of the Polynesian derived founding population as a whole. This initial dog population was then subsumed by a new wave of dogs introduced during the European colonisation of New Zealand in the 19th century. Analyses of dog bone from archaeological sites provide the only method of investigating the arrival and dispersal of the first dogs to arrive in New Zealand.

The limited number of mitochondrial lineages present at Wairau Bar suggests that only a few dogs may have been introduced to New Zealand within the first few years of settlement, despite the large numbers of dog bones recovered from early archaeological sites throughout the country (Davidson, 1987). However, it is also possible that the founding population came from a source population with very limited diversity, or that a combination of both factors contributed to the observed lack of variation. Historical observations of animal management practices in New Zealand and elsewhere in Polynesia, and the lack of animal control measures visible in the archaeological record, indicate that the majority of the dogs at Wairau Bar and other early archaeological sites were likely to be free-ranging. The carrying capacity of free-ranging dogs is closely related to the human environment that they occupy (Wandeler et al., 1988), and it is possible that a large dog population grew rapidly from a small number
of dogs following their arrival in New Zealand, given the quantity and quality of food sources available.

For populations such as that of Wairau Bar, which represent relatively recent migration events, the use of data solely from the mitochondrial control region is insufficient to address questions about population structure and founding events. Our analyses demonstrate the benefits of adopting a more fine-grained approach that identifies sufficient diversity to discriminate between individuals that would otherwise be assigned the same haplotype, and to clarify their relationships with each other. As with the dingo in Australia (Savolainen et al., 2004), the founding population of dogs in New Zealand is likely to have been drawn from a group of dogs that had passed through a series of genetic bottlenecks at times during their movement across the Pacific, hence arriving in New Zealand with limited genetic variation. The use of mitogenomes, however, enables us to identify five haplotypes within the Wairau Bar sample, whereas only one (A75) could have been identified had we used only the control region.

The analysis of mitogenomes of the Wairau Bar dogs also contributes to our understanding of the dispersal of dogs throughout the Asia-Pacific region, prior to their arrival in New Zealand. Pang and colleagues’ (Pang et al., 2009) analyses of the major dog clade identified ten sub-groups within Haplogroup A, with distinct geographical patterning. However, the majority of published mitogenomes from their study were from mainland China, with a few from Thailand, but none from Island South East Asia. The Wairau Bar dogs’ mitogenomes sit within sub-group a2, which has an East Asian distribution and includes the A29 control region haplotype found in dogs and dingoes (Pang et al., 2009). Further analyses of full mitogenomes from Island South East Asian village dogs and aDNA analyses of dog bone from archaeological sites could assist with understanding the distribution of dog lineages in this region, as human migrations from Island South East Asia likely played a pivotal role in the dispersal of dogs out into the Pacific.

The Wairau Bar dogs all carry the control region haplotype A75 (if indels are included in the analysis, the Wairau Bar dogs correspond to the control region haplotype A192) and the
short Arc2 haplotype. The A75 control region haplotype and the Arc2 short haplotype have a distribution across Mainland China, Island South East Asia and the Pacific, but to date have not been found in modern dogs in Taiwan or the Philippines (Oskarsson et al., 2012). The A192 haplotype is found in modern village dogs from Bali, Indonesia (Oskarsson et al., 2012 Supplementary Information). The settlement of Remote Oceania is associated with expansion of Austronesian language speakers; these languages are thought to have originated in Taiwan, and were then introduced through the Philippines and southwards through ISEA and into the Pacific (Blust, 2013). Based on the presence of the A75 (A192) haplotype in the Wairau Bar dogs and the modern distribution of this haplotype, it appears that at least one of the dog dispersals into the Pacific followed a more south-western route through Indonesia. Interestingly, this has parallels with the early movement of pigs and rats, based on genetic and geometric morphometric studies (Larson et al., 2007, Matisoo-Smith and Robins, 2004).

The short haplotype Arc2 was carried by all of the Wairau Bar dogs analysed in our study. Previous short fragment analysis of other ancient samples from New Zealand and Polynesia found that samples were split between two haplotypes, Arc1 (32%) and Arc2 (68%). At this point it is not known whether the lack of the Arc1 haplotype arises from a sampling bias, or relates to geographical or chronological factors. The Arc1 haplotype has been previously identified in dog samples from three archaeological sites in New Zealand, two of which are offshore islands (Urupukapuka and Harataonga). Further analyses are required to investigate the possible distribution of Arc1 in New Zealand.

5.5 Summary

We sequenced fourteen complete mitogenomes of dogs from one of the earliest archaeological sites in New Zealand and were able to identify five haplotypes, rather than the single haplotype that would have been observable using the control region. These results suggest that sequencing additional complete mitogenomes from dog bones from New Zealand and other Polynesian archaeological sites may be able to refine our understanding
of the dispersal of dogs during the colonisation of New Zealand, and their subsequent movements at local and regional scales.

In the next chapter, I present the results of sequencing mitogenomes from a range of colonisation era archaeological sites in New Zealand and the Cook Islands to examine this possibility further.
Chapter 6
Dog introduction and dispersal during the colonisation of New Zealand

Preface

The majority of this chapter forms the basis of manuscript submitted to the Journal of Pacific Archaeology as follows: Greig, K., Boocock, J., Walter, R. and E. Matisoo-Smith. Dog introduction and dispersal during the colonisation of New Zealand. Supplementary information is provided in Appendix B.

6.1 Introduction

Domesticated animals, including dogs, are part of the archaeological narrative for the human colonisation of Oceania in the Holocene, primarily as a component of a ‘transported landscape’. This concept has been highly influential in Pacific archaeology, and is most closely associated with the movement of people associated with the LCC in the third millennium BC (Kirch, 2000). It describes a process where migrating groups moved across the Pacific taking with them a set of horticultural concepts, plants and animals that were then utilised on the new islands they settled. Three domesticated animals, the pig, dog and chicken, along with the commensal Pacific rat, make up the four animals introduced to the region as part of these Lapita migrations. Understanding these Lapita and subsequent migrations have underpinned the development of the commensal model (Matisoo-Smith, 2009, Storey et al., 2013), whereby patterns of animal genetic variation in domesticated and commensal plants and animals are used to inform and interrogate models and hypotheses about the origin and dispersal of people. Commensal studies in the Pacific have generally focused on the origins of founding populations and dispersals across the region (e.g. Storey et al., 2012, Matisoo-Smith and Robins, 2009, Larson et al., 2007), rather than finer grained studies of relationships between groups at smaller geographical scales (but see Matisoo-Smith et al., 1999).
Only two of these domesticated animals, the dog (kuri) and the Pacific rat (kiore), were successfully established during the human colonisation of New Zealand. Bones and teeth from both these species are found in early sites throughout the two mainland islands, and on some but not all offshore islands (Davidson, 1987). To date, no dog bones have been reported outside of archaeological contexts in New Zealand – that is, there is no evidence for a feral dog population developing in New Zealand prior to the arrival of European dogs in the late eighteenth century. This second wave of dogs was transported to New Zealand by explorers, whalers and sealers from as early as the 1790s. The numbers of these dogs rose quickly as European settlement expanded, to the extent that by the early 1880s the kuri no longer existed as a distinct type (Clark, 1995). Despite this loss, archaeological remains still provide a means of investigating the place of dogs in the early history of New Zealand.

Inherited traits observable in archaeological dog teeth were first used over forty years ago to investigate questions about New Zealand’s settlement history. Allo (1971) suggested that congenital dental abnormalities could be used to distinguish between different dog populations. This approach was applied to several archaeological assemblages in the 1970s and extrapolated to identify interactions between human communities [for example, Motutapu (Smith, 1981); Palliser Bay (Leach, 1979)]. Subsequently Clark (1997) pointed out that these kinds of dental abnormalities are also found in modern populations and may have been caused by interruptions to growth patterns, trauma or poor diet. Anderson and Clark (2001: 162) do, however, note that supernumerary alveoli of the mandibular M3 do appear to occur in greater frequency in kuri than in modern dogs.

The development of aDNA methods and sequencing technologies has enabled the investigation of genetic variation in archaeological samples using molecular techniques. Now, rather than making inferences about settlement history based on inherited morphological traits, these methods can be applied to investigate these relationships using DNA. Savolainen and colleagues reported two ancient Polynesian dog haplotypes based on a small fragment of mtDNA, observed in sequences from archaeological dog remains from Hawaii, the Cook Islands, and New Zealand (Savolainen et al., 2004). Our group sequenced
complete mitochondrial genomes from a sample of dogs from Wairau Bar and identified five different haplotypes, which suggested that further work using complete mitogenomes could be informative about patterns of dog genetic diversity across New Zealand (Greig et al., 2015).

Since the 1970s, ideas about the timing and nature of New Zealand colonisation have also been significantly revised. Excavations at early archaeological sites such as Wairau Bar and Shag Mouth, combined with refinements in radiocarbon dating, have resulted in a much shortened chronology for colonisation, and the characterisation of the human colonisers as highly mobile (Anderson and Smith, 1996, Jacomb et al., 2014). Migrants are thought to have departed from an East Polynesian homeland zone where interactions were maintained across several island groups (Walter, 1996). Dog bones are present in early New Zealand archaeological sites in high numbers. With the development of aDNA techniques, it is timely to revisit the possibility of using genetic markers to investigate the origins and relationship between groups of dogs from early New Zealand archaeological sites, and possible links with the East Polynesian homeland.

This study presents the results of aDNA analysis of complete mitochondrial genomes of 35 dogs from the colonisation era of New Zealand. We describe the mitochondrial genetic diversity of dogs sampled from five early archaeological sites in New Zealand and one in the Southern Cook Islands, and consider the implications of these results for the current model for New Zealand colonisation.

6.2 Materials and methods

6.2.1 Samples and sites

Samples from New Zealand were obtained from five early archaeological sites: Houhora (Mt Camel), Wairau Bar, Shag Mouth, Pleasant River, and Pounawea, geographically spread across the country from north to south (Table 6.1, Figure 6-1). Houhora, Wairau Bar and Shag Mouth are colonisation phase sites that date to the first decades of New Zealand settlement. They contain the bones of moa and other extinct birds and artefacts with
distinctly East Polynesian stylistic affinities. Pleasant River and Pounawea also contain artefacts and extinct faunal material indicative of colonisation era occupation. The assemblages of dog bones from Wairau Bar and Houhora were identified to skeletal element, using the reference collection at the University of Otago Archaeological Laboratories. The large number of dog bones excavated from these two sites, as well as Shag Mouth, enabled specimens to be selected of the same element from each site assemblage, ensuring that each specimen represented a different individual. We can be sure therefore that no dog from these three archaeological sites was sampled more than once. Specimens from Pleasant River and Pounawea, however, comprise a variety of skeletal elements that could come from the same individual.

**Houhora (Mt Camel)**

The Houhora archaeological site is situated near the tip of the Aupouri Peninsula, just inside the entrance to the Houhora Harbour, at the top of the North Island. The site lies on a coastal terrace below high stable dunes at the base of the prominent landmark of Mt Camel. Excavations were carried out in the mid-1960s and early 1970s by archaeologists from the Anthropology Department of the University of Auckland. Although several theses were written about aspects of the excavations, an overall synthesis of the results was not produced for over thirty years (Furey, 2002).
Table 6.1: New Zealand and Southern Cook Island archaeological specimens sequenced for this study. Archaeological site where the specimen was obtained and when site was occupied, specimen number, and mitogenome coverage (% of genome).

<table>
<thead>
<tr>
<th>Archaeological Site</th>
<th>Time of occupation</th>
<th>Laboratory Specimen No.</th>
<th>Mitogenome coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Houhora</td>
<td>Early 14th century (Furey, 2002)</td>
<td>MS10071, MS10072, MS10073, MS10074, MS10075, MS10076, MS10077, MS10078, MS10079</td>
<td>100, 100, 100, 100, 100, 99.67, 99.86, 99.56, 99.74</td>
</tr>
<tr>
<td>Pleasant River</td>
<td>Late 14th and 15th centuries (Smith, 1999)</td>
<td>MS10093, MS10094, MS10095</td>
<td>99.12, 99.98, 99.32</td>
</tr>
<tr>
<td>Pounawea</td>
<td>14th to 15th century (Smith and James-Lee, 2010)</td>
<td>MS10096, MS10097, MS10098</td>
<td>99.99, 99.99, 99.96</td>
</tr>
<tr>
<td>Shag Mouth</td>
<td>Early 14th century (Anderson et al., 1996)</td>
<td>MS10111, MS10112, MS10113, MS10114, MS10115, MS10124</td>
<td>98.98, 99.57, 100, 99.41, 99.09, 98.77</td>
</tr>
<tr>
<td>Wairau Bar</td>
<td>First half of the 14th century (Jacomb et al., 2014)</td>
<td>MS10062, MS10065, MS10066, MS10068, MS10069, MS10070, MS10129, MS10130, MS10131, MS10132, MS10133, MS10135, MS10136, MS10137</td>
<td>99.95, 100, 99.58, 99.88, 100, 99.99, 98.68, 99, 99.72, 99.72, 99.98, 99.91, 100, 99.25</td>
</tr>
<tr>
<td>Ureia (AIT-10)</td>
<td>Mid 14th century (Allen and Wallace, 2007)</td>
<td>MS10326, MS10327</td>
<td>95.67, 96.07</td>
</tr>
</tbody>
</table>
Figure 6-1: Map of Polynesia, showing New Zealand and the Cook Islands, inset Southern Cook Islands.

The site contains some ill-defined relatively recent horizons containing European artefacts and a much better defined layer that Furey (2002) interprets as a small village occupied in the early fourteenth century. The site contains artefacts with stylistic affinities with East Polynesian forms, the bones of extinct birds such as moa, and numerous sea mammal bones. Features excavated include artefact manufacturing zones, cooking areas, waste midden dumps and postholes from structures. The site is considered to be a village with year round occupation, rather than a short duration camp, on the basis of the amount of archaeological material excavated, the variety of artefacts, the size and extent of cooking features, and the
depth of the stratigraphy (Furey, 2002 :122). The dog bone sample used for this study came from excavation areas spread across the site from layers that are thought to correspond to the earliest occupation levels.

**Wairau Bar**

Wairau Bar lies at the mouth of the Wairau River, on the east coast of the South Island. The site was first formally excavated in the 1950s and 1960s by the Canterbury Museum, and again more recently in 2009 when the Museum repatriated koiwi tangata (human remains) that had been previously taken from the archaeological site for reburial. Jacomb and colleagues (2014) used aDNA and Bayesian analysis of radiocarbon dates to develop a high precision chronology for deposition of a large cooking feature (Oven Pit 1) that was excavated during the 2009 fieldwork. This showed that the oven had been filled with midden during a single event that took place in the first half of the fourteenth century AD. The dog bones used in this study were all excavated from this feature, and the rapid deposition indicates that the dogs were likely to have been a part of the dog population living at Wairau Bar immediately prior to their death and incorporation in the oven fill.

**Shag Mouth**

The archaeological site at Shag Mouth is located at the mouth of the Shag River, just south of Oamaru, on the east coast of the South Island. Anderson and Smith (1996) describe Shag Mouth as a ‘transient village’; a settlement established close to rich resources such as seals and moa, occupied briefly, and then abandoned once these resource were depleted. The suite of radiocarbon dates from the excavations show that occupation began in the early fourteenth century AD and spanned 40 to 50 years (Anderson et al., 1996).

The archaeological site lies in dunes immediately above the beach, and comprises features interpreted as representing a range of structures facing a central activity area, with waste being discarded around the periphery (Anderson et al., 1996). The dog bones used for this study were excavated in 1989 from two parts of the site, SM/B and SM/D.
Pleasant River

The archaeological site at Pleasant River comprises a complex of archaeological features spread along the estuarine margins and dunes at the mouth of the river, 20 km south of Shag Mouth, and about 50 km north of Dunedin. Excavations in 1991–1993 showed the site contained a similar faunal assemblage to Shag Mouth, however, the stratigraphy, structural features and artefact assemblage were quite different (Smith, 1999). The site contains a wide range of fauna, including moa bones and eggshell, seal, dog, bird, fish and shellfish remains, but lacks evidence for permanent structures and distinct activity areas. Blades and flakes associated with food processing dominate the artefact assemblage, although other types are also present but not in large numbers.

The radiocarbon chronology suggests a series of short-term occupations spanning several hundred years, but occurring in two phases (Smith, 1999). The first phase of occupation appears to be contemporary with the major settlement at Shag Mouth, when temporary campsites at Pleasant River may have functioned as satellite settlements of the Shag Mouth community. The later phase of occupation in the late fourteenth and fifteenth centuries may have been part of a larger and more dispersed settlement system. The sample for this study was obtained from Area 1 Layer 2, which on the basis of the radiocarbon chronology dates to later than the main occupation at Shag Mouth, around the late fourteenth to early fifteenth century, but nonetheless still contains a similar faunal assemblage including moa and sea mammals indicative of the colonisation era.

Pounawea

The Pounawea archaeological site, in the Catlins area of coastal Otago, is located off Manuka Point at the confluence of the Catlins and Owaka Rivers, a few hundred meters from the river mouth. The archaeological site is interpreted as a coastal village, similar to Shag Mouth, and contains a wide range of fauna including moa and seals, and abundant artefacts with stylistic affinities to other early assemblages (Hamel, 2001). Smith and James-Lee’s (2010) review of the radiocarbon dates indicate the site was occupied between the fourteenth and fifteenth centuries. The location of the site near the river mouth was subject to increasing erosion in the 1970s, so a salvage excavation was carried out in 1979 to obtain information
about the environment, economy, material culture and any structural remains prior to the loss of the site (Hamel, 1980). Dog bones, included the sample used in this study, were recovered from across the site.

**Ureia, Southern Cook Islands**

In addition to the New Zealand samples, we included two specimens excavated from the Ureia archaeological site (AIT-10) on Aitutaki in the Southern Cook Island group. These specimens date to around the mid-14th century (Allen and Wallace, 2007), so are similar in age to the New Zealand samples. Ureia, however, is thought to post-date the initial colonisation of the Cook Islands, as it lacks large numbers of indigenous bird species and sea turtle commonly found in early sites (Allen and Wallace, 2007). Conversely, the presence of pig, chicken, dog, and Polynesian rat throughout the sequence are suggestive of an established agricultural economy.

### 6.2.2 aDNA methods

Forty-one dog teeth were sampled for this study. All DNA extraction and sequencing library preparation before PCR amplification was carried out at the Ancient DNA Laboratory at the University of Otago, where stringent procedures are in place to avoid contamination (Knapp et al., 2012a). Extraction, library preparation, and raw data processing were undertaken as described in Chapter 5. In brief, we carried out silica-based extractions (Rohland and Hofreiter, 2007b), from which sequencing libraries were prepared as described by Knapp and colleagues (Knapp et al., 2012c) for Illumina sequencing, with slight modifications. Raw sequence reads were processed using a purpose-built pipeline, and consensus sequences obtained for all specimens, and these have been deposited in GenBank. In these GenBank sequences non-variant sites that were supported by fewer than three reads were changed to 'Ns'.

We investigated the genetic population structure of the dogs by constructing a median-joining network using Popart (v1.7.1) ([http://popart.otago.ac.nz/](http://popart.otago.ac.nz/)) with default settings. To investigate possible population expansion or contraction based on deviation from a neutral model of evolution, we also estimated Tajima D again using PopArt. We used Bayesian
methods implemented in Beast to investigate demographic history using the skyline plot function (Drummond et al., 2005). For this aspect of the analysis we incorporated sequences from GenBank from modern dogs in Indonesia and Thailand that have previously been shown to be similar to sequences obtained from Wairau Bar specimens (Greig et al., 2015). In addition, all ancient haplotypes were aligned to the ancient Polynesian Arc1 and Arc2 short control region fragments (15458–15720) using MUSCLE (v3.8.31) (Edgar, 2004) to investigate whether these short haplotypes appear within our dataset.

6.3 Results

6.3.1 DNA preservation, sequence recovery and authenticity

Complete mitogenomes were obtained from eight specimens, with another 27 specimens having nearly complete coverage (Appendix B). Libraries constructed from six specimens failed to generate consensus sequences of a sufficient length (greater than 95% of the reference sequence) so were discarded from further analyses. As with the previous analysis of mitogenomes from Wairau Bar dogs described in Chapter 5, the control region generally had proportionally lower read depth than elsewhere in the mitogenome. The median average read depth for the 35 specimens was 42x, varying across specimens from between 6x and 432x (Appendix B). The six discarded specimens had much reduced coverage in comparison with other samples and also substantially lower read depth. Damage patterns such as short fragment lengths and deamination patterns were consistent with those expected from aDNA (Sawyer et al., 2012). Throughout DNA extraction and library preparation, blank extractions were processed alongside samples to provide negative controls. These did not contain any reads mapping to the dog reference genome.

6.3.2 Genetic population structure

The median-joining network (Figure 6-2) reveals a striking lack of genetic diversity in the sequences from the New Zealand dog specimens, and a close relationship between these sequences and those from the Southern Cook Islands. The New Zealand sequences form a central haplotype comprising over 70% of the total number of dogs sampled, and with representatives from all five archaeological sites. The nodes with a single point mutation
radiating off this central node may comprise rare mutations that were carried only by individual dogs and possibly their descendants, but not at a level that relates to population structure. None of these singletons are shared between dogs from different archaeological sites.

Figure 6-2: Locations of archaeological sites in New Zealand and the Southern Cook Islands and a median-joining haplotype network of mitogenomes of dogs. The archaeological sites and corresponding network nodes are coloured as shown in the legend.

There are three branches with terminal nodes that differ from the central haplotype by two or more mutations. One is a sequence from Shag Mouth, while the other comprises a sequence from Wairau Bar, which has an intermediate sequence from Shag Mouth also on this branch. The two sequences from the Southern Cook Islands are more divergent than the New Zealand sample, differing by six mutations from each other and by one and seven mutations respectively from the central New Zealand node. In respect of the short Arc1 and Arc2 haplotypes reported by Savolainen and colleagues in New Zealand and the Cook Islands (2004), only Arc2 was present in the dogs that we analysed (Appendix B).
The New Zealand sequences display a ‘starburst’ pattern, with one central node or haplotype, surrounded by other nodes radiating outwards (Figure 6-2). This starburst pattern is associated with a recent population expansion from a small number of founders (Avise, 2000). The results of the Tajimas D test calculated in Popart gave a value of -2, which is significant and indicates deviation from neutral evolution. The distinctive starburst network and the Tajima’s D value indicate a past population expansion, which may predate the appearance of dogs in high numbers in Polynesian sites. Further attempts to infer possible parameters of this process using a Bayesian skyline plot were unsuccessful, and were likely confounded by the short time frame, low sample numbers and lack of diversity.

6.4 Discussion

When people arrived in New Zealand, they brought their dogs with them. To investigate the introduction and dispersal of dogs during the East Polynesian colonisation of New Zealand, we generated 37 complete or nearly complete mitogenomes, from dogs sampled from five early archaeological sites in New Zealand and from one archaeological site from the Southern Cook Islands. The earliest layers of the New Zealand archaeological sites (Houhora, Wairau Bar, Shag Mouth, Pleasant River and Pounawea) all date to within the first hundred years of human arrival in the early fourteenth century AD, while Ureia in the Southern Cook Islands is slightly older. The New Zealand archaeological sites are spread across the country from the top of the North Island, to the lower South Island, spanning a distance of nearly 2,000 kilometers.

6.4.1 Composition of the founding population

Our analysis of the genetic population structure of the New Zealand dogs shows that the founding population had extremely limited genetic diversity, with no observable geographic patterning. The genetic structure of the sampled dogs is dominated by one haplotype, which is carried by nearly three quarters of the dogs sampled. Of the 10 dog specimens that did not carry the founding haplotype, eight differ by only one mutation and two by two mutations. This pattern is the same as we observed previously in a smaller sample of dogs from Wairau
Bar, where we suggested that further sampling may increase the number of haplotypes shared within the founding population. This has proven not to be the case.

Although colonisation era archaeological sites in New Zealand contain large numbers of dog bones, the mitogenomic homogeneity makes estimating the size of the founding population difficult. The founding human population of New Zealand has been estimated to be high—perhaps involving as many as 200 women (Whyte et al., 2005)—and the migration involved multiple voyages so it seems possible that large numbers of dogs may have been brought to New Zealand during this process. The introduction of dogs to New Zealand may have also occurred over several decades and from multiple points, rather than comprising a single founding event. This is consistent with the current model for the colonisation of New Zealand, discussed in more detail below.

6.4.2 Origins of New Zealand’s first dogs

The sample of dogs’ teeth used for this study was obtained from archaeological contexts dating to the first fifty to hundred years of dogs’ arrival in New Zealand. Given this chronology it is highly likely that the majority, if not all, of the single point mutations observed in the sample arose prior to this colonisation period. This raises the possibility of identifying the source population for the New Zealand dogs, where these mutations are also present. Tracking the origin of the New Zealand dog population in this way requires comparative data on the genetic makeup of potential source populations. So far, such data is scarce.

An analysis of a short fragment of mitochondrial DNA from Pacific archaeological sites generated by the Matisoo-Smith laboratory and first reported in Savolainen et al. (2004) was able to identify two short haplotypes present in the dogs sampled. These haplotypes were termed Arc1 and Arc2, and were observable in samples from archaeological sites in New Zealand, the Cook Islands and Hawaii. Only one of these haplotypes (Arc2) is present in the samples we have analysed from New Zealand and the Southern Cook Islands. While the distribution of these haplotypes across East Polynesia indicates a shared ancestry for dogs, they lack the level of discrimination necessary to address inter-island relationships.
In contrast to the dogs’ mitogenomic lineages, previous mitochondrial analysis of four human burials from a cluster at Wairau Bar showed that the individuals were not closely maternally related (Knapp et al., 2012b). This was the earliest burial cluster at the archaeological site and was also distinctive from other groups as stable isotope analysis revealed that these individuals were not local, having had different diets and childhood places of residence from individuals in other burial groups (Kinaston et al., 2013). It is possible that some of these individuals were first generation immigrants, born and raised elsewhere in East Polynesia. The lack of a close maternal genetic relationship was interpreted to mean that the founding population of New Zealand was derived from a number of different communities within an interaction zone that spanned several islands or archipelagos (Knapp et al., 2012b). This is consistent with the ‘Hawaiiki Zone’ hypothesis that argues on the basis of material culture and linguistics that New Zealand was settled from a zone that included the Southern Cook Islands and western French Polynesia, which was linked by regular interaction and exchange (Walter, 1994).

The mitogenomes of the dogs we sequenced from the Southern Cook Islands carry two distinct haplotypes, differing from each other by four mutations. One is very similar to the founding New Zealand haplotype, differing only by one mutation. It is possible that the lineages from New Zealand and the Southern Cook Islands were part of a larger source population of dogs spanning the wider Polynesian homeland zone. Alternatively, with further sequencing of East Polynesian samples, the New Zealand haplotype may be found to have a more localised geographical distribution. Further analysis of archaeological dog remains from East Polynesia is necessary to investigate possible origins in more detail.

6.4.3 Dog dispersal during the colonisation phase of New Zealand

Once established in New Zealand, the lack of mitochondrial genetic diversity of the dogs makes identifying connections between archaeological communities through shared dog mtDNA lineages impossible, as we cannot discriminate between dogs carrying the central haplotype. What we can say however from the distribution of dog bone in archaeological sites, is that dogs from the founding population and their descendants were rapidly
transported throughout New Zealand during the colonisation era. Given the homogeneity of the mitochondrial lineages, questions about the movement of closely-related dogs between archaeological sites in New Zealand may be better addressed using nucDNA markers, if these were able to be obtained. Stable isotope analysis may also assist with identifying local and non-local dogs, as has been carried out with a sample from Wairau Bar (Kinaston et al., 2013). In that study, the five specimens analysed all possessed strontium values consistent with the local environment, raising the possibility that while humans moved around the colonisation phase landscape, dogs remained fairly localised.

6.4.4 Introduction and establishment of Oceanic dogs

The mitochondrial make up of New Zealand’s first dogs also offers insights into the process of dog introduction across the wider region. The presence of dog bones in archaeological sites in the Pacific is patchy and discontinuous, and molecular genetic analyses can provide complementary information about the timing and trajectory of dog dispersal. The arrival of people and dogs in East Polynesia marks the last leg of a major dispersal by people across the previously uninhabited islands of Remote Oceania. The ancient mitogenomic lineages of the New Zealand and Southern Cook Islands dogs are very similar and suggest that the dogs being moved around by the East Polynesian human migrants may have come from populations with limited mitochondrial genetic diversity.

This limited diversity implies that during the movement of dogs across the Pacific there were events, or processes, that resulted in a reduction in canid genetic diversity. There are several putative scenarios that may have contributed to this reduction, some linked to human interactions with dogs and the migration process between islands. In the first scenario, a gradual process that involved a subset of dogs being sequentially taken from the existing population could result in a series of founder effects. This scenario could occur if dogs accompanied people in a series of linked migrations across the region. Favourable traits possessed by some of dogs may have been under human selection, contributing to a reduction in the overall genetic diversity. Alternatively, rather than a gradual process, a single or very limited movement involving a small number of dogs at a critical point could have
resulted in a major reduction in genetic diversity. This scenario could be related to human migrations, or alternatively through the movement of dogs via other mechanisms such as their incorporation into exchange networks. In the final scenario, a catastrophic event may have had a major effect on the dog population at some point.

In apparent contradiction to the reduction of mitochondrial genetic diversity, the starburst pattern of the haplotype network indicates a recent population expansion from a small number of founders (Avise 1987). While we are unable to date this expansion using molecular genetic methods, it would potentially be associated with a key period in the dispersal process, signaled in the archaeological record by the appearance of significant numbers of dog bones in Remote Oceanic archaeological sites around 2,000 years ago (summarised in Chapter 2).

The low mitochondrial genetic diversity observed in New Zealand dogs parallel findings about the mtDNA composition of dingoes introduced to Australia in the mid-Holocene. Although separated in time by at least 4,000 years, there appears to be similarities in the initial introduction of dogs to both locations, with founding populations being made up of an extremely restricted number of founding maternal lineages, yet dogs being successfully established in both places. The trajectory followed by the dogs introduced to Australia involved the separation of a group of dogs from human communities. This ultimately resulted in the dingo, a feral dog successfully adapted to living separately from people. The timeframe for this transition is not known but may have spanned several thousand years. During the much shorter time period in New Zealand, between the introduction of Polynesian dogs and the later arrival of European dog breeds, there is no evidence that a feral dog population became established. The only mitogenomes of dingoes sequenced to date are from modern specimens, and it is possible that ancient but now extinct dog lineages may have also existed that could reveal new relationships between Pacific dog populations.
6.5 Summary

Our study of mitogenomes from dogs from New Zealand and the Southern Cook Islands has revealed the strikingly limited mitochondrial diversity of dogs sampled from the colonisation era. While the uniformity of sequences from New Zealand archaeological sites has constrained our ability to investigate interactions between human communities, it does support the colonisation process being undertaken by highly mobile groups of people and occurring rapidly. The ancient mitogenomic lineages of the New Zealand and Southern Cook Islands dogs are very similar and suggest that the dogs being moved around by the East Polynesian human migrants may have come from populations with limited mitochondrial genetic diversity. The mitochondrial homogeneity in the dogs sampled from Wairau Bar differs from the mtDNA diversity observed in the people buried at Wairau Bar. The arrival of people and dogs in New Zealand marks one of the last legs in a major dispersal across the previously uninhabited islands of Remote Oceania, yet molecular genetics reveals differences in the genetic signatures of this dispersal in dogs and people that would not otherwise be known from the archaeological record.
Chapter 7

Ancient mitogenomes of South East Asian and Pacific dogs shows complex regional history

Preface

The chapter forms the basis of an unpublished manuscript to be submitted to a refereed journal for publication as follows: Greig, K., Boocock, J., McDonald, K., Addison, D. J., Allen, M. S., David, B., Gibbs, M., Higham, C. F. W., McNiven, I. J., O'Connor, S., Tsang, C. H., Liu, F., Walter, R. & Matisoo-Smith, E. Ancient mitogenomes of South East Asian and Pacific dogs shows complex regional history. Supplementary information is provided in Appendix C. The modern dingo DNA extraction and sequencing was carried out by one of the co-authors, but is provided in Appendix C for information purposes.

7.1 Introduction

Despite the status of dogs as the first species to emerge from the domestication process at least 15,000 years ago, the location, timing and nature of this process is the subject of continuing debate (Larson et al., 2012). Archaeological and genetic research has identified several possible regions for domestication: Central Asia, East Asia, the Middle East and Europe, with time frames varying between archaeological and molecular dating methods (Pang et al., 2009, Savolainen et al., 2002, Shannon et al., 2015). Two independent domestication events in East Asia and Western Eurasia (Europe and the Middle East) have recently been proposed (Frantz et al., 2016) and this research, along with analyses using nuclear markers (Pilot et al., 2015, Sacks et al., 2013), highlight the likelihood of the replacement of some early Eurasian dog lineages by those with East Asian ancestry.

While domestication marks the beginnings of the human relationship with dogs, this relationship is a continuing and dynamic one. It can take a range of forms, with differing degrees of interaction between people and dogs observable throughout the world today. Dogs can be companion animals and working dogs, live a free-ranging village existence, or
even become feral populations that have lost all connection with people. It is likely too that these contemporary forms of the relationship do not reflect the range that was present in the past, just as the genetic composition of modern dog populations may not be representative of past populations. Dog remains from archaeological sites therefore provide an important source of information not only about the dog domestication process but also about the ongoing relationship between dogs and people.

Archaeological evidence indicates that the first dogs introduced to Australasia and the Pacific region were the ancestors of the dingoes of Australia. The earliest dingo bones have been found in an archaeological context in the Madura Cave in South Australia dated to 3,500 years ago (Milham and Thompson, 1976). Molecular analyses of mtDNA and Y-chromosome markers suggest an Asian origin and give an earlier date range of between 4-18,000 ya (Savalainen et al., 2002, Oskarsson et al., 2012, Ardalan et al., 2012). Dingoes subsequently adapted to living independently from people and are now considered to be one of the few truly feral dog populations found today (Reponen et al., 2014), nonetheless, some individuals may still be incorporated into Aboriginal communities as companions and hunting aides (Smith and Litchfield, 2009). NGSDs, rarely found in the New Guinea Highlands, are genetically very closely related to dingoes (Savalainen et al., 2004). The history of these dogs is however little known. Further to the east across the Pacific Ocean, dogs were part of the last major pre-industrial human migration that resulted in the colonisation of the islands of East Polynesia (Kirch, 2000). Previous genetic studies have also indicated a likely Asian origin for these dogs (Savalainen et al., 2002, Oskarsson et al., 2012), however, limited archaeological and aDNA evidence from the Pacific, and no aDNA from Island South East Asian dogs, meant that little could be said about the timing and nature of this process.

Here we present complete or near complete mitogenomes from archaeological dog specimens from Thailand, ISEA and Pacific islands, and modern dingoes. When combined with additional complete mitogenome sequences from modern dogs from Southeast Asia,
Australia and New Guinea, a complex history of dog-human interactions and multiple introductions into the region is revealed.

7.2 Materials and Methods

7.2.1 Ancient DNA extraction, library preparation, sequencing and data processing

Twenty one dog specimens comprising bone and teeth from thirteen archaeological sites were used for this study (Table 7.1, Appendix C), including one specimen published in Greig et al. (2015). The mitogenomes reported here were generated using hybridisation capture and sequenced using NGS on the Illumina MiSeq platform in accordance with protocols to ensure the authenticity of aDNA sequences (as described in Chapter 5). Throughout DNA extraction and library preparation, blank extractions were processed alongside samples to provide negative controls.

We processed the raw sequence reads using a computational protocol that we have previously described (see Chapter 5). A notable feature of this pipeline is the use of a composite mitochondrial reference genome (made up of likely contaminant genomes) to remove and check for contamination. Coverage and read depth plots were created for all samples (Appendix C), and plots of read fragment length (Appendix C) were created for the merged reads using the ggplot2 package for the R programming language (Wickham, 2009, RDC Team, 2008). Reads displayed the expected characteristics of aDNA, including fragment length (Appendix C) and deamination patterns (Appendix C). To maximize sequence usability in downstream analyses imputation was performed and sites with greater than 80% confidence were retained. Consensus sequences containing indels (insertions and/or deletions) were produced for each sample that passed quality control, and these will be deposited in GenBank. In these GenBank sequences, non-variant sites that were supported by fewer than three reads were changed to 'N's.
Table 7.1: Ancient Mainland and Island Southeast Asian and Pacific samples sequenced for this study. Specimen numbers, mitogenome coverage (% of genome), location and date of archaeological sites (years before present) and source of date.

<table>
<thead>
<tr>
<th>Laboratory Specimen No.</th>
<th>Mitogenome coverage (%)</th>
<th>Archaeological Site</th>
<th>Location</th>
<th>Age (years before present)</th>
<th>Source of date</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS10015 MS10016 MS10017</td>
<td>54.25 98.33 71.36</td>
<td>Fale Islet (TU-2)</td>
<td>Atafu, Tokelau</td>
<td>500–310 BP</td>
<td>Addison et al., 2009</td>
</tr>
<tr>
<td>MS10023</td>
<td>43.06</td>
<td>Bogi 1</td>
<td>South Coast, Papua New Guinea</td>
<td>2900–2500 BP</td>
<td>McNiven et al., 2012</td>
</tr>
<tr>
<td>MS10031</td>
<td>17.20</td>
<td>Edubu 1</td>
<td>South Coast, Papua New Guinea</td>
<td>&lt;2350 and 2500 BP</td>
<td>McNiven et al., 2012</td>
</tr>
<tr>
<td>MS10043 MS10044</td>
<td>98.45 44.04</td>
<td>Goemu</td>
<td>Torres Strait</td>
<td>950–1000 BP</td>
<td>McNiven et al., 2003</td>
</tr>
<tr>
<td>MS10062</td>
<td>99.95</td>
<td>Wairau Bar</td>
<td>New Zealand</td>
<td>650 BP</td>
<td>Jacomb et al., 2014</td>
</tr>
<tr>
<td>MS10090</td>
<td>83.71</td>
<td>Graciosa Bay</td>
<td>Santa Cruz, Solomon Islands</td>
<td>350 BP</td>
<td>Gibbs et al., 2015</td>
</tr>
<tr>
<td>MS10324 MS10325 MS10664 MS10665</td>
<td>27.14 68.78 100.00 33.22</td>
<td>Nu’alolo Kai</td>
<td>Kauai, Hawaii</td>
<td>250–150 BP</td>
<td>Graves et al., 2005</td>
</tr>
<tr>
<td>MS10326 MS10327</td>
<td>95.67 96.07</td>
<td>Ureia (AIT-10)</td>
<td>Aitutaki, Cook Islands</td>
<td>750–550 BP</td>
<td>Allen and Wallace, 2007</td>
</tr>
<tr>
<td>MS10329</td>
<td>90.50</td>
<td>Taurama</td>
<td>New Guinea</td>
<td>1000–2000 BP</td>
<td>Sutton et al., 2015</td>
</tr>
<tr>
<td>MS10330</td>
<td>98.19</td>
<td>Matja Kuru 2</td>
<td>East Timor</td>
<td>2900–3075 BP</td>
<td>Gonzalez et al., 2013</td>
</tr>
<tr>
<td>MS10331</td>
<td>66.2</td>
<td>Ban Chiang</td>
<td>Thailand</td>
<td>3558–3392 BP</td>
<td>Higham et al., 2011</td>
</tr>
<tr>
<td>MS10332</td>
<td>79.36</td>
<td>Nong Nor</td>
<td>Thailand</td>
<td>2,650–2,450 BP</td>
<td>Higham et al., 2011</td>
</tr>
<tr>
<td>MS10333 MS10334</td>
<td>96.71 99.97</td>
<td>Shisanhang (SSH)</td>
<td>Taiwan</td>
<td>1,800–500 BP</td>
<td>Pietrusewsky and Tsang, 2003</td>
</tr>
</tbody>
</table>
7.2.2 Modern dingo mitogenomes

Blood samples from four dingoes from Wellington Zoo were collected by the zoo veterinarian staff and sent to the University of Otago. DNA was extracted using the MagJET magnetic bead (ThermoFisher) protocol (as per manufacturer’s instructions). Four long-range primer pairs were used to amplify the complete mitochondrial genome (Appendix C). Blunt end repair, ligation of sequencing adaptors and barcoding were carried out following (Kircher et al., 2012), with modifications for Illumina sequencing adaptors. Sequencing was carried out on the Illumina MiSeq platform. Reads were processed using a modified version of the aDNA pipeline described above, with the omission of the mapDamage step.

7.2.3 Phylogenetic reconstructions

We used jmodeltest2 (v2.1.7) (Posada and Crandall, 1998) to assess different models of evolution for the archaeological dog samples, modern dingoes and published modern sequences used in this study. The best model was then selected based on the Bayesian information criterion (BIC). According to this criterion a HKY+I model was most appropriate. Using this model, Bayesian trees were created with BEAST (v1.8.2) (Drummond et al., 2012) using 10,000,000 generations and sampling every 1000th generation. The maximum credibility tree was determined using TreeAnnotator with 25% of the trees being discarded as burn-in. Convergence diagnostics were investigated using Tracer (v1.4.0) and FigTree (v1.4.0). We used MEGA7 (Kumar et al., 2016) to generate a maximum likelihood (ML) tree with 10,000 bootstrap replicates. A median joining network of the Pacific samples, modern dingoes, NGSD, and modern dogs from GenBank was created using Popart (v1.7.1) (http://popart.otago.ac.nz/) with default settings.

7.3 Results and discussion

We generated complete and partial mitochondrial genome sequences from 21 ancient dog specimens from archaeological sites from Thailand, ISEA and the Pacific (Table 1, Appendix C), and a sample of four modern dingoes from the Wellington Zoo in New Zealand. Published sequences of modern dogs from ISEA, Papua New Guinea and Australia were also used for comparison (Pang et al., 2009).
7.3.1 East Asian ancestry for Australasian and Pacific dogs

Phylogenetic analyses of mitogenomes of modern dogs throughout the world have shown that over 65% of the dogs sampled fall within Haplogroup A (Duleba et al., 2015). There is geographical structure within this large haplogroup, with the greatest mitochondrial diversity being found in East Asia (Pang et al., 2009, Duleba et al., 2015, Fregel et al., 2015). Duleba and colleagues (2015) further identified several sub-clades that are present almost exclusively in East Asian modern dogs. These sub-clades were estimated to have the oldest evolutionary ages (between 15 kya to 38.7 kya) and are hypothesised to represent the initial founding gene pool. Here we follow Duleba’s phylogeny and nomenclature for dog haplotypes, which has the greatest resolution for sub-clade A2, which is the lineage carried by the majority of dogs we sequenced in our study.

All the specimens from which we obtained complete or near complete mitogenome sequences belong to Haplogroup A, with the exception of one specimen from Taiwan which belongs to Haplogroup B (see Appendix C for a list of variable sites relating to haplotype assignment). The median joining network in Figure 7-1 shows the relationship between the haplotypes. The network was produced using Popart (v1.7.1) (http://popart.otago.ac.nz). Within Haplogroup A, the majority of sequences from the ancient Pacific dogs belong to Dubela’s (2015) haplotype A2b2, sharing most, but not all, of the defining mutations with a specimen on one of A2b2 terminal branches. The dingoes belong to A2b3, and share only a few defining mutations with the specimen on one of the two A2b3 terminal branches. The second ancient specimen from Taiwan has a sequence that also sits within this A2b3 group. One ancient specimen, from East Timor, belongs to haplotype A4’5, and has some of the defining mutations for sub-clade A5.
Figure 7-1: Median joining network of mitogenome sequences from ancient Island Southeast Asian and Pacific dogs (filled circle), modern dingoes sequenced for this study (crossed circle) and dingoes, NGSD and Southeast Asian dogs from GenBank (hollow circle), produced using Popart (v1.7.1) (http://popart.otago.ac.nz/). Assignment to haplogroup after Duleba et al., 2015. The geographic origin of the sequences is shown by node colour.
These ancient mitogenomes represent a series of individual insights into the history of dog dispersal eastwards beyond East Asia prior to the influence of modern European dog breeds. We carried out a Bayesian analysis implemented in BEAST (v1.8.2) (Drummond et al., 2012), incorporating the archaeological dates for each sequence as priors, in order to investigate the divergence times between the ancient specimens carrying A and B haplotypes and subsequent splits (Appendix C). The tree however lacks sufficient resolution to be informative about such events, with overlapping date ranges for most splits. The topology of the tree remains the same whether or not tip dates are included, and is consistent with the tree generated using the maximum likelihood method (Figure 7-2).

7.3.2 **Multifaceted introduction of dingoes to Australia**

The use of control region motifs to assign haplotypes has previously been shown to provide insufficient discriminatory power to differentiate between several mitogenomic sub-clades, including A2 (Fregel et al., 2015), where dingoes and NGSDs haplotypes are located. Previous studies have identified one main mtDNA control region haplotype for dingoes, separated by one mutation from that carried by a small sample of NGSDs (Savolainen et al., 2004), and two Y chromosome haplotypes shared by dingoes and NGSDs (Ardalan et al., 2012, Brown et al., 2011). These results, along with results from blood haemoglobin (Runstadler et al., 2006), suggest an extremely limited introduction of dingoes, most likely via New Guinea or nearby islands. The analysis of whole mitogenomes enables a further avenue of investigation of the genetic relationships within the dingo population, and between dingoes and NGSDs. The dingoes we sampled, and the dingo and NGSD sequences we obtained from GenBank, are located within a yet to be defined sub-clade of A2b3 (Duleba et al., 2015).
Figure 7-2: Molecular phylogenetic analysis by maximum likelihood method, implemented in MEGA7 (Kumar et al., 2016). The evolutionary history shown inferred by using the maximum likelihood method based on the Hasegawa-Kishino-Yano model (Hasegawa et al. 1985). The tree with the highest log likelihood (-25257.5243) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix
of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.0500)). The rate variation model allowed for some sites to be evolutionarily invariant ([+I], 0.0010% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 45 nucleotide sequences. There were a total of 16774 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

The median joining network (Figure 7-1) and the ML tree (Figure 7-2) both differentiate between two distinct groups of dingoes. This is consistent with a previous analysis that found the same geographical distribution of dingo mtDNA lineages, split between the northwest and southeast of Australia (Cairns, 2014). This is also observable in results from nuclear and Y chromosome analyses (Cairns, 2014, Brown et al., 2011, Ardalan et al., 2012). These findings support the hypothesis that there were two independent and possibly spatially discrete dingo introductions to Australia. In addition, the dingo groups, while more closely related to the NGSDs than to any other dogs, do diverge from the NGSD sequence which may indicate a period of isolation between the two groups of dogs. The only genetic data from NGSDs has been obtained from the captive population derived from eight founders (Koler-Matznick et al., 2007). Village dogs in New Guinea however possess many, if not all, of the characteristics ascribed to NGSDs (Dwyer and Minnegal, 2016) and it is possible that sampling from populations away from main centres may also be informative about NGSD dog and dingo ancestral history.

Interestingly, a sequence we generated from a specimen excavated from a Taiwanese Iron Age archaeological site is situated in the network (Figure 7-1) and ML tree (Figure 7-2) between the two groups of dingoes. A similar link to Taiwanese dog lineages has been observed in Y chromosome markers (Sacks et al., 2013), as modern dingoes possess Y haplotypes closely related to those found in modern Taiwanese dogs. Sacks and colleagues (2013) raise the possibility of a dispersal event from Taiwan, or the movement of dogs with Austronesian language speakers out of China to the south through the Tonkin Gulf, in addition to the Austronesian expansion via Taiwan, to explain this observation. Molecular analysis of dogs from mainland archaeological sites during this critical period of the Neolithic expansion could greatly assist with tracking dog dispersals.
Although dates obtained from genetic analyses have also raised the possibility of an early pre-Holocene introduction of dogs to Australia (Oskarsson et al., 2012), there is currently no archaeological evidence for dingoes prior to about 3,500 years ago. Balme and O’Connor (2016) hypothesise that increasing amounts of small game animals observable in the archaeological record in the mid-Holocene may be indicative of the arrival and use of dingoes to assist with hunting. The mid-Holocene timing for the introduction of dingoes to Australia, based on archaeological dates and changes in faunal composition, along with molecular evidence indicating at least two introductions, suggests that dogs were being transported between islands by at least some groups of maritime people in ISEA by this time.
7.3.3 New lineage of Late Lapita dogs

Explanations for the Holocene human colonisation of Remote Oceania draw on archaeology and linguistics to model the movement of Austronesia language speakers out of Taiwan and into the Pacific, via the Philippines, followed by the emergence of the LCC which spread across much of Near and Remote Oceania (Bellwood, 2011, Kirch, 2000). Archaeological evidence for the LCC is found from New Guinea to West Polynesia. Various scenarios have been suggested, incorporating archaeological, linguistic and biological data, to account for the dispersal process and the extent to which interactions occurred between incoming Austronesian speakers and existing human groups in Near Oceania (for a summary see Matisoo-Smith, 2015).

Dogs are often thought to have been moved across the islands of Pacific by people associated with the LCC (Kirch, 2000). There is only limited archaeological evidence for dogs though in early eastern Lapita archaeological sites, and no dog bone reported in Lapita archaeological sites beyond the Bismarck Archipelago (for reviews see Matisoo-Smith, 2007, Specht et al., 2014). The presence of drilled dog teeth in Near Oceanic and Lapita archaeological sites, along with the scarcity of other skeletal elements, also indicates that dog remains may have been incorporated into archaeological contexts by mechanisms unrelated to the transport of living dogs, such as the exchange of valuables. Although mtDNA control region sequences indicate a route through ISEA for the ancestors of Polynesian dogs (Oskarsson et al., 2012, Savolainen et al., 2004), a possible link between Pacific dogs and modern Taiwanese dogs has also been suggested (Sacks et al., 2013).

Although we were unable to generate complete mitogenomes from the two archaeological specimens from the Late Lapita archaeological sites in the Papuan Gulf, we were able to obtain limited coverage of parts of the genome (43% and 17%). Compared to the rest of the ancient mitogenomes we sequenced, the Late Lapita specimens carry diagnostic mutations at several positions indicating that they fall within the B haplogroup, and also lack mutations for haplogroup A (Appendix C). The majority of positions used for B haplogroup assignment are found outside the control region. As observed previously in sequences from
ancient New Zealand dog specimens (Greig et al., 2015), the control region of the Late Lapita sequences has proportionally lower read depth than elsewhere in the genome. This characteristic, combined with the low coverage generally, means that we were unable to assign the Late Lapita sequences to a control region haplotype (i.e., after Savolainen et al., 2004). These are the first specimens that have potentially been identified with a B lineage in the Pacific, and suggest a possible link with the Austronesian dispersal from Taiwan. The two nearly complete ancient mitogenomes from dogs from the Iron Age Shisanhang Taiwanese are highly divergent compared to other ancient sequences, comprising both A2b3* and B haplogroups. The greatest genetic diversity of modern dogs is found in MSEA (Pang et al., 2009, Wang et al., 2016), and it is perhaps not unexpected to observe these divergent lineages in Taiwan adjacent to the mainland. Additional mitogenomes from dogs from archaeological sites in Taiwan, and samples from the Philippines and other ISEA sites would be extremely valuable for tracing the Neolithic dispersal of dogs beyond MSEA.

7.3.4 Later introduction of Remote Oceanic dogs

Despite the scarcity of dog remains in Lapita archaeological sites in the western Pacific, there is considerable archaeological evidence for people taking dogs with them during the colonisation of East Polynesia that began around 1,000 years ago (Wilmshurst et al., 2011). Linguistic evidence, in addition to modelling human migrations, is also informative about the nature of interactions between people and dogs during the colonisation of the Pacific (see Chapter 2). Words for items that were a common part of Austronesian, Lapita and Polynesian cultures are transmitted consistently to daughter languages, and often shared by languages across a wide area. Terms for dog however are not consistently shared across the Pacific and show considerable variability. There is no widely shared cognate for dog in either Proto-Austronesian or Proto-Oceanic (associated with the LCC), and in Oceanic languages the term for ‘dog’ is variable. By the time people were settling the islands of Polynesia the situation appears to have changed. There is a Proto-Polynesian term for dog, which alongside the terms for chicken and pig, is a unique innovation. These terms first appear in Proto-Polynesian, and regular cognates consistently appear in daughter languages, including the Polynesian outlier languages to the west (Osmond and Pawley, 2011: 242). This contrast
suggests that, by the time that the colonisation of East Polynesia commenced, the Polynesians had developed biological and behavioural innovations that resulted in a highly successful method of transport and establishment of domestic animals not seen previously in the region.

The complete and near complete ancient mitogenomes we sequenced from dogs from archaeological sites from East Polynesia (the Cook Islands, Hawaii and New Zealand), and Atafu, an atoll of Tokelau in West Polynesia, form a distinct haplogroup A2b2* separate from the A2b3* cluster of dingo dogs and the NGSD (Figure 7-2). The subgrouping in the ML tree (Figure 7-1) is only weakly supported. Ancient sequences from dog specimens from Taurama, in the Papuan Gulf of New Guinea, and the Torres Strait are also part of this Pacific haplogroup, rather than the B lineage observed in the Late Lapita dogs from earlier Papuan Gulf archaeological sites. The Taurama archaeological site dates to between 2,000 and 1,200 BP, and the dog burial was found in an archaeological context associated with Early Papuan Pottery (EPP) (Sutton et al., 2015). EPP archaeological sites appear along the Papuan coast around 2,000 BP and provide evidence of a further maritime migration of pottery-making people with a mixed fishing, hunting and horticultural economy into the area (Allen et al., 2011). Further work is necessary to explore links between EPP communities and the process of dog dispersal in New Guinea and the wider Pacific.

The appearance of dogs in much higher numbers in archaeological sites particularly across Polynesia, carrying a new genetic lineage A2b2*, suggests there may have been a significant change in the relationship between dogs and people in the Pacific around 2000 years ago. We have previously suggested that this may signal changes in technology and animal husbandry that facilitated the successful movement of dogs further into the Pacific than ever before, possibly also linked with the impetus for the colonisation of East Polynesia (see Chapter 2). Other possibilities involve a movement of new people and their dogs into the region (Addison and Matisoo-Smith, 2010) or the incorporation of new dogs into existing cultures via exchange networks.
7.3.5 Origins of Pacific dogs

Our study has identified at least three introductions of dogs to the wider Pacific region. Firstly, the appearance of dingoests and NGSDs (A2b3*), linked by close maternal genetic relationships and becoming visible in the archaeological record in Australia around 3,500 years ago. Secondly, a different mtDNA lineage is detectable in the Late Lapita dogs from archaeological sites on the south coast of New Guinea. Finally, we found evidence of a geographically widespread Pacific haplogroup (A2b2*), appearing in archaeological sites dating from the last 2,000 years. Despite tantalising evidence for the mid-Holocene dispersal of dogs resulting in the introduction of dingoasts and NGSDs to Australia and New Guinea, we are no closer to tracing the origins and possible route for this movement. Although dingoasts and NGSDs, and ancient and modern Taiwanese dogs, show similarities in mtDNA (this study), and modern Y chromosome markers (Sacks et al., 2013), these connections may reflect the genetic diversity of a shared ancestral population on mainland China, rather than a direct dispersal event linking the two populations.

The ancient sequences from the late Pacific clade appear, at this point, to be most similar to partial sequences obtained from dogs from Bronze Age archaeological sites in Thailand (this study). Although we were unable to obtain complete mitogenomes from the archaeological Thai specimens, we were able to examine some positions with adequate coverage that suggest that they are consistent with the A2 sub-clade (Duleba et al., 2015) (Appendix C). Compared to the wider dataset of modern dog mitogenomes, the ancient Pacific dog mitogenomes are part of the A2b2 sub-clade (Duleba et al., 2015), also shared with modern dogs sampled in Thailand (Pang et al., 2009), Indonesia (Cairns, 2014) and Fiji (Duleba et al., 2015). This indicates that dogs from the Pacific clade may have ultimately been derived from MSEA dogs, rather than moving down from Taiwan during the Austronesian expansion. A similar south western route has been suggested for pigs (Larson et al., 2007). High frequencies of two short control region haplotypes, found in ancient Polynesian dogs, have also been observed in modern Indonesian dogs (Oskarsson et al., 2012). It is possible that some modern village dogs in ISEA carry ancient lineages associated with the dispersal of dogs into the Pacific, but as yet there is very little comparative data for dog mitogenomes in the region.
Figure 7-3: Distribution of ancient dog lineages sequenced from MSEA, ISEA and Pacific archaeological sites. Symbol shape denotes haplogroup, red symbols indicate where complete mitogenomes were able to be produced, and blue symbols are partial sequences.

The mitogenomic sequence from the 3,000 year old dog from East Timor carries the A4’5 haplotype (after Duleba et al., 2015), which is not shared by any other ancient specimen we sequenced. This is one of the earliest recorded instances of dogs in ISEA, but to date we have found no evidence of this lineage being carried further out into the Pacific. The beginning of the Neolithic in ISEA has been described by Spriggs (2011) as a cultural ‘hot spot’, involving the spread of new languages, ideas and identities, with both migration and recruitment from local communities. It is likely that as more samples of dogs from Neolithic archaeological sites across the region become available, a more detailed picture of the history of ISEA dogs will be revealed.

7.4 Summary

Dogs were introduced to the islands of Australasia and the Pacific during human colonisation and migrations, but the timing and dispersal routes are unclear. The ancient mitogenomes sequenced from archaeological dog specimens from ISEA, Australasia and the
Pacific presented here offer a novel series of individual insights into the history of dog dispersal out from East Asia as it occurred prior to the influence of modern European dog breeds. We generated 19 complete and partial mitogenomes from ancient MSEA, ISEA and Pacific dogs, and four modern dingoes. Our results reveal levels of discontinuity and complexity in the introduction and movement of dogs, which are mirrored in the archaeological and linguistic evidence. Our molecular genetic analysis suggests that there were at least two spatially discrete introductions of the dingo to Australia in the mid-Holocene, during a period of widespread environmental, technological and cultural change. At least one other different lineage is present in ISEA around this time, observed in the dog burial from East Timor. Archaeological evidence for the introduction of dogs to the Pacific as part of the LCC is extremely limited, and suggests that while dogs may have been present in Lapita sites in Near Oceania, they were not transported in sufficient numbers to establish viable populations during the colonisation of Remote Oceania. Nonetheless, we demonstrate that mitogenomes from dogs in Late Lapita archaeological sites show affinities with Iron Age dogs from Taiwan, raising the possibility of at least one introduction of dogs during the Austronesian expansion from the north. Finally, we have identified a major late introduction of dogs across the region, culminating in the establishment of dog populations in colonisation era sites throughout East Polynesia.
Chapter 8
Discussion

This thesis set out to investigate the introduction and dispersal of dogs throughout the Pacific, using archaeological and aDNA methods based on the production of ancient mitogenomes from archaeological samples. Ancient DNA was extracted from three sets of samples from archaeological sites in Mainland and Island South East Asia, the Pacific, and sequenced using NGS technology. Each set of samples differed in terms of their geographical extent, dates and time span. The first set was obtained from a sample of fourteen dogs from a single archaeological site, that of Wairau Bar in New Zealand (Chapter 5). Wairau Bar is one of the earliest dated archaeological sites in New Zealand, and is thought to have played a significant role in the colonisation of the country (Jacomb et al., 2014). The second set of mitogenomes was obtained from archaeological dog specimens from Wairau Bar and a further four colonisation era archaeological sites from across New Zealand. This set also included two specimens from an archaeological site at Ureia, on the coast of the island of Aitutaki in the Southern Cook Islands (Chapter 6). The final set of mitogenomes was generated from a diverse range of samples from archaeological sites of different time periods from across Southeast Asia and the Pacific (Chapter 7). Because of the size and composition of the different sample sets, different phylogenetic methods were used for the analyses.

In this chapter, I review the outcomes of the research in relation to the three research questions posed in the first chapter. 1) I evaluate the utility of ancient mitogenomes to investigate genetic relationships and ancestry at an archaeological site, island or archipelago level; 2) I examine how the results of the aDNA analyses articulate with archaeological models for the movements of people and animals in the region; and 3) I consider the ways in which aDNA results can contribute towards understanding dog-human mobility and interactions in the Pacific.
8.1 Genetic variation in ancient mitogenomes of Pacific dogs and the investigation of genetic relationships and ancestry

This section responds to my first research question: Do complete mitogenomes provide sufficient variation to investigate genetic relationships and ancestry at an archaeological site, island or archipelago level? The three sample sets used in this thesis form a sliding scale in terms of the space and time of the underlying archaeological assemblages, thus enabling this question to be addressed. The Wairau Bar sample is tightly constrained to a particular time of only a few hundred years (if that) and a single place (Jacomb et al., 2014). The New Zealand colonisation era and Southern Cook Island sample set spans a similar time frame, but sample locations are separated by over 3,000 kilometres. In extreme contrast, the Asia Pacific sample has a time depth of over 3,000 years and sample locations span 10,000 kilometres. In addition to sample selection, the usefulness of aDNA data generated during this research also relates to sequence recovery and authenticity, genetic diversity, and the outcomes of molecular genetic analyses.

This section begins with a discussion of the nature of the sequences that were generated during this research, then looks at the genetic variation between specimens, archaeological sites, islands and island groups. This leads to an evaluation of the results in relation to the utility of the sequences to investigate genetic relationships and ancestry at different scales.

8.1.1 Sequence recovery and authenticity

One of the first challenges in aDNA research is obtaining sufficient authentic DNA from specimens to enable molecular genetic analyses to be undertaken. The aDNA in-solution silica extraction and double-stranded capture protocol used for this research was developed in the Ancient DNA laboratory at the University of Otago (Chapter 5), in accordance with published protocols being used at the time. Strict procedures were followed to control contamination and authenticate aDNA, including laboratory-based measures to avoid contamination with exogenous DNA, and bioinformatic procedures to check for contamination, evaluate reads for characteristic patterns of aDNA damage, and generate sequences in accordance with appropriate quality control measures. In this fast moving field,
over the course of this thesis, new techniques have been proposed to enable more efficient extraction and sequencing of aDNA, such as silica column extraction (Dabney et al., 2013), single-stranded library preparation (Gansauge and Meyer, 2013), target enrichment (Cruz-Davalos et al., 2016) and whole-genome capture methods (Ávila-Arcos et al., 2015).

The initial study presented in Chapter 5 demonstrated that DNA was present in the dog teeth from Wairau Bar and could be extracted and sequenced to produce complete mitogenomes. Complete mitogenome sequences were able to be obtained for three specimens, and near complete sequences for another 11 specimens (greater than 95% of the reference genome covered by a read depth of greater than two). This excellent level of DNA recovery differs from the mtDNA sequences obtained from four of the Wairau Bar human burials (Knapp et al., 2012b), however, the dog sample was excavated in 2009, whilst the human burials were excavated over 50 years ago and held in museum storage prior to aDNA analysis. The protocol used for the human burials involved 454 sequencing, while the Wairau Bar dogs were sequenced using Illumina, so the technology used may have also played a part in the differing coverage and read depth of the two sample sets.

The protocol was then applied to samples from colonisation era archaeological sites across New Zealand and a comparable archaeological site on Aitutaki, and achieved similar results in terms of coverage and read depth (Chapter 6). Some of the samples had been excavated up to 40 years previously, however, mitogenomes were able to be sequenced and authenticated.

Finally, the protocol was applied to a variety of samples excavated over the last fifty years from Southeast Asian and Pacific archaeological sites. The recovery rate of aDNA from some of the tropical Pacific specimens differed considerably from the New Zealand samples. This was not unexpected, as problems with DNA preservation in warm and wet environments are a recognised limitation of aDNA in tropical areas (Matisoo-Smith and Horsburgh, 2012). Nonetheless, the near complete mitogenome generated from a specimen of the dog buried at Matja Kuru cave in East Timor demonstrates the preservation that is possible in cave environments in the region (as predicted by Hofreiter et al., 2015). Only partial coverage was
able to be obtained from some of the oldest dog specimens from the Late Lapita sites on the south coast of Papua New Guinea, but some diagnostic point mutations (SNPs) were able to be observed which meant that some high level indication of haplotypes was possible.

8.1.2 Genetic diversity

The analytical value of previous studies of mtDNA from archaeological dog samples from New Zealand and tropical East Polynesia was limited by the homogeneity of the sequences observed in a small section of the control region. Two haplotypes, Arc1 and Arc2, were able to be distinguished from the initial aDNA study of archaeological dog remains from Polynesia (Savolainen et al., 2004). The study used a 263 bp sequence from the control region and compared the ancient haplotypes with sequences from modern village dogs. The ancient haplotypes were detected in modern dogs from China, Mainland and Island SEA, but the study lacked sufficient resolution to be able to address questions about population structure and founding events.

One of the main drivers for this thesis was to investigate if complete mtDNA sequences from Pacific dogs possess sufficiently greater levels of genetic diversity to enable differentiation between different specimens, archaeological sites and island groups. If so, this would provide a more refined dataset to address questions about dog movements and human interactions across the region. Analysis of the complete mitogenomes obtained in this research did enable a greater level of discrimination between mtDNA haplotypes (discussed in Chapters 5 and 6). But similar problems stemming from a lack of differentiation between specimens were still apparent in the analysis of mtDNA from New Zealand dogs, despite targeting the complete mitogenome. The results clearly demonstrate the influence of genetic and demographic history in the mtDNA of Pacific dogs, and the resulting utility of mitogenomes to address particular research questions.

The mtDNA network produced for the colonisation era New Zealand dogs (presented in Chapter 6) was very similar to that produced for the Wairau Bar dogs alone (Chapter 5), and shows that all the sampled New Zealand dogs carried extremely similar mtDNA haplotypes, only varying by three point mutations (SNPs) across the network. As dog bones are
numerous in New Zealand colonisation era archaeological sites, specimens for aDNA sequencing were able to be selected from the same element. This meant that individual dogs were only sequenced once. It is possible, however, that they could be related e.g. mother and offspring. Given the haplotype network produced, even if this was the case, it is unlikely to significantly alter the overall result.

The sequencing of mitogenomes did enable a greater level of discrimination than that previously obtained using very short control region fragments. But the lack of diversity in the founding population of New Zealand dogs observed in the mitogenomes, in combination with the relatively recent history of dogs in the Pacific, meant that there was insufficient variation to differentiate the population on the basis of mtDNA haplotypes. As a result, it was not possible to examine relationships between dogs from different archaeological sites from the colonisation era. Because of this limited genetic diversity, adding additional numbers of samples from across New Zealand from similar time periods to the Wairau Bar dogs did not alter this pattern. Sequencing autosomal nuclear markers could overcome problems with distinguishing closely related individuals and hence provide a greater level of detail about lineage sharing between archaeological sites, which could in turn be informative about the connections between human communities.

Although the lack of genetic diversity prevented inter-archaeological site comparisons, the homogeneity of the mtDNA observed in New Zealand colonisation era dogs is nonetheless informative, suggesting founders with very limited genetic diversity. The variation between the two mitogenomes from the Southern Cook Island specimens and the New Zealand and Hawaiian samples suggests that sequencing samples from across central East Polynesia may be able to provide further information about the possible source of the founding population of New Zealand dogs. At this inter-island level, mitogenomes have the potential to provide valuable information about connections between New Zealand and other parts of Polynesia.

In comparison to the New Zealand sample, ancient mitogenomes obtained from dog bones from archaeological sites in Southeast Asia and the Pacific show much greater levels of divergence. Several of the sample locations in the Pacific also contained much less dog bone
compared to the large amounts found in early archaeological sites in New Zealand (Davidson, 1987). In two archaeological sites, Matja Kuru in East Timor (Gonzalez et al., 2013), and Taurama in New Guinea (Bulmer, 2001), samples were obtained from individual dog burials, rather than midden contexts. Infrequent dog bone was also present in the Late Lapita archaeological sites on the South Coast of New Guinea (McNiven et al., 2012), and only partial aDNA sequences were able to be generated.

Individual specimens of this nature cannot therefore be considered representative of the dog population that could have been present in associated human settlements in the past. Rather, they are an individual snapshot of the genetic makeup of one or two dogs at a particular place at a particular point in time. There does seem to be temporal and geographical trends in mtDNA diversity in the dogs sampled, with decreasing diversity moving eastward paralleled by the increasing ages of the archaeological sites. This trend is consistent with studies of global domestication and dispersal, where the region south of the Yangtze River possesses the greatest mitochondrial diversity observed in modern dogs, and this diversity decreases in relation to the distance from this centre (Pang et al., 2009). The use of modern data to address questions about the past does however need to be undertaken with caution, as modern genetic pattern may not be representative of past events (Storey et al., 2012, Leonardi et al., 2016). Nonetheless, it seems reasonable that levels of diversity may decrease in this way.

The longer time scale, the greater geographical spread, and apparent differences in demographic history of the dog samples, are all likely to have contributed to the greater divergence between samples in the west. How this divergence may relate to patterns of human mobility, and dog human interaction, is discussed in more detail in sections 8.2 and 8.3 below.

8.1.3 Phylogenetic, population genetic and phylogeographic analyses

The composition of each of the three sample sets (Wairau Bar, New Zealand colonisation era and Southern Cook Islands, and the South East Asia and Pacific region), meant that several different approaches were used to analyse the mtDNA data generated. Networks and trees
were used to infer phylogenies for each sample set. Median joining haplotype networks using Popart (v1.7.1) (http://popart.otago.ac.nz) were produced to visualise phylogenies for all the sample sets, and examine the relationships between haplotypes. Tajima’s D was also calculated using Popart. Maximum likelihood trees were produced using MEGA (Tamura et al., 2013) and MEGA7 (Kumar et al., 2016), and Bayesian trees were generated with Beast (v1.8.2) (Drummond et al., 2012) for the colonisation era New Zealand and the Southeast Asian and Pacific sample sets, to investigate if different tree-based methods were in agreement with the results of the network method, and to see whether Bayesian methods were informative about the timing of the appearance of particular lineages.

At the outset of this study, it was hoped that an increased sample size, along with the use of the whole mitogenome, as opposed to a small portion of the control region, would enable population genetic analyses to be carried out at Wairau Bar, and in the sample set from across New Zealand. The results could then be used to investigate the genetic structure of the founding dog population in New Zealand, and to identify and explore possible connections between human communities during the colonisation era. From a phylogeographic perspective, the shape of the haplotype network was consistent with the outcome of a recent population expansion from a limited number of founders (Avise, 2000). But the extremely limited genetic diversity of the New Zealand dogs constrained any further use of population genetic analyses, such as statistics or Bayesian tree building methods.

The result from the mtDNA analysis of the Wairau Bar dogs contrasts with those achieved from population genetic analysis of aDNA from a number of indigenous bird and marine mammal taxa in New Zealand, where the structuring of the genetic diversity present in natural population has enabled the identification of processes of extinction and replacement (e.g. Grosser et al., 2016, Rawlence et al., 2015, Rawlence et al., 2014). This highlights one of the challenges of working with domesticated populations with reduced genetic diversity, and demonstrates the potential impact of human actions on the genetic makeup of these populations.
The observed genetic diversity in the two sequences from Shi San Han in Taiwan, with two different major dog haplogroups represented in only two specimens, indicates a far greater genetic diversity may have been present in the dog population there and that a much greater sample size would be required to capture the diversity present. This in turn may make mtDNA population genetic analyses more feasible. The use of nuclear markers would also assist with investigating population structure.

The next section of this chapter examines the results of the phylogenetic analyses in relation to archaeological models for colonisation in Pacific.

8.2 Phylogenetic analyses compared with archaeological models for human colonisation in the Pacific

This section addresses my second research question: How do the outcomes of phylogenetic analyses compare to archaeological models for colonisation era human and animal mobility and interactions in Oceania and the Pacific? The production of ancient mitogenomes for this thesis were targeted towards two particular periods of colonisation that took place during the movement of people across the Pacific. Firstly, the colonisation of New Zealand, one of the last large land masses in the world settled by people, at around 1250 AD (Wilmshurst et al., 2011), and secondly, the appearance and spread of the LCC, which extended from the Bismarck Archipelago to West Polynesia by about 3,000 years ago (Kirch, 2000). Although the results from the analysis of the ancient mitogenomes are quite different from each other, they are consistent with the overall shape of each respective archaeological model. Questions do arise where some aspects of the mitogenomic analyses are inconsistent with the archaeological model, or highlight parts of the model that have not be elaborated in sufficient detail to explain the mechanisms involved in some of the findings. In this section I look firstly at the colonisation of New Zealand, then at the emergence and expansion of the LCC, followed by a discussion of alternative archaeological models that may also be useful for examining the movement of dogs in the Pacific.
8.2.1 New Zealand

Refinements in radiocarbon dating have driven a reappraisal of the dates for the colonisation of New Zealand, commencing with Anderson’s (1991) review. The beginning of the colonisation period, marked primarily by the presence of a distinctive artefact assemblage with East Polynesian affinities and a suite of extinct fauna, is now thought to have occurred no earlier than 1200 AD (Wilmshurst et al., 2011). Native biota, such as moa, were hunted to extinction within as little as two centuries following the arrival of people (Holdaway et al., 2014). The number and distribution of archaeological sites from this colonisation period, spread from one end of the country to the other, suggests a highly mobile population of several hundred people was required to account for this archaeological site distribution (Weisler and Walter, 2017).

The genetic composition of dog mitogenomes from New Zealand colonisation era archaeological sites is strikingly similar. This homogeneity supports the hypothesis of a mobile human colonising population which established breeding groups of dogs from the same founding stock in the settlements they established throughout the country during the colonisation process. Although the limited mtDNA variation precluded the identification of possible interactions between people from different colonisation era communities during the settlement process, the pattern of genetic diversity, or lack thereof, does provide some information about the demographic history of Polynesian dogs prior to the colonisation of New Zealand.

The mitogenome sequences from dogs from the Ureia archaeological site in the Southern Cook Islands are closely related to those of the New Zealand and Hawaiian dogs. It is possible that the lineages from these island groups were part of a large source population extending across the wider Polynesian homeland zone. Alternatively, the New Zealand haplotypes may have a more restricted geographic distribution and hence be informative about possible departure points of the New Zealand colonists. Further analysis of dogs from additional archaeological sites in East Polynesia is required to explore possible origins in more detail.
The other two Pacific domesticates, the pig and chicken, are absent in pre-contact New Zealand archaeological sites (Davidson, 1987), so it is not possible to determine if the patterns of genetic variation observed in New Zealand dogs are similar to the other domestic introductions. Whether the lack of pigs and chickens in New Zealand archaeological sites represents a failed attempt at transporting or establishing these animals, or that they weren’t brought at all, is not known. Other East Polynesian islands and island groups however do have archaeological evidence for the introduction of all three domesticated animals (Kirch, 2000), and the implications of this are discussed in more detail in the sections below.

8.2.2 Lapita Cultural Complex

The second period of colonisation investigated in this thesis research relates to the appearance and spread of the LCC. Archaeological evidence for the LCC is distributed over a far greater distance than that involved in the colonisation of New Zealand. In addition, the colonisation process during the LCC involved the movement of new groups of people speaking different languages into Near Oceania where people had already been living for thousands of years, followed by the colonisation of islands in Remote Oceania that were previously uninhabited (Kirch, 2000).

The complexity of the LCC is apparent in Green’s Triple I model which provides for processes of intrusion, integration and innovation to explain the appearance of a distinctive set of archaeological evidence first appearing around 3,500 years ago in the Bismarck Archipelago (Green, 1991b, 2003). The movement of Austronesian language speakers southwards into Near Oceania from Taiwan, via the Philippines and ISEA and out into Near then Remote Oceania represents the intrusion of a new culture in the region. The Lapita peoples integrated aspects of the material culture already present in the region, and also developed technological innovations during their migrations. Green (2003) further elaborates his model with the addition of a ‘voyaging corridor’ that recognises the movement of people, languages and ideas through the islands of Wallacea, that link ISEA with Near Oceania.
Two of the ‘intrusive’ characteristics of the LCC observed in the archaeological record—agricultural practices and domesticated plants and animals—are associated with the concept of the transported landscape (Kirch, 2000). This concept describes a process where migrating groups moved across the Pacific taking with them a set of horticultural concepts, plants and animals that were then utilised on the new islands they settled. Lapita peoples were not the first people to move animals around in the Pacific, with translocation of some marsupials occurring in Near Oceania in the Pleistocene (Allen et al., 1989). However, the movement of domesticated animals (dogs, pigs and chickens) into Remote Oceania is largely, if not exclusively, associated with the LCC.

The mitogenomes obtained from the sample of archaeological dog teeth and bones from across the ISEA and the Pacific do appear broadly congruent with Green’s VC Triple I model. In comparison with the closely related mtDNA of ancient dogs from East Polynesia presented in Chapters 5 and 6, as a group the mitogenomes obtained from ancient dogs from ISEA and Papua New Guinea archaeological sites possess much greater genetic diversity, although they also span a much larger geographical area and time period. This suggests multiple introductions of dogs into the area, although archaeological evidence from Neolithic sites is scarce, so the viability of these early introductions is not clear. At least one lineage carried by the 3,000 year old dog buried at Matja Kuru (East Timor) (Gonzalez et al., 2013) has not yet been observed in any other dogs any further eastwards in the Pacific.

Although only partial coverage from two Late Lapita dogs was able to be generated (Chapter 7), analyses suggests that these dogs may carry B haplotypes. A B haplotype was also generated from a dog from the Shi San Han Iron Age archaeological site in Taiwan. This was a surprising finding, as the possible presence of a B lineage found in both Taiwan and late Lapita archaeological sites is the first ancient molecular genetic evidence for the introduction of dogs into the region by Lapita peoples with links to Taiwan. This northern link is consistent with Sacks and colleagues (Sacks et al., 2013) observation of a possible ancestral connection between modern Taiwanese dogs and Australian dingoes based on Y chromosome markers. In addition, studies of mtDNA control region and complete
mitogenomes indicate a southern ISEA route for the introduction of dogs to East Polynesia carrying A haplotypes.

This lack of continuity between Late Lapita and East Polynesian dog lineages appears inconsistent with the expectations of the VC Triple I model (Green, 2003). Kirch and Green contend that Lapita peoples are ancestral to the Polynesians, and that distinctive Polynesian culture arose in West Polynesia from Lapita antecedents (Kirch and Green, 2001). Recently Matisoo-Smith and Addison (2010) have pointed out that there are aspects of Polynesian bioanthropology and archaeology that are inconsistent with a Lapita-only model for Polynesian origins. These include differences in human skeletal morphology, subsistence and material culture. In addition, they suggest that the appearance of new lineages of dogs, chickens and rats around 2,000 years ago could be associated with a second population movement out of Asia. Matisoo-Smith and Addison (2010) propose that the colonisation of Polynesia may be more complex than initially thought, and suggest that West Polynesian ‘Triple I’ may be more appropriate than a Lapita-only model.

Given the inconsistencies between Green’s VC Triple I model and the archaeological, linguistic and genetic data for the genetic distribution of mtDNA haplotypes observed in this thesis, I now turn to another class of archaeological models that may offer an alternative framework for considering the movement of dogs in the Pacific.

8.2.3 Alternative archaeological models

Applications of the commensal model and phylogeographic studies incorporating aDNA conceptualise animals as co-migrants with people. In the Pacific, Green’s (2003) VC-Triple I model is concerned with explaining the appearance of the LCC as a cultural complex which comprises a group of people, their languages, material culture and cultural practices. Looking at the distribution of domesticated animals at this time, is it appropriate to adopt this model with a one-to-one correspondence between people and dogs, where people and dogs are co-migrants, travelling together, in a manner that means that their distribution can be modelled in the same way?
Storey and colleagues (2013) caution that the characteristics of the species being transported must be taken into account when using a commensal approach. The relationship between people and dogs is quite different to that of rats, which were the first species used in the ‘commensal model’ (Matisoo-Smith, 2009). Rats were used as a proxy for people because of specific characteristics that meant that rat phylogenies were considered likely to mirror initial human colonisation events (Matisoo-Smith, 1994, 1998).

In Chapters 2 and 3, my review of the archaeological evidence, ethnographic and contemporary accounts about village dogs has demonstrated the diverse ways in which people and dogs may have interacted in the region over several thousands of years. Given that dogs do not appear to have formed feral populations beyond the large islands of New Guinea and Australia, are highly dependent on people for survival in small island contexts, and may have been extirpated and introduced more than once to islands, it may not be all that useful to draw on the same models as those developed for people (and rats) to explain dog introduction and dispersal in ISEA and the Pacific.

As a complicating factor, dogs may have been exchanged while alive, but also on their deaths as either food or ornaments. Although there has been a great deal of interest in the role of pigs in social exchange and gifting, particularly in Near Oceania, the corresponding role of dogs has received much less attention—with the notable exceptions of Luomala (1960b) and Titcomb (1969). Archaeologists have also been interested in the presence of pigs, as they are seen as intimately bound up with agricultural intensification. Dogs however may also be a part of this type of exchange, as documented by Chowning (1991). Some of the earliest evidence for dogs in Lapita archaeological sites is in the form of drilled teeth (Matisoo-Smith, 2007) which may be evidence of the exchange of valuables, rather than living dogs.

It is possible that archaeological models developed to explain the exchange of material culture, that is the movement of things rather than people, may provide an alternative source of information for investigating aspects of the processes by which dogs may have been dispersed throughout the Pacific. These models attempt not only to identify instances of exchange but to understand the underlying social and economic structures and processes
involved. The redistribution of prestige items in Polynesia, for example, may be associated with the emergence of chiefly elites (Kirch, 1989, Clark et al., 2014).

Archaeological models of trade and exchange in the Pacific have been developed primarily for stone and ceramic items, as these durable materials usually survive well in the archaeological record (Summerhayes, 2008). Chemical characterisation is carried out to identify groups of items with similar chemical signatures, and if possible, to determine a source location. Finding a material’s source location requires field survey, sampling and characterisation of possible sources. The distribution of an item from its source to where it ended up, and changes to this pattern, can be used to model the kind of exchange that took place. For example, obsidian from the Bismarck Archipelago in Near Oceania has been found in archaeological sites dating to around 3,000 years across a distance of over 6,500 km, from Borneo in the west to Fiji in the east (Summerhayes, 2008). Conversely, the raw materials used to manufacture Lapita ceramics have been shown to be predominantly of local origin, however, the motifs that decorate the pots appear to be part of an aesthetic tradition that also spanned several thousand kilometres (Summerhayes, 2008).

When a commodity is only available in a specific location, generally it is more abundant in archaeological sites closest to its source, and its frequency will decrease with distance. If the frequency of an item is plotted against the distance from the source a ‘fall-off curve’ can be produced (Renfrew, 1977). Different modes of exchange produce different shaped fall-off curves. Several modes of exchange have been identified such as down-the-line, reciprocity and central place. In the Pacific, two different modes of exchange of obsidian can be observed (Summerhayes, 2008). Between the late Pleistocene and mid-Holocene the distribution of obsidian shows an exponential fall-off in abundance away from the source, and fits a ‘down-the-line’ mode of exchange, between semi-sedentary communities. But during the first phase of colonisation by the Lapita peoples the nature of the exchange differed. Although still decreasing in frequency from the source, the composition of the early LCC obsidian assemblages were different, as large and relatively unworked pieces were present in early sites. This pattern is linked to a formal exchange network, where exchange is an adaptive
mechanism which acts as a ‘lifeline’ for colonists back to a homeland and as a means of maintaining social ties (Summerhayes, 2008).

Walter and colleagues (2010) use two modes of exchange based on studies of obsidian distribution on the Papuan Coast and Bismarck Archipelago (Irwin, 1991, Specht, 2002) to examine early period obsidian and later greenstone exchange in New Zealand. The ‘coloniser’ mode, as described above, involved a pulse of exotic items during a sustained phase of colonisation of New Zealand. In this mode, the movement of exotic items may have been a side effect of the contacts maintained between colonising groups. The ‘trader’ mode developed much later as wide-spread trading network for finished adzes.

The distribution of dogs in archaeological sites in the Pacific is not uniform, and the genetic makeup of dogs in the region is also varied. Archaeozoology, in combination with aDNA, may therefore offer a way of identifying possible ‘source’ locations for domesticated animals, and for describing patterns of distribution through time and space. These patterns could then be compared with archaeological modes of exchange. Assuming that a colonising group possessed a reasonable number of dogs and the colonisation process occurred relatively quickly, the expectations of the coloniser mode of exchange would be that there would be dogs with a similar genetic makeup to the source population found in most, if not all, new colonies. As long as contact with a ‘homeland’ or with other communities was maintained, any difficulties in establishing breeding groups of dogs could be offset by recruitment from other communities, and breeding between dogs from different communities could also occur. This pattern is apparent for dogs from the colonisation era archaeological sites in New Zealand, where dogs are plentiful and have very similar mtDNA.

This is not the pattern however for dog bone in early Lapita archaeological sites. The ‘homeland’ for the LCC is likely to have been the Bismarck Archipelago. As discussed in Chapter 2 archaeological evidence for dogs in the archipelago is slight, and decreases to the east until becoming absent altogether from Lapita sites in Vanuatu, New Caledonia and further eastward. This is unlikely to be due to taphonomic processes affecting the survival of skeletal material, as large quantities of pig bone have been recovered from the Teouma
archaeological site (Valentin et al., 2010). It may be that there were simply insufficient numbers of dogs to establish viable populations on newly colonised islands during the LCC migrations. Alternatively, did dogs arrive in Near Oceania and become incorporated into Lapita and later communities via a different mechanism?

Down-the-line and trader modes of exchange of living dogs are likely to give rise to different patterns than the coloniser mode, however, distinguishing between them may be challenging. Where dogs are being exchanged or traded, it might have been difficult to establish viable populations, so there may be low numbers of dogs in archaeological sites, and possibly evidence for multiple introductions from different source populations as people tried to maintain their dog numbers. This is more like the pattern observed in Near Oceania around 2,000 to 3,000 years ago. Depending on the source population, crossing the Near-Remote Oceanic boundary may have been the end of the line for Lapita dogs. Renfrew (1977) point out that in relation to fall-off curves, ‘distance’ means effective distance which is not necessarily the same as the distance between two points. Topography may increase the distance, and the means of transport is critical in determining distance. He goes on to state that “the development of new travel technologies …fundamentally alter effective distance” (Renfrew, 1977: 72-73). I will return to the issue of transporting dogs in section 8.3.3 below.

In addition to aDNA, stable isotopes have also been used to investigate whether domesticated animal remains from archaeological sites in the Pacific are of local or non-local origin, and the implications of this for understanding human mobility and migration. Kinaston and colleagues (2013) study of isotopes from dogs from the Wairau Bar archaeological site found that all the dog samples had local signatures, in comparison to several of the human burials which had non-local signatures. Shaw and colleagues (2009) identified a non-local Lapita era pig in their strontium isotope analysis of human and pig teeth from the archaeological site of Kamgot, in the Anir Islands of the Bismarck Archipelago. They also questioned whether pig mobility could be used as a proxy for human migration, and suggest that rather than migration per se, human mobility might be more appropriately indirectly inferred from pig mobility. They also suggest that there is sufficient
strontium isotope variation between Pacific islands to facilitate the use of this isotope to reconstructing human migration in the western Pacific. Future studies combining stable isotopes and aDNA analyses may be informative about patterns of exchange of domesticated animals in the Pacific.

One of the issues with using molecular phylogenies to inform archaeological models is apparent here. As discussed above, dogs (and pigs) may be moved around by people and traded between different groups. These movements may relate more to the localised mobility of human groups for a variety of purposes, rather than large scale processes of migration. While large population-based studies may be able to characterise the genetic make-up of dogs present at a particular place and time, a small sample may not adequately capture the lineages present, or conflate those from different time periods. Furthermore, a one to one correspondence between particular mtDNA dog haplotypes and human groups, where dogs are only moved in association with particular peoples, in relation to migration may not necessarily accurately represent the ways in which dogs may have been incorporated into human communities, and ultimately archaeological sites. The links between dogs and human communities therefore need to be explicitly considered when considering dog and human movements. The following section addresses how this research contributes to understanding such links.

8.3 Contribution of ancient mtDNA for understanding dog and human mobility and interactions in Australasia and the Pacific

My third research question asks: Are phylogenetic patterns of dog diversity and relatedness based on mitogenomes informative about dog-human mobility and interactions in the Pacific? The comparison of the ancient mitogenome analyses and archaeological models for human colonisation (discussed in Section 8.2 above) address this question quite broadly, in terms of the contribution that the aDNA results make in relation to tracking dog mobility associated with human migrations across the region. An appraisal of the results of the ancient mitogenome analyses in relation to the archaeological models has shown that in the two colonisation periods examined, the aDNA results are largely consistent with the
archaeological models. But the aDNA results have raised some questions that would not be visible using standard archaeozoological techniques: in particular, the diversity of dog lineages in ISEA and Near Oceania, and the discontinuity between Lapita era and East Polynesian dog lineages.

When considering interactions between dogs and people it is worthwhile keeping in mind that these interactions occur within the context of a relationship between two parties. Although in Section 8.2.3 above I considered the applicability of archaeological models for trade and exchange for domesticated animals, a dog is not an artefact and is capable of influencing the world around it, its relationship with people and the archaeological record in many ways. Dog mobility involves more than simply the act of moving a dog from one location to another. The characteristics of the dogs themselves and interactions between dogs and people are an important aspect of dog mobility. In addition to providing information about dog mobility, patterns of genetic diversity observed in dog mtDNA from archaeological samples have highlighted other aspects of people’s interactions with dogs, dog ecology and behaviour. In most cases, this results in more questions than answers, but does demonstrate the potential for aDNA to contribute towards understanding dog-human interactions more generally.

This section begins with a discussion of the interactions between dogs and people particularly in relation to dog mobility, then considers possible interactions from a broader perspective. Possible changes in the relationship between people and dogs during the colonisation of the Pacific are then considered, and finally the contribution of the aDNA data on a global scale in the context of dog domestication and dispersal are explored.

8.3.1 Subject and object

The distribution of populations across geographical space has formed the basis of phylogenetic and phylogeographic studies of wild populations, commensal species such as rats, and domesticated species, including dogs. Each of these groups of animals have different genetic, behavioural and ecological characteristics that influence their relationship with people. Populations in the wild have little to no interactions with people, commensal species
may be transported by people but once introduced have little or no human intervention in their reproduction and ecology apart from inadvertently providing a domestic niche (O’Connor, 2010), while people may play an active and ongoing role in the ecology of domestic species such as dogs.

Recent trends in the humanities, including the development of interdisciplinary human-animal studies, have seen renewed engagement with questions about human-animal interactions and dynamics. Serpell (2009: 633) recognises an ambiguity that exists in relation to animals, which occupy an intermediate place between the ‘world of human subjects and the world of insensate objects or things’. Influenced by the ideas of anthropological theorists such as Tim Ingold (e.g. 1986, 1994) with regard to relationality, some archaeologists now recognise that animals may have played subjective, agential roles in past societies, rather than being reconstructed as objects, such as food resources, prey or a source of raw materials (Hill, 2013). For example, Orton (2010) in his study of cattle in the Neolithic, argues that they are ‘sentient property’, despite being objects within property relations between people, they are also subjects with their own social world that overlaps with that of people.

The philosopher Helen Steward argues for the following concept of agency for animals: an agent can move the whole, or at least part, of its body; it possesses some form of subjectivity; it has at least some types of rudimentary intentional state (such as trying, or wanting), and it is a “settler of matters concerning certain of the movement of its own body….attributed always first and foremost to the agent, and only secondarily to environmental impacts or triggers” (2009: 226). This concept clearly recognises the animals may be subjects which act intentionally. As archaeologists, the challenge is therefore to recognise aspects of animal agency in our reconstructions of the past, and the implications of this for understanding human-animal interactions.

A more detailed understanding of the genetic variation and relatedness of dogs at a particular place and time, for example, colonisation era sites in New Zealand, may provide insights into aspects of dog mobility and human-dog interactions, beyond reconstructions of phylogenies over large spatial and temporal scales. Dog movement in and around settlements is relevant
to a range of questions about human activities and dog behaviour, such as the access to, or provision of, foodstuffs and shelter, whether dogs are free to roam wherever they wish, and the extent, if any, to which people are exerting over dogs’ spatial movements and reproduction. Mobility is also associated with movements of animals between settlements, and this may involve deliberate human action or not. These kinds of interactions are often fundamental to archaeological interpretations of the place of animals in human societies, including domestication, animal management practices, and the functional roles of dogs. Conversely, aspects of dog behaviour unrelated to human activities may also be implicated in phylogenetic and phylogeographic patterns.

In many parts of the world dogs in free-ranging populations are associated with human habitations but are not subject to direct human control. Archaeological evidence of dog gnawing on bones in New Zealand archaeological sites and historical observations suggests that in some locations dogs were free-ranging and therefore likely to be able to interbreed freely. The genetic homogeneity of the mitogenomes from the dogs at Wairau Bar precludes the use of mtDNA to investigate patterning that could be indicative of selective or controlled breeding, but future studies incorporating nuclear markers could investigate this issue. NGSDs, still rarely observed in the New Guinea Highlands, have been argued to be a separate breed to lowland village dogs (Koler-Matznick et al., 2007). Yet it has recently been suggested that lowland dogs possess many, if not all, of the phenotypic characteristics associated with NGSDs (Dwyer and Minnegal, 2016). NGSDs are very closely related to the Australian dingo, however there is no comparative genetic data from lowland New Guinea dogs.

Despite a frequent lack of direct involvement with free-ranging dogs, interactions between dogs and people can have negative consequences, as well as positive ones. Ecological studies and ethnographic accounts both describe issues for people living with dogs, including dog bites and predation on other livestock (Chowning, 1991). Living in proximity to dogs may also have exposed people to disease and parasites where dogs either suffered from the disease or were carriers. The majority of contemporary ecological studies reviewed for this research were produced as part of investigations into the control of rabies, which is a major health
issue in many parts of the world (Wandeler et al., 1993). Given the risks inherent with keeping dogs, what was the place of dogs in human communities in the Pacific?

### 8.3.2 The place of dogs in human communities

Ethnographic and ecological studies of dog populations have demonstrated the variation in people’s attitudes towards free-ranging dogs in contemporary societies (see Chapter 3). Dogs may be incorporated into human communities in different ways, with a free-ranging village population providing individual dogs for particular purposes, such as hunting, watch dogs, companions, or as a source of food for everyday or ceremonial occasions. Some aspects of these interactions may be detectable in the archaeological record, but we cannot assume that people’s interactions with dogs today will be the same as those in the past. It is likely too that peoples’ relationships with dogs over time were subject to change.

One of the oldest forms of interaction between people and dogs that is detectable in the archaeological record is one arising from death, rather than life. The act of burial may relate to a bond between the animal and people, or the use of the dog for ceremonial or ritual purposes. Morey (2010) goes so far as to argue that the bond between people and dogs was a critical factor in the domestication process. Deliberate burial of dogs has taken place around the world, with the oldest archaeological instances dating to over 12,000 years ago (Morey 2010). The deliberate burial of dogs (and pigs) has also been recorded across Polynesia (Clark, 1996). Two of the specimens used for aDNA analysis in this study came from complete dog burials, from the Matja Kuru cave in East Timor (Gonzalez et al., 2013), and from under the house floor at Taurama on the South Coast of Papua New Guinea (Bulmer 1975, 1979). While it is not clear whether the Matja Kuru dog was deliberately buried, this dog carried a mtDNA lineage that has not been found in other ancient specimen sequenced during this research. The dog at Taurama appears to have been buried deliberately under a house floor, and has the Pacific lineage observed throughout Oceanic sites.

While contemporary Western cultures do not consider dog to be a food animal, in some parts of the world this is not the case. The practice is found in parts of Asia, including South Korea, China, Vietnam, Cambodia and Thailand (Podberscek, 2009, Wu, 2002), Africa and
parts of Polynesia (Serpell, 1995b). Evidence for this practice is also found in the archaeological record in ISEA, and out into Remote Oceania. Butchery marks on dog bones have been recorded from archaeological assemblage from Philippines (Amano et al., 2013). The presence of dog remains in archaeological sites in Remote Oceania, particularly East Polynesia, is often viewed from a dietary perspective, as these are often recovered from midden contexts associated with food preparation and disposal. As well as dogs being used as food, they were also a source of raw materials. The three Pacific domesticates are frequently conceptualised as part of the set of animal resources exploited by people. There can be a bias towards viewing domesticated animals as products, derived from their consumption as subsistence resources (Bogucki, 1993), rather than as a living part of the domestic environment, interacting with people and the world around them.

Contemporary Aboriginal attitudes in Australia do not view dogs as a food source, and there is little or no evidence for the consumption of dingoes in Australian archaeological sites. Ethnographic evidence from Melanesia indicates that in the recent past dogs were also not considered an appropriate food, which Bulmer (2001) draws on to explain the limited numbers of archaeological dog bones found in Lapita and New Guinea archaeological sites. Dog teeth were however perforated and used to make necklaces and anklets, and are also recorded in ethnographies as minor valuables (e.g. Chowning, 1991).

Although dogs may have been valued in some parts the Pacific, the complete absence of dogs in published archaeological contexts in Vanuatu and New Caledonia is puzzling, especially given the large numbers of pigs present from colonisation era sites such as Teouma onwards (Valentin et al., 2010). Given the numbers of pig bones and other faunal remains at Teouma, it would be expected that if dogs had been present their bones would have also survived in the archaeological site. In Section 8.2.3 above, I suggested that perhaps dogs were not easy to obtain and maintain during the LCC colonisation phase in these parts of Remote Oceania, and I now consider the implications of this in more detail.
8.3.3 An innovation in animal management practices?

One of the challenges that people faced in the colonisation of the islands of the Pacific was the changes in biogeography moving through ISEA, beyond the large and inter-visible islands of the Bismarck and Solomon archipelagos, and out into island Oceania. The archaeological and linguistic data (reviewed in Chapter 2) suggests that moving domesticated animals, particularly dogs, across the longer water barriers and successfully establishing them on the smaller and increasingly depauperate islands may have been a difficult task initially. Although the presence of domesticated animals, including dogs, in colonisation era Lapita sites is the subject of ongoing debate, there is archaeological evidence for low numbers of domesticates in western Lapita sites, and greater numbers of pigs and chickens do appear just across the border to Remote Oceania in early sites Reef Santa Cruz and Vanuatu. In Fiji and West Polynesia pigs and dogs appear to be later introductions (Anderson, 2009).

By the time people were moving into central East Polynesia around 1500 years later, they appear to have acquired all three domesticated animals and mastered their transportation and establishment. This seems counter-intuitive, the further east one travels the more difficult the environmental conditions become: The islands are smaller, the distances between them greater, and the terrestrial biodiversity decreases significantly. Yet despite these more challenging conditions, colonisation era archaeological sites in central East Polynesia often contain dogs, pigs and chickens. This suggests that a major innovation or change in people’s interactions with these animals occurred prior to East Polynesian colonisation.

This is likely to have involved two aspects, human (and possibly animal) behavioural adaptations, and biological adaptations in animals. There are cultural and biological implications for this innovation: In terms of human behaviour did people invest increasingly more energy into the transport and establishment of animals, given the increasingly impoverished environment? We know very little though about animal management during Lapita colonisation and subsequently. Ethnographic accounts of extirpation of pigs and the
loss of dogs does suggest some of the challenges of keeping animals on islands with limited resources (Giovas, 2006). From a biological perspective, did the island environment act as a selective pressure on the animals themselves? Did this result in the survival of hardier animals better adapted to island life? Did people also select particular animals with desirable behaviours or characteristics? If people were selecting animals for phenotypic traits, this could influence the mtDNA genetic diversity of these animals and their offspring.

It is possible that the animal management practices involved in the dispersal of domesticated animals across the Pacific were developed as part of an ongoing process that took place during the colonisation of the Pacific. Although the start may have been during the Lapita expansion, one of the major jumps in innovation appears to have taken place in West Polynesia, just prior to the colonisation of East Polynesia and the Polynesian outliers. If so, this innovation appears to be tied up with the set of circumstances that led to the major migratory push out into East Polynesia that occurred around 1000 BP.

Explanations for the pause between the colonisation of West and East Polynesia are varied. Anderson (2013) identifies the main factors as pressure from population growth in West Polynesia or from further west via Micronesia, the appearance of double-hulled sailing canoes, and climatic changes related to El Nino weather patterns affecting both sailing conditions and horticultural productivity. Doubled-hulled sailing canoes may have facilitated the successful transport of domesticated animals given their storage capacity and swiftness between island groups.

### 8.3.4 Dog domestication and dispersal

The dispersal of dogs across ISEA and Oceania, and ultimately to New Zealand, represents one endpoint of a trajectory beginning with the domestication of wolves, probably in southern East Asia at the end of the Pleistocene (Savolainen et al., 2002, Pang et al., 2009, Wang et al., 2016). Molecular genetic analyses continue to unravel the expansion of dogs beyond southern East Asia based on genomic approaches sampling modern dogs and wolves. Most recently, using whole genomes Wang and colleagues (2016) argue for the divergence of dogs around 33,000 years ago with an expansion beginning around 15,000 years ago to
the Middle East, Africa and Asia. This process however is yet to be documented in the archaeological record.

What happened next during the Holocene is of particular relevance to tracing the shared history of dogs and people in ISEA and Oceania. Recent studies have shown that dogs possess a greater number of copies of the amylase gene, involved in starch digestion, than wolves and other wild canids (Freedman et al., 2014). This shift in copy number occurs at the same time as a similar adaptation in humans. The shift from hunting to food production may have also had significant impact on the relationships between dogs and people. Perhaps no longer as valuable as a hunting aid, or associated with communities with decreasing amounts of time for hunting, dogs themselves become another source of food.

Dog remains have been excavated from archaeological sites in China between 8,000 to 10,000 years ago, where people were subsisting on a mixed economy of hunting and cultivated plants (Chi and Hung, 2012). Deliberate dog burials have also been excavated, and the importance of dogs in these communities as hunting aides has been suggested (Liu and Chen, 2012). By the mid-Holocene, communities in China and Vietnam were adopting food production technologies, and these innovations began spreading into Southeast Asia. The remains of dogs begin to appear in archaeological sites in Vietnam and Thailand around 4,000 years ago (see Chapter 3). In the archaeological site of An Son in Vietnam, dog bones have evidence of butchery marks consistent with being a food item (Bellwood et al., 2011). Similar evidence has been recorded from Nagasabaran in the Philippines (Amano et al., 2013). The LCC expansion into Oceania has also been associated with the spread of people with food producing economies (Bellwood and Diamond, 2005). However, without comparative ancient genomes, it is difficult to place the ISEA and Pacific dog sequences generated by this research within the broad geographical parameters of this expansion.

The timing for the introduction of dingoes to Australia around 4,000 years ago based on archaeological evidence places their introduction around the same time as dogs were being incorporated into archaeological sites in MSEA. Yet how dogs reached Australia at this time remains unclear. MtDNA from dogs from Late Lapita archaeological sites associated with
the movement of food producing Lapita peoples southwards from the Philippines, that were produced as part of this thesis, indicate quite different mtDNA lineages to those carried by dingoes.

Two thousand years later there is archaeological and genetic evidence for another significant introduction and expansion of dogs across the Pacific. This thesis research has documented dogs with a ‘Pacific’ lineage in Papua New Guinea which is also spread across Remote Oceania, ultimately reaching the island groups at the periphery of the Polynesian triangle. The two partial, ancient Thai sequences are similar to this group. This route from MSEA via Indonesia into the Pacific parallels the findings for pigs and rats, based on genetic and morphometric studies (Larson et al., 2007, 2010, Matisoo-Smith and Robins, 2004).

The combination of aDNA and archaeology presented in this thesis therefore contributes to the wider picture of dog domestication and dispersal, in addition to providing new data to inform current archaeological models about colonisation in the Pacific. In the next chapter, I conclude with a summary of my findings, discuss the main limitations of the research and finally make some suggestions for future research.
Chapter 9
Conclusion

Dogs are one of the most successful domesticated animal species, and have spread with people across the world. Dogs are generalised scavengers, and have proved highly adaptable to different environments associated with human communities. The dispersal of dogs across the Pacific region is inseparably linked to the relationship between dogs and people. Unlike movement across continental landmasses, dogs must have been transported by people across the water barrier that separate islands. Despite a patchy and discontinuous distribution in the archaeological record, by the time European explorers reached the Pacific region, dogs had accompanied people to the furthest points of Remote Oceania, including New Zealand.

This research has contributed novel information derived from the production and analysis of ancient dog mitogenomes, about the origins, introductions and dispersal of dogs throughout the Pacific region. Previous studies of a small portion of the mitochondrial genome obtained from archaeological dog specimens and modern village dogs had identified two haplotypes found in East Polynesia that were also present in ISEA modern dogs, indicating a dispersal route through south western ISEA (Oskarsson et al., 2012). As a result of the data generated by this thesis research new haplotypes have been observed in archaeological specimens, with implications for understanding the human colonisation of the region and the means by which dogs may have been transported into the Pacific.

This thesis has focused on two periods associated with human colonisation in the Pacific which are well documented in the archaeological record. Firstly, the appearance of the LCC spread across the islands of Near and Remote Oceania 3500 years ago (Kirch, 2000), and secondly, the more recent colonisation of New Zealand by East Polynesians around 700 years ago (Jacomb et al., 2014). In addition to archaeological and aDNA analyses, ethnographic and ecological studies of dog-human interactions have also been drawn upon to enable a consideration of the ways in which dog behaviour and the context of archaeological specimens may influence the outcomes of the research.
It is likely that most Pacific dogs formed free-ranging populations associated with human communities. Individual dogs may have been selected as companions, for hunting or for food. There may have been little, if any, control exerted over dogs’ mobility or reproduction in these circumstances. The carrying capacity of these free-ranging dogs is closely related to that of the human population, which may be a contributing factor to the numbers of dogs found in Pacific archaeological sites.

Samples from archaeological sites, however, may differ from those from a natural population. Individuals may not be identifiable and may be represented by multiple bones. Archaeological assemblages may be the result of cultural practices that bias particular age or sex classes. They are likely to be a subset of the animal bones deposited in the archaeological site, and may have been subject to a range of taphonomic processes. Furthermore, the assemblage may span a time period of several hundred years.

By specifically focusing on two colonisation processes it has been possible to explore the ways in which the aDNA results may be in agreement, contradict or offer new information about the current archaeological models for colonisation. The influences exerted by the archaeological context, demographic and genetic history of the samples being used and hence the kinds of analyses that are possible have also been highlighted. The thesis has also examined the implications of using larger data sets in combination with single specimens, and working across different spatial and temporal scales.

The outcomes of the aDNA analyses of dog specimens from colonisation era archaeological sites in New Zealand has demonstrated the impact that the interactions between people and dogs can have on molecular diversity, and hence the usefulness of studies using solely maternal markers. This pattern however is informative about the dogs brought to New Zealand and the colonisation process. The homogeneity of dog lineages in New Zealand archaeological sites suggests a single founding population, and supports the current archaeological model of rapid and strategic colonisation.
Complete and partial ancient mitogenomes were also generated from archaeological specimens throughout Southeast Asia and the wider Pacific to investigate the place of dogs in the Lapita expansion. These results have revealed the complexity of dog introductions and dispersals in the region, that were not previously apparent in the studies that targeted the control region only, or from the results of standard macroscopic archaeozoological methods.

9.1 Limitations of the research

The laboratory work for this thesis commenced at an exciting time following the emergence of aDNA methods, as new molecular genetic techniques had begun to be applied to archaeological specimens. Over the course of the thesis, refinements to extraction, sequencing and analytical methods were being published on a regular basis. Within the context of this research, however, it was necessary to adopt and use a particular protocol consistently throughout the research to facilitate the comparison of results. The method used for DNA recovery and sequencing of the relatively recent samples from temperate New Zealand proved to give excellent results. The method was less successful for some of the older samples from archaeological sites closer to the equator where environmental conditions would have been warmer and more humid. It is possible that some of the new methods such as single stranded library preparation could give a better result, and these should be considered for this work in future.

It was hoped that sampling a large cohort from colonisation era sites in New Zealand would enable some population genetic analyses to be undertaken. Given the homogeneity of the mitogenomes across the sample, this proved impossible to do. The influence of demographic and genetic history on the composition of the sample set should not be overlooked when planning this kind of study. The incorporation of nuclear markers is feasible, given the excellent preservation of many of the ancient specimens, and this approach could enable inter-archaeological site analysis.

One of the biggest challenges to overcome for this thesis related to the surfeit of molecular data produced, rather than the lack of it, and the appropriate methods to analyse it. The raw
9.2 Implications for future research

Dogs have a special place in human society, their appearance as the first domesticated animal species marks a turning point in our relationships with other animals. There is still much to learn about the ongoing process of dog domestication and its consequences. These are entangled with human history, and the study of dogs can be informative about many aspects, including social life, economics and migration. There are many avenues of future research that have been highlighted over the course of the research, which is to be expected considering such a wide geographical area and spanning much of the human history of the region. The now familiar call of aDNA researchers for more samples for analysis, particularly from key archaeological sites associated with the Lapita expansion, is to be expected. Notwithstanding this request, below I outline three particular areas of future research arising from the findings of this thesis.

9.2.1 New Zealand dogs

Dogs were the only domesticated animal introduced to New Zealand by the first human colonists from East Polynesia. Although the limited mtDNA diversity observed in New Zealand dogs as a result of this research precluded population-level studies, the excellent preservation of aDNA in some specimens suggests that obtaining nucDNA may be possible in the future. Studies incorporating nuclear data may enable a range of questions to be addressed, including population structure among closely related individuals, possible connections between dogs from different archaeological sites, and instances of selection for particular traits. For example, were people in New Zealand selectively breeding dogs for particular coat colours (Ollivier et al.) and hair length for use in the production of cloaks and
decorative tassels? Ethnographic accounts from East Polynesia record the desirability of long hair, and white and black hair, for use in dog skin cloaks, and include descriptions of long-haired dogs in New Zealand and the Tuamotu Islands (Luomala, 1960b) which could be investigated using aDNA methods. Osteological studies of the New Zealand dog have also identified possible variations in cranial and long bone proportions (Clark, 1995) that could be examined. Ancient DNA analyses may also be used to investigate possible phenotypic differences between deliberate dog burials and dogs used for food.

It is thought that human mobility reduced greatly following the initial colonisation process, and this may be reflected in the genetic makeup of later dog populations. Sampling and sequencing aDNA of dogs from later period archaeological sites would enable an exploration of changes in population structure over time. In addition, the decline in dog numbers in later archaeological sites (Davidson, 1987) could be related to changes in the carrying capacity of the domestic environment. Bioarchaeological indicators of animal health could also informative in this regard.

It appears that there were no feral dog populations in New Zealand prior to European settlement, contra the trajectory of the dingo, the ancestors of which were likely to have been dogs from Southeast Asia. This suggests that the environment in New Zealand may not have been conducive to the survival of dogs independently of people. The demise of the Polynesian dog, through its replacement with European dogs introduced during voyages of exploration, and subsequent arrival by sealers, whalers and European settlers is also not well understood.

9.2.2 Interactions between people and dogs

From the beginnings of domestication, whether in the Middle East, Europe, or East Asia, and then with their spread through Mainland and ISEA and the Pacific, it is clear that dogs have had varying interactions with people. Dogs continue to perform a wide range of functions in Asian and Pacific societies, as a food item, guard dogs, hunting dogs, pets and companions, totems and spirit guides. The use of dogs as a food item is strongly associated with East and
Southeast Asian and some Pacific communities; this appears to be of great antiquity and may be pertinent to understanding the dispersal of dogs and people in the Neolithic.

Exploration of the domestic relationship between dogs and people in Pacific prehistory is presently lacking. There is little or no understanding of how dogs were cared for, how they may have been part of human communities and social life, and how this may have changed through time. Archaeological explanations that explicitly address the presence of dog remains in archaeological sites in the Pacific often incorporate economic assumptions, where dogs are seen as part of the available resources used by people for food, and, in some cases, for the production of items of material culture. While possible, and in some cases highly likely, this is only one aspect of the relationship and contrasts with the complex prehistoric human-dog relationships described elsewhere in the world. Although there have been a number of studies about the social significance of the pig, particularly in Near Oceania, the corresponding role of the dog has received less scrutiny. Dogs, as well as pigs, are part of social interactions involving issues of wealth and prestige throughout the region. This appears most pronounced in Remote Oceania, where by the time of European contact, living dogs, and dog products transformed into food and items of adornment and cloaks, and could be associated with people of chiefly rank and status.

The relationship between human migration, colonisation, animal movements and the incorporation of animals into social, economic and political networks in the Pacific is an area that deserves further study.

9.2.3 Dog domestication and dispersal

The process of domestication and its consequences are entangled with human history. The study of Southeast Asian and Pacific dogs can be informative about many aspects of this shared past, including social life, economics and migration. The area south of the Yangtze River is a pivotal region for understanding dog domestication and subsequent dispersals. Archaeological information has the potential to clarify the timing of the introduction of dogs into the region, the nature of the hunting and gathering communities that had dogs, and also the processes by which dogs became associated with food producing groups. Archaeological
investigations, in addition to providing valuable information about the context of dog remains in sites, can also provide material for molecular studies.

Recent developments in methods of aDNA extraction and sequencing that enable the use of expanded genetic markers, including nucDNA, may assist significantly with evaluating the appropriateness of various genetic models and with understanding the process of domestication and possibly human selection for particular traits. In Southeast Asia, whole genome studies on both ancient and modern dogs may be used to test various hypotheses regarding the process of domestication; for instance, when do changes in the amylase gene associated with starch digestion (Axelsson et al., 2013) appear in Southeast Asian dogs?

The islands of Wallacea, Australia and New Guinea have the oldest evidence for dogs in the Pacific and also based on the limited samples presented here, the most mitochondrial genetic diversity. Archaeozoological and molecular genetic analyses of archaeological samples from ISEA and New Guinea and dingoes from Australia, possibly combined with modern samples from Papua New Guinea village dogs, would greatly assist with better understanding of the initial movements of dogs into the Pacific. The possible links between Near Oceanic dog mitogenome and Y-chromosome sequences, and those from Taiwan need further investigation. The context of the archaeological specimens and the information that they may provide in relation to animal-human interactions, including trade and exchange, and the place of dogs in these changing human communities are also important aspects to be investigated. Archaeological research programmes generating more fine-grained models for the emergence of the Neolithic in ISEA and the Pacific may also offer alternative ways of thinking about the movement of people and dogs in this region.

The relationships between people and dogs across the Asia Pacific region are not uniform, and have been influenced by a range of social, technological and environmental factors through time. Studies of dog populations and their interactions with people at particular places and times still have the potential to provide new information about the subtleties of this enduring relationship. The diversity of people’s interactions with dogs in the Asia Pacific
region serve as a reminder of the complexity of our close relationship with dogs since its commencement thousands of years ago.
References

Acosta-Jamett, G., Cleaveland, S., Cunningham, A. A. & Bronsvoort, B. M. 2010. Demography of domestic dogs in rural and urban areas of the Coquimbo region of Chile and implications for disease transmission. Preventive Veterinary Medicine, 94, 272-281.


Dwyer, P. D. & Minnegal, M. 2016. Wild dogs and village dogs in New Guinea: were they different? *Australian Mammalogy,* 38, 1-11.


Ingold, T. 1994. What is an animal?


Appendices
Appendix A
Supplementary Information—Chapter 5

List of Tables

Table A.1: Wairau Bar dog sample showing specimens used for analyses, GenBank Accession Numbers, coverage and read depth statistical summary, and percentage of imputed bases. .........................................................200

Table A.2: Levels of contamination observed in Wairau Bar dog sample. .............200

Table A.3: Informative sites for Arc1 and Arc2 short control region fragment. ....201

List of Figures

Figure A.1: Graph of mitogenome coverage of the Wairau Bar dog sequences. ....202

Figure A.2: Read depth for Wairau Bar dog sequences. ................................203

Figure A.3: Size distribution of merged endogenous mtDNA fragments from Wairau Bar dogs. ..........................................................204

Figure A.4: Wairau Bar dog sample damage patterns......................................205
Table A.1: Wairau Bar dog sample showing specimens used for analyses, GenBank Accession Numbers, coverage and read depth statistical summary, and percentage of imputed bases.

<table>
<thead>
<tr>
<th>Wairau Bar Dog ID No.</th>
<th>GenBank Accession No.</th>
<th>Coverage (%)</th>
<th>Mean read depth</th>
<th>Standard deviation of read depth</th>
<th>Imputed variable sites (%)</th>
<th>Used in final analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS10062</td>
<td>KT168373</td>
<td>99.928</td>
<td>94.11221</td>
<td>30.62578</td>
<td>0.01887</td>
<td>y</td>
</tr>
<tr>
<td>MS10064</td>
<td>n/a</td>
<td>45.328</td>
<td>0.76284</td>
<td>1.09689</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10065</td>
<td>KT168371</td>
<td>100</td>
<td>189.50487</td>
<td>28.39455</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10066</td>
<td>KT168372</td>
<td>99.647</td>
<td>66.28379</td>
<td>23.89453</td>
<td>0.01887</td>
<td>y</td>
</tr>
<tr>
<td>MS10067</td>
<td>n/a</td>
<td>78.311</td>
<td>1.86035</td>
<td>1.62678</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10068</td>
<td>KT168369</td>
<td>99.857</td>
<td>27.68781</td>
<td>9.32758</td>
<td>0.03774</td>
<td>y</td>
</tr>
<tr>
<td>MS10069</td>
<td>KT168370</td>
<td>100</td>
<td>424.91863</td>
<td>80.98381</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10070</td>
<td>KT168374</td>
<td>99.988</td>
<td>239.79584</td>
<td>64.56409</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10129</td>
<td>KT168375</td>
<td>98.655</td>
<td>11.61177</td>
<td>5.95921</td>
<td>0.13208</td>
<td>y</td>
</tr>
<tr>
<td>MS10130</td>
<td>KT168378</td>
<td>99.002</td>
<td>14.86602</td>
<td>6.95936</td>
<td>0.09434</td>
<td>y</td>
</tr>
<tr>
<td>MS10131</td>
<td>KT168379</td>
<td>99.707</td>
<td>46.41711</td>
<td>11.28736</td>
<td>0.01887</td>
<td>y</td>
</tr>
<tr>
<td>MS10132</td>
<td>KT168376</td>
<td>99.713</td>
<td>21.39953</td>
<td>6.92298</td>
<td>0.03774</td>
<td>y</td>
</tr>
<tr>
<td>MS10133</td>
<td>KT168377</td>
<td>99.976</td>
<td>46.25426</td>
<td>12.09964</td>
<td>0.01887</td>
<td>y</td>
</tr>
<tr>
<td>MS10135</td>
<td>KT168382</td>
<td>99.88</td>
<td>37.31058</td>
<td>9.70366</td>
<td>0.03774</td>
<td>y</td>
</tr>
<tr>
<td>MS10136</td>
<td>KT168380</td>
<td>100</td>
<td>206.85479</td>
<td>42.26062</td>
<td>0.01887</td>
<td>y</td>
</tr>
<tr>
<td>MS10137</td>
<td>KT168381</td>
<td>99.229</td>
<td>10.89687</td>
<td>4.46003</td>
<td>0.07547</td>
<td>y</td>
</tr>
</tbody>
</table>

Table A.2: Levels of contamination observed in Wairau Bar dog sample.

<table>
<thead>
<tr>
<th>Wairau Bar Dog ID No.</th>
<th>Human (NC_012920.1)</th>
<th>Chicken (NC_001323.1)</th>
<th>Cow (NC_006853.1)</th>
<th>Pig (NC_012095.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS10062</td>
<td>0.0</td>
<td>0.0</td>
<td>0.000356071010733</td>
<td>5.08672872476e-05</td>
</tr>
<tr>
<td>MS10065</td>
<td>0.00410667207463</td>
<td>0.0</td>
<td>0.00258568241736</td>
<td>2.53498276212e-05</td>
</tr>
<tr>
<td>MS10066</td>
<td>7.57735008921e-05</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10068</td>
<td>0.000342583076396</td>
<td>0.0</td>
<td>0.000171291538198</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10069</td>
<td>4.76043129508e-05</td>
<td>0.0</td>
<td>0.00174945850094</td>
<td>2.38021564754e-05</td>
</tr>
<tr>
<td>MS10070</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00122447194647</td>
<td>2.18655704727e-05</td>
</tr>
<tr>
<td>MS10129</td>
<td>0.000414765657404</td>
<td>0.0</td>
<td>0.000414765657404</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10130</td>
<td>0.00480923372876</td>
<td>0.0</td>
<td>0.00032061558197</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10131</td>
<td>0.000300751879699</td>
<td>0.0</td>
<td>0.00220551378446</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10132</td>
<td>0.00108014689999</td>
<td>0.0</td>
<td>0.00021602979996</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10133</td>
<td>0.00129844186976</td>
<td>0.0</td>
<td>0.0061925689173</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10135</td>
<td>0.00038333759264</td>
<td>0.0</td>
<td>0.00332225913621</td>
<td>0.0</td>
</tr>
<tr>
<td>MS10136</td>
<td>0.000481894533939</td>
<td>0.0</td>
<td>0.00117031529671</td>
<td>2.2947358759e-05</td>
</tr>
<tr>
<td>MS10137</td>
<td>0.0075055187638</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table A.3: Informative sites for Arc1 and Arc2 short control region fragment.

<table>
<thead>
<tr>
<th>Dog mitochondrial reference sequence (NC_002008)</th>
<th>15630 A</th>
<th>15642 T</th>
<th>15648 A</th>
<th>15655 G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arc1</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>Arc2</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10062</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10064</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10065</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10066</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10068</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10069</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10070</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10129</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10130</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10131</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10132</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10133</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10135</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>MS10136</td>
<td>G</td>
<td>A</td>
<td>G</td>
<td>A</td>
</tr>
</tbody>
</table>
Figure A.1: Graph of mitogenome coverage of the Wairau Bar dog sequences. Each sample is represented by a horizontal bar. The length of each bar on the x-axis shows the proportion of the mitochondrial genome covered by at least one read. The read depth is shown by the colour gradient given in the legend.
Figure A2: Read depth for Wairau Bar dog sequences. The x-axis represents the length of the dog mitochondrial genome in base pairs. The y-axis is the number of reads covering each site of the mitogenome.
Figure A.3: Size distribution of merged endogenous mtDNA fragments from Wairau Bar dogs. Number of reads and fragment length are provided.
Figure A4: Wairau Bar dog sample damage patterns. Top plots: Base frequency 5’ and 3’ of strand breaks. The gray brackets indicate start and end of molecules (strand breaks). Purines (A and G) show an elevated frequency before strand breaks. Bottom plots: C to T and G to A nucleotide misincorporations at the first and last 25 bases of endogenous mtDNA fragments from the Wairau Bar dog sample (Sawyer et al., 2012).
Appendix B
Supplementary Information—Chapter 6

List of tables

Table B.1: Table with specimens used for analyses, GenBank Accession Numbers, coverage and read depth statistical summary, and percentage of imputed bases. ..........207

Table B.2: Levels of contamination observed in the New Zealand samples.........................209

List of figures

Figure B.1: Graph of mitogenome coverage for New Zealand samples.........................211

Figure B.2: Read depth for New Zealand samples..........................................................212

Figure B.3: Size distribution of merged endogenous mtDNA fragments for the New Zealand Samples.................................................................................................213
Table B.1: Table with specimens used for analyses, GenBank Accession Numbers, coverage and read depth statistical summary, and percentage of imputed bases.

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Archaeological site</th>
<th>Archaeological provenience</th>
<th>GenBank Accession No.</th>
<th>Coverage (%)</th>
<th>Mean read depth</th>
<th>Standard deviation of read depth</th>
<th>Imputed variable sites (%)</th>
<th>Used in final analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS10071</td>
<td>Houhora</td>
<td>D9 3b</td>
<td>KU215692</td>
<td>100</td>
<td>286.0175764</td>
<td>61.39303444</td>
<td>0.015625</td>
<td>y</td>
</tr>
<tr>
<td>MS10072</td>
<td>Houhora</td>
<td>D7 3</td>
<td>KU215691</td>
<td>100</td>
<td>261.239493</td>
<td>57.81743223</td>
<td>0.015625</td>
<td>y</td>
</tr>
<tr>
<td>MS10073</td>
<td>Houhora</td>
<td>B6 3</td>
<td>KU215690</td>
<td>100</td>
<td>250.4369582</td>
<td>48.4396315</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10074</td>
<td>Houhora</td>
<td>A6 3</td>
<td>KU215689</td>
<td>100</td>
<td>173.1709213</td>
<td>37.28464276</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10075</td>
<td>Houhora</td>
<td>C10 2b</td>
<td>KU215688</td>
<td>100</td>
<td>307.7526155</td>
<td>70.29256448</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10076</td>
<td>Houhora</td>
<td>C10 2b</td>
<td>KU215696</td>
<td>99.67</td>
<td>51.51318228</td>
<td>12.83643932</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10077</td>
<td>Houhora</td>
<td>C8 3a</td>
<td>KU215694</td>
<td>99.86</td>
<td>76.30561368</td>
<td>18.62196162</td>
<td>0.046875</td>
<td>y</td>
</tr>
<tr>
<td>MS10078</td>
<td>Houhora</td>
<td>C7 2d</td>
<td>KU215695</td>
<td>99.56</td>
<td>24.781491</td>
<td>8.020079539</td>
<td>0.09375</td>
<td>y</td>
</tr>
<tr>
<td>MS10079</td>
<td>Houhora</td>
<td>C7 2d</td>
<td>KU215687</td>
<td>99.74</td>
<td>29.6989866</td>
<td>14.72041935</td>
<td>0.0625</td>
<td>y</td>
</tr>
<tr>
<td>MS10080</td>
<td>Houhora</td>
<td>C8 3b</td>
<td>n/a</td>
<td>32.82</td>
<td>0.40252286</td>
<td>0.639244926</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10093</td>
<td>Pleasant River</td>
<td>PLR 150, BQ1</td>
<td>KU215701</td>
<td>99.12</td>
<td>15.53362827</td>
<td>7.866011576</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10094</td>
<td>Pleasant River</td>
<td>PLR 50, BQ6</td>
<td>KU215698</td>
<td>99.98</td>
<td>32.42559933</td>
<td>10.70820898</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10095</td>
<td>Pleasant River</td>
<td>PLR 50, BQ5</td>
<td>KU215697</td>
<td>99.32</td>
<td>12.2268787</td>
<td>5.35576362</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10096</td>
<td>Pounawea</td>
<td>PN/J18/L2</td>
<td>KU215700</td>
<td>99.99</td>
<td>42.3032125</td>
<td>9.847018347</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10097</td>
<td>Pounawea</td>
<td>PN/J17/L2</td>
<td>KU215699</td>
<td>99.99</td>
<td>42.3273749</td>
<td>9.336570409</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10098</td>
<td>Pounawea</td>
<td>PN/J21/L2</td>
<td>KU215702</td>
<td>99.96</td>
<td>72.1070724</td>
<td>14.74762628</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10110</td>
<td>Shag Mouth</td>
<td>SM/D 620 BQ-1</td>
<td>n/a</td>
<td>55.4</td>
<td>0.834683608</td>
<td>0.939539297</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10111</td>
<td>Shag Mouth</td>
<td>SM/D 650 BQ-1</td>
<td>KU215682</td>
<td>98.98</td>
<td>9.035212531</td>
<td>4.0244103</td>
<td>0.15625</td>
<td>y</td>
</tr>
<tr>
<td>MS10112</td>
<td>Shag Mouth</td>
<td>SM/D 640 BQ-2</td>
<td>KU215683</td>
<td>99.57</td>
<td>15.71614755</td>
<td>6.030554494</td>
<td>0.0625</td>
<td>y</td>
</tr>
<tr>
<td>MS10113</td>
<td>Shag Mouth</td>
<td>SM/B 37 BQ-1</td>
<td>KU215684</td>
<td>100</td>
<td>30.43438752</td>
<td>22.45777488</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10114</td>
<td>Shag Mouth</td>
<td>SM/B 12 BQ-5</td>
<td>KU215685</td>
<td>99.41</td>
<td>10.62360256</td>
<td>4.037299205</td>
<td>0.078125</td>
<td>y</td>
</tr>
<tr>
<td>Code</td>
<td>Area</td>
<td>PIT</td>
<td>Location</td>
<td>Sample Code</td>
<td>Date Code</td>
<td>Age (y)</td>
<td>Biovolume (μm³)</td>
<td>Distance (km)</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>------</td>
<td>---------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>---------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>MS10115</td>
<td>Shag Mouth</td>
<td>KU2</td>
<td>SM/B 22 BQ-2</td>
<td>99.09</td>
<td>12.96873319</td>
<td>5.767303259</td>
<td>0.078125</td>
<td>y</td>
</tr>
<tr>
<td>MS10117</td>
<td>Shag Mouth</td>
<td>n/a</td>
<td>SM/D 385 BQ-3</td>
<td>78.95</td>
<td>1.544987147</td>
<td>1.24759772</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10123</td>
<td>Shag Mouth</td>
<td>n/a</td>
<td>SM/C 19 BQ-2</td>
<td>41.18</td>
<td>0.519878041</td>
<td>0.703297632</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10124</td>
<td>Shag Mouth</td>
<td>n/a</td>
<td>SM/D 1 BQ-1</td>
<td>98.77</td>
<td>6.711603994</td>
<td>3.311381942</td>
<td>0.234375</td>
<td>y</td>
</tr>
<tr>
<td>MS10062</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-30; BQ-5</td>
<td>99.95</td>
<td>96.12100197</td>
<td>31.14564181</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10064</td>
<td>Wairau Bar</td>
<td>n/a</td>
<td>WB-CSA-65; BQ-12</td>
<td>46.16</td>
<td>0.783224726</td>
<td>1.114350772</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10065</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-56; BQ-2</td>
<td>100</td>
<td>187.1910085</td>
<td>30.23237431</td>
<td>0</td>
<td>y</td>
</tr>
<tr>
<td>MS10066</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-84; BQ-1</td>
<td>99.58</td>
<td>67.55999283</td>
<td>24.3031636</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10067</td>
<td>Wairau Bar</td>
<td>n/a</td>
<td>WB-CSA-84; BQ-2</td>
<td>78.85</td>
<td>1.896155916</td>
<td>1.64493616</td>
<td>* not used</td>
<td>n</td>
</tr>
<tr>
<td>MS10068</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-56; BQ-1</td>
<td>99.88</td>
<td>28.3347807</td>
<td>9.463862894</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10069</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-68; BQ-7</td>
<td>100</td>
<td>432.7004842</td>
<td>82.056367</td>
<td>0.015625</td>
<td>y</td>
</tr>
<tr>
<td>MS10070</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-CSA-68; BQ-7</td>
<td>99.99</td>
<td>244.3329348</td>
<td>65.60904926</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10129</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-27</td>
<td>98.68</td>
<td>11.9106833</td>
<td>5.712001547</td>
<td>0.125</td>
<td>y</td>
</tr>
<tr>
<td>MS10130</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-27</td>
<td>99</td>
<td>15.2641836</td>
<td>7.12935422</td>
<td>0.09375</td>
<td>y</td>
</tr>
<tr>
<td>MS10131</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-27 BQ</td>
<td>99.72</td>
<td>47.69139714</td>
<td>11.60070205</td>
<td>0.046875</td>
<td>y</td>
</tr>
<tr>
<td>MS10132</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-85 BQ</td>
<td>99.72</td>
<td>21.9852924</td>
<td>7.088684989</td>
<td>0.046875</td>
<td>y</td>
</tr>
<tr>
<td>MS10133</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-73</td>
<td>99.98</td>
<td>47.66802176</td>
<td>12.47607665</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10135</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-C-142 BQ</td>
<td>99.91</td>
<td>38.34602738</td>
<td>9.876219431</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10136</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-D-31 BQ</td>
<td>100</td>
<td>212.509655</td>
<td>43.18645993</td>
<td>0.03125</td>
<td>y</td>
</tr>
<tr>
<td>MS10137</td>
<td>Wairau Bar</td>
<td>KT1</td>
<td>WB-D-6 BQ</td>
<td>99.25</td>
<td>11.17636157</td>
<td>4.538749511</td>
<td>0.078125</td>
<td>y</td>
</tr>
<tr>
<td>MS10326</td>
<td>Ureia</td>
<td></td>
<td>TP8, Zone H, level 8, Sample bag #248</td>
<td>in progress</td>
<td>95.67</td>
<td>4.029652657</td>
<td>2.631127563</td>
<td>Y</td>
</tr>
<tr>
<td>MS10327</td>
<td>Ureia</td>
<td></td>
<td>TP5, Zone E, level 5, Sample bag #149</td>
<td>in progress</td>
<td>96.07</td>
<td>5.22705805</td>
<td>3.461024973</td>
<td>Y</td>
</tr>
</tbody>
</table>
Table B.2: Levels of contamination observed in the New Zealand samples.

| Sample ID | Human  (gi|251831106|ref|NC_012920.1|) | Chicken  (gi|5834843|ref|NC_001323.1|) | Cow (gi|60101824|ref|NC_006853.1|) | Pig  (gi|223976078|ref|NC_012095.1|) |
|-----------|-----------------|-------------------|-----------------|-----------------|-----------------|
| MS10062   | 0               | 0                 | 0.000406938     | 5.09E-05        |
| MS10065   | 0.020531799     | 0                 | 0.005272363     | 2.53E-05        |
| MS10066   | 7.76E-05        | 0                 | 0.000232702     | 0               |
| MS10068   | 0.001028101     | 0                 | 0.00017135      | 0               |
| MS10069   | 4.76E-05        | 0                 | 0.002034988     | 2.38E-05        |
| MS10070   | 0               | 0                 | 0.001311992     | 2.19E-05        |
| MS10071   | 0.000590688     | 0                 | 0.004951355     | 0               |
| MS10072   | 0.000715889     | 0                 | 0.001818745     | 0               |
| MS10073   | 0.000860417     | 0                 | 0.002522586     | 0               |
| MS10074   | 0.000961104     | 0                 | 0.007773632     | 0               |
| MS10075   | 0.000460981     | 0                 | 0.001514653     | 0               |
| MS10076   | 0.001018226     | 0                 | 0.004887486     | 0.000203645     |
| MS10077   | 0.006907246     | 0.000211446       | 0.007259656     | 0.000140964     |
| MS10078   | 0.002877698     | 0                 | 0.002877698     | 0               |
| MS10079   | 0.004848628     | 0                 | 0.00307474      | 0.000236518     |
| MS10111   | 0.039303762     | 0                 | 0.000561482     | 0               |
| MS10112   | 0.003910068     | 0                 | 0.000977517     | 0               |
| MS10113   | 0.010119357     | 0                 | 0.0495589       | 0.000129735     |
| MS10114   | 0.034778681     | 0                 | 0.000903342     | 0               |
| MS10115   | 0.000793021     | 0                 | 0.000396511     | 0               |
| MS10124   | 0               | 0                 | 0               | 0               |
| MS10131   | 0.001102646     | 0                 | 0.002906977     | 0               |
| MS10132 | 0.0077777058 | 0 | 0.000216029 | 0 |
| MS10133 | 0.0071899934 | 0 | 0.008188536 | 0 |
| MS10135 | 0.00319448 | 0 | 0.003705597 | 0 |
| MS10136 | 0.001078526 | 0 | 0.001743999 | 2.29E-05 |
| MS10137 | 0.080353201 | 0 | 0 | 0 |
| MS10095 | 0 | 0 | 0.005197055 | 0 |
| MS10094 | 0.058010606 | 0 | 0.003374578 | 0 |
| MS10097 | 0.005611125 | 0 | 0.00634301 | 0.000121981 |
| MS10096 | 0.001702128 | 0 | 0.001823708 | 0.000243161 |
| MS10093 | 0.000343289 | 0 | 0.008582218 | 0.000686577 |
| MS10098 | 0.000289394 | 0 | 0.001446969 | 0 |
| MS10130 | 0.012508018 | 0 | 0.000320718 | 0 |
| MS10129 | 0.006639004 | 0 | 0.000414938 | 0 |
Figure B.1: Graph of mitogenome coverage for New Zealand samples. Each specimen is represented by a horizontal bar. The length of each bar on the x-axis shows the proportion of the mitochondrial genome covered by at least two reads. The read depth is shown by the colour gradient given in the legend.
Figure B.2: Read depth for New Zealand samples. The x-axis represents the length of the dog mitochondrial genome in base pairs. The y-axis is the number of reads covering each site of the mitogenome.
Figure B.3: Size distribution of merged endogenous mtDNA fragments for the New Zealand samples. Number of reads and fragment length are provided.
Appendix C
Supplementary Information—Chapter 7

Modern dingo extraction and sequencing.................................214

List of Tables

Table C.1: Ancient specimens used for analyses, coverage and read depth statistical summary...................................................................................................................217
Table C.2: Variable sites relating to haplotype assignment.........................218
Table C.3: Defining SNPs observed in Late Lapita ancient dog sequences.........225

List of Figures

Figure C.1: Bayesian tree created with BEAST......................................227
Figure C.2: Graph of mitogenome coverage of the ancient dog sequences........228
Figure C.3 Read depth for the ancient dog sequences..................................229
Figure C.4: Ancient dog damage patterns..............................................230
Modern dingo DNA extraction and sequencing (carried out by Kate Mcdonald, Department of Anatomy, University of Otago, Dunedin)

Four dingo blood samples (1.5mL each) were collected by Dr. Lisa Argilla at the Wellington Zoo (DW, DY, DB, DK) in August 2015. These were stored in EDTA for delivery to the University of Otago immediately following collection. Blood samples were kept in the freezer (-25°C) following arrival.

All DNA extraction and sequencing library preparation was carried out in a University of Otago modern DNA laboratory. 1000µL of blood was processed for each sample. As the blood was stored in EDTA, cells were extracted from the solution using 1x SSC buffer to ensure the EDTA would not inhibit the PCR. DNA was extracted from the cells using the MagJET magnetic bead (ThermoFisher) protocol (as per manufacturer’s instructions). Samples were digested with Proteinase K for 4 hours prior to extraction. DNA was precipitated with isopropanol, washed with ethanol-based wash buffers and eluted in Elution Buffer (provided in Qiagen MinElute PCR Purification Kit).

Four long-range primer pairs (designed by Dr. Ann Horsburgh) were used to amplify the complete mitochondrial genomes of the Wellington Zoo dingoes. Table 1 gives the primer pairs used for amplification of each sample along with their sequences and binding positions.

Table 1. Primers used for long-range amplification of blood and tissue samples, and their relative position within the mitochondrial genome.

<table>
<thead>
<tr>
<th>Samples: Wellington Zoo dingoes</th>
<th>Primer Pairs:</th>
<th>Primer Sequence:</th>
<th>5’-3’ Binding Position:</th>
<th>Expected Length (bp):</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR1F LR1.5R</td>
<td>TGGTATCACCCACCTAGGACC</td>
<td>15,947-15,967</td>
<td>3457</td>
<td></td>
</tr>
<tr>
<td>LR1.5F LR1R</td>
<td>CCTAACAAGGCCCCCTTTCCTC</td>
<td>2656-2676</td>
<td>3036</td>
<td></td>
</tr>
<tr>
<td>LR2F LR2R</td>
<td>CAGCCCTCCCCAGTACGAAAG</td>
<td>2544-2564</td>
<td>5581-5601</td>
<td></td>
</tr>
<tr>
<td>LR3F LR3R</td>
<td>CGGACTAGTCGTTCCTCCAAA</td>
<td>10,514-10,534</td>
<td>5162</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TGAAGCAACACTGATTCCGAC</td>
<td>10,569-10,589</td>
<td>5439</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GACGGCCATATCCCTGGAGG</td>
<td>15,998-16,018</td>
<td>5729</td>
<td></td>
</tr>
</tbody>
</table>

Long-range PCR products were amplified using KAPA Biosystems LongRange HotStart PCR Protocol (as per manufacturer’s instructions). 2.0µL of template DNA was used in a 30.0µL reaction, with 1.5µL of each of the corresponding primers. All PCR reactions were run in a Bio-Rad iCycler. PCR products were viewed by gel electrophoresis using 2.5µL of PCR product on a 1% agarose gel to ensure DNA was present in the amplified fragments.

PCR products that showed expected lengths of DNA in the gel were purified using Qiagen MinElute silica spin columns, as the number of samples was small. This was done following the manufacturer’s instructions.
Quantification of purified PCR products was done using Qubit dsDNA HS (High Sensitivity) Assay Kit and Qubit 3.0 fluorometer (Life Technologies). For each sample, 199µL of Qubit dsDNA HS Buffer (0.2-100ng) buffer and 1µL of HS dsDNA dye. As the dye is light sensitive solution was made immediately prior to use. Standards are supplied in the kit.

Blunt end repair, ligation of sequencing adaptor and barcoding were carried out following Kircher at al. 2012, with modifications for Illumina sequencing adaptors. Samples were sonicated using diagenode Bioruptor Pico Sonication System to produce fragments c. 500bp in length. Purification of tissue sample libraries was performed using AMPure XP magnetic beads.
Table C.1: Ancient specimens used for analyses, with coverage and read depth statistical summaries.

<table>
<thead>
<tr>
<th>Laboratory Specimen No.</th>
<th>Location</th>
<th>Coverage (%)</th>
<th>Mean read depth</th>
<th>Standard deviation of read depth</th>
<th>Included in phylogenetic analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS10015</td>
<td>Atafu, Tokelau</td>
<td>54.25</td>
<td>0.802893525</td>
<td>0.899637217</td>
<td>N</td>
</tr>
<tr>
<td>MS10016</td>
<td>Atafu, Tokelau</td>
<td>98.33</td>
<td>6.987505231</td>
<td>3.360716471</td>
<td>Y</td>
</tr>
<tr>
<td>MS10017</td>
<td>Atafu, Tokelau</td>
<td>71.36</td>
<td>1.472649011</td>
<td>1.320041998</td>
<td>N</td>
</tr>
<tr>
<td>MS10023</td>
<td>Bogi1, Papua New Guinea</td>
<td>43.06</td>
<td>0.590960722</td>
<td>0.826636889</td>
<td>N</td>
</tr>
<tr>
<td>MS10031</td>
<td>Edubu1, Papua New Guinea</td>
<td>17.2</td>
<td>0.191606385</td>
<td>0.44289477</td>
<td>N</td>
</tr>
<tr>
<td>MS10043</td>
<td>Goemu, Torres Strait</td>
<td>98.45</td>
<td>10.34435344</td>
<td>7.609122939</td>
<td>Y</td>
</tr>
<tr>
<td>MS10044</td>
<td>Goemu, Torres Strait</td>
<td>44.04</td>
<td>0.644586597</td>
<td>0.911223393</td>
<td>N</td>
</tr>
<tr>
<td>MS10062</td>
<td>Wairau Bar, NZ</td>
<td>99.95</td>
<td>96.12632271</td>
<td>31.14222377</td>
<td>Y</td>
</tr>
<tr>
<td>MS10090</td>
<td>Gracios Bay, Santa Cruz</td>
<td>83.71</td>
<td>2.127697734</td>
<td>1.645368195</td>
<td>Y</td>
</tr>
<tr>
<td>MS10324</td>
<td>Kauai, Hawaii</td>
<td>27.14</td>
<td>0.320918276</td>
<td>0.57693473</td>
<td>N</td>
</tr>
<tr>
<td>MS10325</td>
<td>Kauai, Hawaii</td>
<td>68.78</td>
<td>1.399772822</td>
<td>1.432187493</td>
<td>N</td>
</tr>
<tr>
<td>MS10665</td>
<td>Kauai, Hawaii</td>
<td>100</td>
<td>152.955999282597</td>
<td>56.1334554625085</td>
<td>Y</td>
</tr>
<tr>
<td>MS10666</td>
<td>Kauai, Hawaii</td>
<td>33.22</td>
<td>0.390566150535063</td>
<td>0.623172029128382</td>
<td>N</td>
</tr>
<tr>
<td>MS10326</td>
<td>Aitutaki, Cook Islands</td>
<td>95.67</td>
<td>4.029652657</td>
<td>2.631127563</td>
<td>Y</td>
</tr>
<tr>
<td>MS10327</td>
<td>Aitutaki, Cook Islands</td>
<td>96.07</td>
<td>5.22705805</td>
<td>3.461024973</td>
<td>Y</td>
</tr>
<tr>
<td>MS10329</td>
<td>Taurama, New Guinea</td>
<td>90.5</td>
<td>3.725473785</td>
<td>3.237475759</td>
<td>Y</td>
</tr>
<tr>
<td>MS10330</td>
<td>Matja Kuru 2, East Timor</td>
<td>98.19</td>
<td>6.897351587</td>
<td>3.166819041</td>
<td>Y</td>
</tr>
<tr>
<td>MS10331</td>
<td>Ban Chiang, Thailand</td>
<td>66.2</td>
<td>1.093860226</td>
<td>1.076768681</td>
<td>N</td>
</tr>
<tr>
<td>MS10332</td>
<td>Nong Nor, Thailand</td>
<td>79.36</td>
<td>2.247205117</td>
<td>1.97359114</td>
<td>N</td>
</tr>
<tr>
<td>MS10333</td>
<td>Shisanhang, Taiwan</td>
<td>96.71</td>
<td>5.698810307</td>
<td>2.920348464</td>
<td>Y</td>
</tr>
<tr>
<td>MS10334</td>
<td>Shisanhang, Taiwan</td>
<td>99.97</td>
<td>36.19824236</td>
<td>17.7893809</td>
<td>Y</td>
</tr>
</tbody>
</table>
Table C.2: Variable sites relating to haplotype assignment (after Duleba et al. 2015).

<table>
<thead>
<tr>
<th>Dog mitochondrial reference sequence (NC_002008)</th>
<th>Wairau Bar</th>
<th>Kauai</th>
<th>Aitutaki</th>
<th>Tokelau</th>
<th>Tokelau</th>
<th>Santa Cruz</th>
<th>Torres Strait</th>
<th>Torres Strait</th>
<th>Taurama</th>
<th>South Coast</th>
<th>East Timor</th>
<th>Thailand</th>
<th>Thailand</th>
<th>SSH</th>
<th>SSH</th>
</tr>
</thead>
<tbody>
<tr>
<td>32C</td>
<td>32T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>381A</td>
<td></td>
</tr>
<tr>
<td>54C</td>
<td>54T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82C</td>
<td>82T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>381T</td>
<td>381A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>381A</td>
<td></td>
</tr>
<tr>
<td>642A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>642T</td>
<td></td>
</tr>
<tr>
<td>645G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>645A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>648A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>648T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>649A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>649T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1081G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1081A</td>
<td>1081A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1134A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1134G</td>
<td>1134G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1196G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1196A</td>
<td>1196A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1454G</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td>1454A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1725G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1725A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1756C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1756T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2185T</td>
<td>2185C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2185C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2190A</td>
<td>2190G</td>
<td>2190G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2232A</td>
<td>2232G</td>
<td>2232G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2233G</td>
<td></td>
<td>2233A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2388A</td>
<td></td>
<td>2388T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2636G</td>
<td></td>
<td>2636A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
<td>2678.1G</td>
</tr>
<tr>
<td>2638G</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
<td>2683A</td>
</tr>
<tr>
<td>2800C</td>
<td></td>
<td>2800T</td>
<td>2800T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2812C</td>
<td></td>
<td>2812T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2962C</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
<td>2962T</td>
</tr>
<tr>
<td>3028A</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
<td>3028C</td>
</tr>
<tr>
<td>3196T</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
<td>3196C</td>
</tr>
<tr>
<td>3238G</td>
<td></td>
<td>3238A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3257C</td>
<td></td>
<td>3257T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3406C</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
<td>3406T</td>
</tr>
<tr>
<td>3469G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3598G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3692C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3875C</td>
<td></td>
<td>3875T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4204G</td>
<td>4204A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4277A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4390T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4466G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4517G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4646T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5009T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13660C</td>
<td>13660T</td>
<td>13660T</td>
<td></td>
<td>13660T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13708C</td>
<td>13708T</td>
<td></td>
<td></td>
<td>13708T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13777T</td>
<td></td>
<td>13777A</td>
<td></td>
<td>13777A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13864C</td>
<td>13864T</td>
<td>13864T</td>
<td>13864T</td>
<td>13864T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14329A</td>
<td>14329G</td>
<td>14329G</td>
<td>14329G</td>
<td>14329G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14329C</td>
<td>14383G</td>
<td>14383G</td>
<td>14383G</td>
<td>14383G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14467C</td>
<td></td>
<td>14467T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14580C</td>
<td></td>
<td>14500T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14650C</td>
<td></td>
<td>14650T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14671C</td>
<td></td>
<td>14671A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14711C</td>
<td></td>
<td>14711T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14927C</td>
<td></td>
<td>14927T</td>
<td></td>
<td>14927T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14930T</td>
<td></td>
<td>14930C</td>
<td></td>
<td>14930C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14998A</td>
<td>14998G</td>
<td>14998G</td>
<td>14998G</td>
<td>14998G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14998C</td>
<td>14998G</td>
<td>14998G</td>
<td>14998G</td>
<td>14998G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15172C</td>
<td>15172T</td>
<td>15172T</td>
<td>15172T</td>
<td>15172T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15172A</td>
<td>15172D</td>
<td>15172T</td>
<td>15172T</td>
<td>15172T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15214G</td>
<td>15214A</td>
<td>15214A</td>
<td>15214A</td>
<td>15214A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15214C</td>
<td>15214T</td>
<td>15214A</td>
<td>15214A</td>
<td>15214A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15316A</td>
<td>15316T</td>
<td></td>
<td></td>
<td>15316T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15435G</td>
<td></td>
<td></td>
<td></td>
<td>15435G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15461C</td>
<td></td>
<td></td>
<td>15461D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15464C</td>
<td></td>
<td></td>
<td>15464T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15562C</td>
<td></td>
<td></td>
<td></td>
<td>15562C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15595C</td>
<td></td>
<td></td>
<td></td>
<td>15595C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15612T</td>
<td></td>
<td></td>
<td></td>
<td>15612T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15627A</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15627C</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>156272</td>
<td>15627D</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>156273</td>
<td>15627E</td>
<td>15627G</td>
<td>15627G</td>
<td>15627G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15632C</td>
<td></td>
<td></td>
<td></td>
<td>15632C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15639T</td>
<td>15639A</td>
<td>15639A</td>
<td>15639A</td>
<td>15639G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15643A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15643G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15645A</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td>15645G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15652G</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td>15652A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15800T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15800C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15814C</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td>15814T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15815T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15815C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15912C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15912T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15955C</td>
<td>15955T</td>
<td>15955T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15955T</td>
<td>15955T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16003A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16003G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16025T</td>
<td>16025C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16025C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16033A</td>
<td>16033G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16033G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C.3: Defining SNPs for Haplogroups A & B (after Duleba et al. 2015) observed in ancient sequences from dog remains excavated from Late Lapita archaeological sites on the South Coast, Papua New Guinea.

<table>
<thead>
<tr>
<th>Position in dog reference genome NC_02008 (Kim et al. 1998)</th>
<th>MS10023</th>
<th>MS10031</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read depth</td>
<td>Nucleotide</td>
</tr>
<tr>
<td><strong>Defining SNPs for BCE haplogroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>381</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td>8221</td>
<td>C</td>
<td>X1</td>
</tr>
<tr>
<td>8760</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td>8877</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td>13618</td>
<td>G</td>
<td>X3</td>
</tr>
<tr>
<td>15955</td>
<td>T</td>
<td>X1</td>
</tr>
<tr>
<td>16003</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td><strong>Defining SNPs for BE haplogroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2812</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>4277</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>4390</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>8569</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>8736</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>8817</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>10440</td>
<td>C</td>
<td>X1</td>
</tr>
<tr>
<td>12665</td>
<td>C</td>
<td>X1</td>
</tr>
<tr>
<td>14671</td>
<td>A</td>
<td>X1</td>
</tr>
<tr>
<td>15632</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>15652</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td><strong>Defining SNPs for B haplogroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2185</td>
<td>C</td>
<td>X1</td>
</tr>
<tr>
<td>4204</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td>6764</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>11948</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td>13594</td>
<td>A</td>
<td>X2</td>
</tr>
<tr>
<td>15595</td>
<td>T</td>
<td>-</td>
</tr>
<tr>
<td>15612</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>15639</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td>15643</td>
<td>G</td>
<td>-</td>
</tr>
<tr>
<td>15815</td>
<td>C</td>
<td>X3</td>
</tr>
<tr>
<td><strong>Defining SNPs for B1 haplogroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>445</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>4646</td>
<td>C</td>
<td>X1</td>
</tr>
<tr>
<td>5219</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>9825</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td>10257</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td>10542</td>
<td>G</td>
<td>X1</td>
</tr>
<tr>
<td>14608</td>
<td>A</td>
<td>-</td>
</tr>
<tr>
<td>14930</td>
<td>C</td>
<td>X2</td>
</tr>
<tr>
<td>Position in dog reference genome NC_02008 (Kim et al. 1998)</td>
<td>MS10023</td>
<td>MS10031</td>
</tr>
<tr>
<td>------------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>Nucleotide (base called if different to 'A' SNP)</td>
<td>Read depth</td>
</tr>
<tr>
<td>1454 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3469 G</td>
<td>x1</td>
<td>0/1 (A)</td>
</tr>
<tr>
<td>3598 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4517 G</td>
<td>X1</td>
<td>0/1 (A)</td>
</tr>
<tr>
<td>5009 T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6257 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6470 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8225 T</td>
<td>X1</td>
<td>0/1 (C)</td>
</tr>
<tr>
<td>8425 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8991 A</td>
<td>X3</td>
<td>0/3 (2x G)</td>
</tr>
<tr>
<td>10404 C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10533 A</td>
<td>X2</td>
<td>0/2 (T)</td>
</tr>
<tr>
<td>11250 T</td>
<td>X1</td>
<td>0/1 (T)</td>
</tr>
<tr>
<td>11402 T</td>
<td>X1</td>
<td>0/1 (C)</td>
</tr>
<tr>
<td>11936 C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12813 G</td>
<td>X1</td>
<td>0/1 (A)</td>
</tr>
<tr>
<td>13261 C</td>
<td>X2</td>
<td>0/2 (T)</td>
</tr>
<tr>
<td>13660 C</td>
<td>X1</td>
<td>1/1 C</td>
</tr>
<tr>
<td>13708 C</td>
<td>X1</td>
<td>0/1 T</td>
</tr>
<tr>
<td>13777 T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15435 G</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15639 T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15800 T</td>
<td>X4</td>
<td>2/4 (C)</td>
</tr>
<tr>
<td>15912 C</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure C.1: Bayesian tree created with BEAST (v1.8.2) (Drummond et al., 2012) using 10,000,000 generations and sampling every 1000th generation. The maximum credibility tree was determined using TreeAnnotator with 25% of the trees being discarded as burn-in. Convergence diagnostics were investigated using Tracer (v1.4.0) and FigTree (v1.4.0).
Figure C2: Graph of mitogenome coverage of the ancient dog sequences. Each sample is represented by a horizontal bar. The length of each bar on the x-axis shows the proportion of the mitochondrial genome covered by at least one read. The read depth is shown by the colour gradient given in the legend.
Figure C.3 Read depth for the ancient dog sequences. The x-axis represents the length of the dog mitochondrial genome in base pairs. The y-axis is the number of reads covering each site of the mitogenome.
Figure C.4: Ancient dog damage patterns. Top plots: Base frequency 50 and 30 of strand breaks. The gray brackets indicate start and end of molecules (strand breaks). Purines (A and G) show an elevated frequency before strand breaks. Bottom plots: C to T and G to A nucleotide misincorporations at the first and last 25 bases of endogenous mtDNA fragments from the ancient dog sample (Sawyer et al., 2012).
MS10326.unmerged

A

A

C

C

G

G

T

T

Frequency

Frequency

Frequency

Frequency

Frequency

Frequency
MS10329.merged

A

C

G

T

Frequency

0.0
0.1
0.2
0.3
0.4
0.5

-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10

0.0
0.1
0.2
0.3
0.4
0.5

-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10

0.0
0.1
0.2
0.3
0.4
0.5

-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10

0.0
0.1
0.2
0.3
0.4
0.5
