Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

<u>**The mitochondrial genome of the little spotted kiwi**</u>

A thesis presented in partial fulfilment of the requirements
for the degree of

Master of Science in Molecular Biosciences

Massey University
Palmerston North, New Zealand

Elizabeth Ann Crimp

2010

ABSTRACT

The complete mitochondrial genome of the little spotted kiwi, *Apteryx owenii*, has been sequenced and submitted to GenBank (acquisition number GU071052). A method was first developed to extract pure mitochondrial DNA from one millilitre of fresh blood; as birds have nucleated erythrocytes/red blood cells. The mitochondrial DNA was extracted from the isolated intact mitochondria and the genome was amplified by long-range PCR as 1- 4kb overlapping fragments. These fragments then became templates for the second, short-range, overlapping PCR amplifications and subsequent DNA sequencing. This procedure was first trialled using two millilitres of chicken blood before being successfully applied to the kiwi blood. The complete mitochondrial genome of the little spotted kiwi is ~ 17,020bp long. The gene order is the standard avian gene order first reported for chicken mitochondrial DNA. Phylogenetic relationships show the kiwi is part of the Australasian ratite group with the emu and cassowary. This mitochondrial sequence has been used as part of a larger study of the relationships of other ratite birds (such as moa, emu, cassowary, rhea and ostrich) and the weakly flying tinamou of South America. The implication of this analysis is that the ancestral paleognath was probably flying and that flight was lost multiple times during ratite evolution.

ACKNOWLEDGEMENTS

This thesis is proof that no matter how circuitous the path or how long the journey, you have to pursue your dreams. Life evolves naturally despite your plans, be confident, have faith in yourself and trust your instincts.

I would like to thank my supervisor, David Penny, for his guidance, wisdom, tolerance and infinite patience throughout this project.

I would like to acknowledge Trish McLenachan for her help and suggestions in the laboratory and improvements to this manuscript, Gillian Gibb and also the staff at the Otorohanga Kiwi House. Thank you for your help in completing this thesis. I also wish to thank those who have helped indirectly with their time and suggestions; it has been very much appreciated.

For their financial support I also thank the New Zealand Marsden Fund.

To my friends and my family that have remained through-out, my grateful thanks to all of you. To my family for their love and support, to my friends for the coffee, the meals, the wine, the company and their time – wouldn't have made it without you all.

Thanks to the Crockers, Des, John, Julz, the Kersels and Mike. To those who have moved on, miss you, miss you heaps.

"Live as if you were to die tomorrow.
Learn as if you were to live forever." **Gandhi**

Sylvia
Aunty Paulette

TABLE OF CONTENTS

Abstract i
Acknowledgements ii
Table of Contents iii
List of Figures and Tables iv

Chapter One: Introduction 1
 Project and aims 1
 Background 3
Chapter Two: Materials and Methods 17
 Introduction 17
 Permits 18
 Method 1: Isolating intact mitochondria from red blood cells 19
  Evaluation 20
 Method 2: Extracting the mtDNA from the intact mitochondria 23
 PCR Amplification 23
 Method 3: DNA extraction from feathers 27
 Sequencing 28
 Mitochondrial Alignment used for primers 28
 Primers 29
 Sequence confirmation and alignment 33
 Phylogenetic analysis (Appendix 2) 33
Chapter Three: Results 34
 Summary 48
Chapter Four: Discussion 50

References / Bibliography 56
Appendix 1: Short-range PCR Primers; LSK mitochondrial genome sequence 64
Appendix 2: Phillips *et al.* (2010). *Syst. Biol.* 59, 90-107. 77

List of Figures

Figure 1: The continent of Zealandia. 7

Figure 2: Geological time-scale for the Cenozoic. 16

Figure 3: Sequencer output of overlapping DNA sequences for the LSK mt genome. 43

Figure 4: Graphical representation of the LSK mitochondrial genome. 44

List of Tables

Table 1: Length (in amino acids), Start and Stop Codons of Paleognath Mitochondrial Protein-Coding Genes. 45

Table 2: Length (in nucleotides) of Paleognath Mitochondrial Control Regions, Intergenic Spaces and Complete Genomes. 46

Table 3: Length (in nucleotides) of Paleognath Mitochondrial tRNA and rRNA Genes 47

Chapter 1 - **INTRODUCTION**

Project

This project set out to sequence the complete mitochondrial genome of the little spotted kiwi, *Apteryx owenii*. The mitochondrial genome of birds is approximately 17,000 base pairs (bp) in length and contains a variety of genes encoding 13 proteins, 22 transfer RNAs (tRNA) and 2 ribosomal RNAs (rRNA). The same set of 37 genes has been conserved during vertebrate mitochondrial evolution to date (Mindell *et al.* 1998a).

The procedure required mitochondrial DNA (mtDNA) to be extracted from 1 millilitre (ml) of fresh blood from a kiwi from the Otorohanga Kiwi House, New Zealand. Intact mitochondria were isolated using centrifugation, and then the mtDNA extracted, thus eliminating the nuclear DNA (nDNA) and any possible nuclear copies of the mtDNA.

Using the polymerase chain reaction (PCR), the mt genome was amplified in 1-4kb long overlapping fragments. These fragments were then used as templates for the amplification of shorter, overlapping fragments and DNA sequencing. Most primers for this work were provided by Alan Cooper from his work sequencing two extinct New Zealand moa taxa (Cooper *et al.* 2001). Once this information is obtained, non-invasive methods can be used for sequencing specific short regions from material held in Museums, etc.

Aims

Phylogenetic relationships among closely related kiwi and among populations of kiwi cannot be determined solely from morphologically based taxonomies. Kiwi show little physical differentiation, yet the brown kiwi has the highest levels of geographic structuring of genetic diversity observed in birds (Baker *et al.* 1995). This flightless, nocturnal bird exhibits a strong subdivided population structure, presumably because it has relatively low dispersal capability (compared to flying birds). In a study by Baker *et al.* (1995) the mtDNA sequences showed, with two minor exceptions, diagnostic genetic markers for each of the populations investigated. In vertebrates, mtDNA appears to evolve at a more rapid rate than nuclear genes and this is helpful in resolving populations and other closer related taxa.

Firstly, by sequencing the complete mitochondrial genome of a kiwi, this project allows the identification of small mtDNA regions that show a high diversity. These will then aid in

measuring the diversity of a population and help assign priority for conservation management. The mtDNA is maternally inherited and very rarely is there recombination. It provides a rich source of genetic markers that can detect recently evolved population structure and help reconstruct intraspecific phylogenies of matriarchal lineage.

Secondly, when the displacement (D)-Loop is sequenced it will allow this highly variable region to be sequenced for a number of birds from several populations. The D-Loop contains variable blocks that evolve ~ 4-5 times faster than the entire mtDNA molecule. This is important from a conservation viewpoint where a major aim is to conserve as much genetic diversity as possible. Populations that are genetically distinct need to be recognized for separate conservation research and management. Thus developing primers for the D-loop will allow research on the other kiwi species using fresh or frozen blood samples, solid tissue and feathers. Extracting DNA from feather tips using proteinase K and a 24 hour digestion produced mtDNA that was suitable for PCR amplifications (see Method 3). This non-invasive method makes collection, transportation and storage simpler and overall costs lower.

Thirdly, this project is part of a larger collaborative study with the University of Adelaide. It will complement a study at Adelaide where Alan Cooper has sequenced two complete extinct moa mitochondrial genomes (Cooper *et al*. 2001). Comparison of complete ratite genomes and their divergence dates can further the study of the origin of all bird life and also help determine the age of modern avian orders. The ostrich (Härlid *et al.* 1997) and the rhea (Härlid *et al.* 1998) mitochondrial genomes have been sequenced and these birds are thought to have arisen in the Late Cretaceous period when dinosaurs existed (Cooper and Penny, 1997; Cooper *et al*. 2001). The kiwi, emu and cassowary first appeared in the southern hemisphere. The kiwi data could help test the hypothesis that Gondwana was the original centre of avian evolution because there was a greater diversity of birds there than in the northern hemisphere (Cracraft, 2001). The earliest fossil records or basal lineages of many avian taxa are Gondwanain (Cooper and Penny, 1997; Cracraft, 2001). This kiwi sequence will also add information to the study of evolutionary relationships and evolutionary processes in New Zealand.

Complete mtDNA genomes combined with morphological analyses, palaeontological studies, geological and ecological data will allow more thorough testing than is currently possible for

a range of evolutionary hypotheses. This international collaboration ensures that we will get the maximum amount of information from one additional sequence.

Background

The earliest bird fossils are dated ~150 million years ago (Mya). The ratites, large flightless birds including the kiwi, are descended from this flying stock and are a very ancient lineage (Stapel *et al.* 1984; Cooper and Penny, 1997), estimated to have a late Cretaceous origin (Cooper *et al.* 2001; Cracraft, 2001). They have retained many of the adaptations associated with flight (Feduccia, 1999) and have a flat breastbone (keel-less sternum). The ratites and the flighted tinamous (Paleognathae) share the unique characteristic of an archaic palate and are taxonomically separated from all other birds (Neognathae). The present flightless ratites are found only in the southern continents. The modern avian orders (Neornithes) are estimated to have originated in the early Cretaceous (Cooper and Penny, 1997).

There are 10 living species of ratites – the greater rhea (*Rhea americana*) and the lesser rhea (*Rhea pennata*) in South America; the ostrich (*Struthio camelus*) in Southern Africa and formerly in Eurasia; the emu (*Dromaius novaehollandiae*) in Australia; three species of cassowary (*Casuarius casuarius, Casuarius bennetti, Casuarius unappendiculatus*) in north-eastern Australia and New Guinea and three generally recognized species of kiwi (*Apteryx mantelli, Apteryx owenii, Apteryx haasti*) in New Zealand. The major groups appear to have distributed and evolved due to the break-up of Gondwanaland (Cracraft, 1974; Hedges *et al.* 1996), which began about 150 Mya (Smith *et al.* 1994). The exceptions may possibly include the ostrich, the kiwi (Cooper *et al.* 2001) and also the moa (Phillips *et al.* 2010). (For a Geological time scale, see p. 16.).

For extinct ratite species, Baker *et al.* (2005) list 14 major monophyletic lineages of moa from New Zealand. In addition there are two genera of elephant-birds (*Aepyornis,* with four species*; Mullerarnis,* with three species) from Madagascar. van Tuinen *et al.* (1998) suggest that morphological and eggshell evidence show a close relationship between elephant-birds and ostriches. Cooper *et al.* (2001) placed the elephant-bird among the derived ratite taxa but not as a recent divergence from the ostrich or any other ratite lineage. A ratite fossil from the Paleogene of Seymour Island, an island located on the north-eastern tip of Antarctic Peninsula (Tambussi *et al.* 1994) may have been an isolated lineage. Seymour Island and South

America probably remained close until 45Mya, permitting discontinuous dispersal between Antarctica and South America.

The phylogeny and classification of the ratite birds has been debated for over a century and was comprehensively reviewed by Sibley and Ahlquist (1990), Lee *et al.* (1997) and Cracraft (2001). The tinamou and the ratites are a monophyletic group (Caspers *et al*. 1994; Stapel *et al.* 1984) and it was thought that the weakly flying tinamou, of Central and South America, were the ratites closest living relative (Prager *et al.* 1976; van Tuinen *et al.* 1998; Sibley and Ahlquist, 1981, 1990). On the basis of morphological and biochemical studies, Cracraft (1974) thought that the tinamou and the ratites had a common ancestry and the tinamou were closest to the common ancestral type. Egg-white protein electrophoresis (Sibley and Frelin, 1972) and mitochondrial 12S rRNA gene analysis (Cooper *et al.* 1992) have also linked the tinamou to Galliformes (chickens, turkeys and relatives). The tinamous are a Gondwanan group and the only contemporary flying birds with a keeled sternum and an archaic palate (Prager *et al.* 1976). Recent analyses of nuclear genes (Harshman *et al.* 2008) and our molecular data (Phillips *et al.* 2010), confirms that the flighted tinamou are nested within the ratites (see Appendix 2).

Phylogenetic analyses of the mitochondrial 12S rRNA gene (~ 400bp) including sequences from four extinct moa species, showed that moa represented one of the earliest divergences among the ratites (Cooper *et al*. 1992). This aspect at least has support from morphology (Cracraft, 1974; Bledsoe, 1988). The moa and the kiwi are thought to have existed together in New Zealand during the Pleistocene, 2 Mya, and the moa and kiwi each represent monophyletic groups (Cooper *et al.* 1992). Kiwi more recently shared an ancestor with the rhea and the Australian ratites (Prager *et al.* 1976; Sibley and Ahlquist, 1990; van Tuinen *et al.* 1998; Phillips *et al.* 2010). Monophyly of these living Australasian ratites also has molecular support from transferrin immunological distance data (Prager *et al.* 1976); DNA-DNA hybridization data (Sibley and Ahlquist, 1981, 1990); 12S rRNA (Cooper *et al.* 1992); 12S rRNA and C-MOS (Cooper and Penny, 1997; Cooper, 1997); cytochrome c oxidase (CO), subunits: I, II, III, Cytb, 16S and 12S rRNA (Lee *et al.* 1997); mtDNA protein-coding sequences (Cooper *et al.* 2001); complete mitochondrial genome sequences (Haddrath and Baker, 2001). A *Rhea* – Australasia connection also agrees more with earth history than does a *Struthio* – Australasia or *Rhea* - *Struthio* relationship (van Tuinen *et al.* 1998).

This realationship would indicate New Zealand was colonized twice by ancestors of ratite birds. While moa may possibly have had a vicariant origin or flown to New Zealand, kiwi arrived later from Antarctica or Australia. Kiwi are unique to New Zealand, however, the oldest kiwi fossil from the Marton Fauna is dated 1Mya (Worthy & Holdaway, 2002). Although New Zealand has a relatively good marine fossil record, there is a paucity of terrestrial animal fossils prior to the Pliocene, probably because much of New Zealand was submerged during the Oligocene (Cooper and Cooper, 1995). The lack of phylogenetic informative fossil data is a disadvantage to the study of their evolution as there are only limited reliable geological calibration points.

The most recent common ancestor (MRCA) of the Australasian group; the kiwi, cassowary and emu, would have been in Australia before its final separation from the Antarctic 33-34Mya (Cracraft, 2001; Exon *et al*. 2009). The emu and the cassowary are a monophyletic group, estimated to have diverged from the kiwi 65–70Mya, the emu / cassowary split 33–39Mya (Cooper *et al.* 2001). The dates are supported by the mid to late Oligocene fossils *Emuarius gidju* and *Emuarius guljaruba,* from Australia, that showed emu and cassowary lineages had separated sometime before 25Mya (Boles, 1992; 1997; 2001), probably around 30-35Mya (personal communication from W. Boles in Cooper *et al.* 2001). Phillips *et al.* (2010) estimated divergence dates of 60Mya and 21Mya respectively. The cassowary migrated as far as New Guinea, before 50,000 years ago, when Australia was connected to New Guinea due to low sea levels during the Ice Ages.

New Zealand drifted away from West Antarctica and the east coast of Australia very early in the break-up of Gondwanaland, 80Mya (Veevers, 1991) to 84Mya (Mayes *et al*. 1990). As New Zealand drifted away from Australia, the last point of separation of the land masses was the north of the North Island. For a limited time there may have existed a pathway between Australia, New Caledonia and New Zealand. This pathway was no longer present at the time of the marsupial migration from South America and the greatest assemblage of marsupials today live in Australia. Herzer (1998) provides data for a possible discontinuous pathway from New Caledonia to the Norfolk Ridge, the Reinga Ridge, then to Northland (New Zealand) during the Oligocene to Middle Miocene (34-14Mya). The sequence of uplift and subsidence along this route may have favoured a one-way, north to south migration.

In principle, if kiwi arrived after New Zealand was fully separated from Antarctica / Australia and kiwi were non-volant (flightless), land bridges or island chains would have been required to facilitate migration. With sea level fluctuations and plate tectonic movement, more land than the present day could have existed between Antarctica, New Zealand and Australia. Possibly the Southeast Indian Ridge, Indian–Antarctic Ridge, South Tasman Rise, Macquarie Ridge and Campbell Plateau were available pathways. The highest sea levels in the Mesozoic and Cenozoic occurred in the Middle to Late Cretaceous (105-70Mya) and this was followed by a sharp increase in the total emergent land area 60–70Mya (Hedges *et al.* 1996). This notable sea-level drop would have exposed the continental shelves and possibly any ridges and sub-aerial islands. Island hopping, and swimming (Peat, 1999), could have occurred either from Antarctica to New Zealand and also to Australia, or from Antarctica to Australia, New Caledonia and then to New Zealand. Lord Howe Rise, the Three Kings, Colville and the Kermadec Ridges (island arcs) may also have been used as possible pathways. Many other terrestrial species have come to New Zealand since the Cretaceous break-up of Gondwanaland (Goldberg *et al.* 2008). Most other non-flying terrestrial vertebrates have been known to disperse over long distances of open-ocean by rafting on flotsam (Diamond, 1987; Hedges, 1996; de Queiroz, 2005) and this could also apply to flightless / weakly flying birds. Circumpolar dispersal by rafting on floating vegetation is known; some 200 species of plants are currently recognized as being native to both Tasmania and New Zealand (Jordon, 2001).

The present New Zealand land area is only a small fraction of the original continental crust that broke away from Gondwanaland (see Figure 1). Zealandia, 65Mya, included New Caledonia, Lord Howe Rise and Norfolk Ridge (northwest), Challenger Plateau (northwest off Nelson), Chatham Rise (southeast), and south past the present day Stewart Island to Campbell Plateau, and New Zealand's sub-Antarctic islands. This area comprised of ~ 10 x the present New Zealand land area. At 61Mya, Zealandia lay at latitude close to that of present day Antarctica and by 60Mya New Zealand was ~ 1500km east of Australia. At 42Mya New Zealand was still 1100km further south, closer to the Antarctic, at the same latitude as the southern tip of South America (Wikipedia).

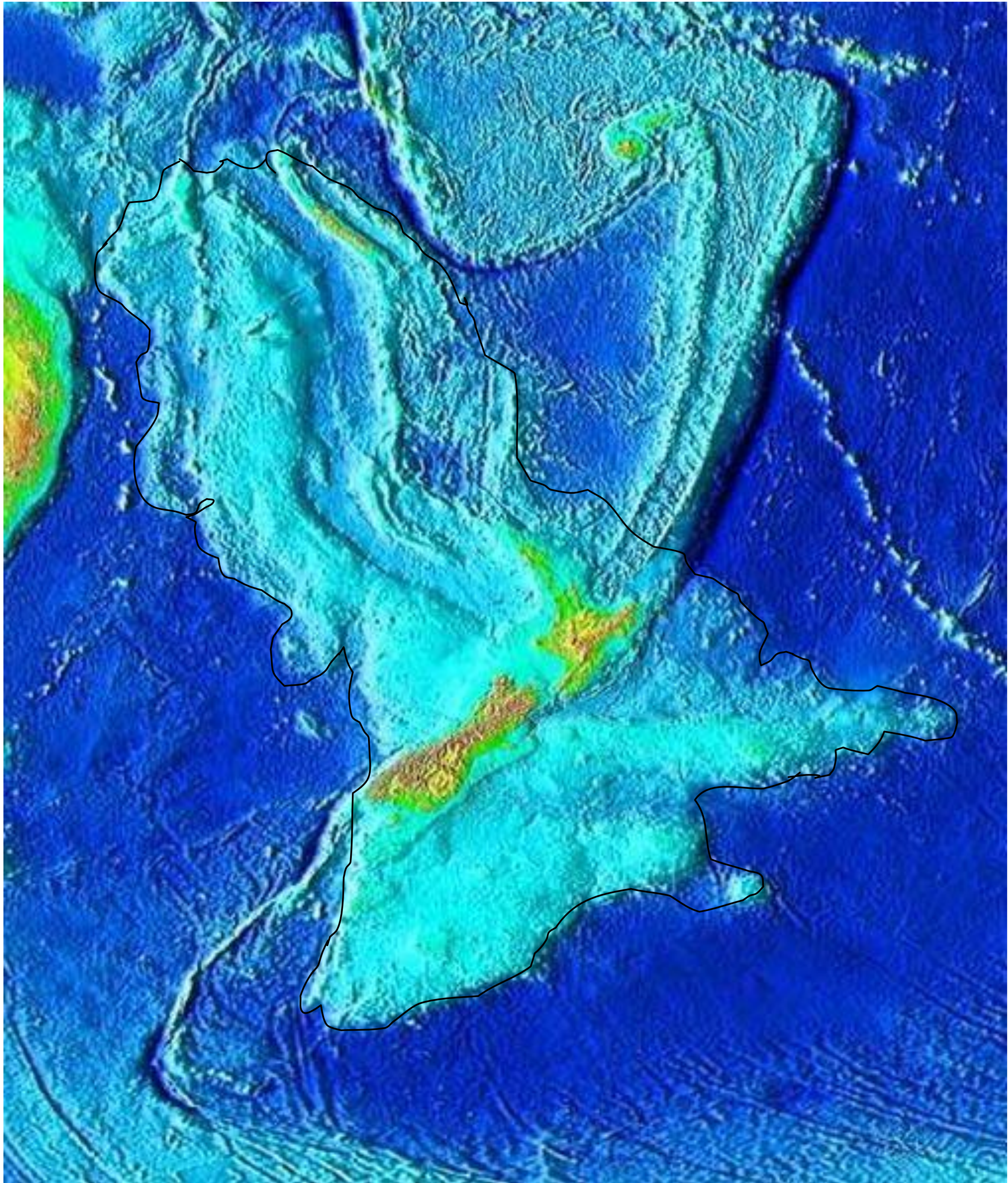

Figure 1:The continent of Zealandia
(Wikipedia,http://en.wikipedia.org/wiki/Zealandia_(continent)).

Only New Zealand, the Chatham Islands, Lord Howe Island, Norfolk Island and New Caledonia remain above the water, 93% of the original landmass is still submerged. Australia (upper left), Fiji and Vanuatu (top centre) and the linear ridges running north and south are not considered part of the continent.

Warm subtropical waters flowed between Australia and New Zealand but with the separation of South America and then Australia's Tasmanian land bridge from Antarctica, 33-34Mya, (Exon *et al.* 2009) the existing ocean current pattern was changed. This separation of the land masses allowed the development of Antarctic circumpolar marine currents (ACC). These well established strong prevailing westerly currents are now a prominent feature of the Southern Ocean. The ACC and associated cold westerly winds led to the thermal isolation of Antarctica and as this intensified the climate cooled. The ice cap rapidly formed close to the Eocene/Oligocene boundary ~34Mya (Zachos *et al.* 2001). In New Zealand the subtropical temperature began slowly cooling then dropped sharply in the late Pliocene, leading to the onset of the Pleistocene ice age.

A ratite / tinamou sister grouping (Phillips *et al.* 2010) and the discovery in Laurasia (North America and Europe) of approximately 50 to 60 million-year-old fossils of small flying paleognathous birds / lithornithids (Houde 1986, 1988) would imply the ancestor of this ratite was flying. If the lithornithids are descended from the ostrich, the rhea or the last common ancestor (LCA) of the moa–tinamou / Australasian clade (Phillips *et al.* 2010), it could imply that flightlessness through evolutionary change evolved several times among the ratites. With the moa and tinamou most likely being a sister group to the Australasian clade, and possibly also the elephant-bird (Phillips *et al.* 2010), this would imply that the moa and the kiwi may have flown over to New Zealand, wind / island assisted, after New Zealand had separated from Antarctica. When a species of land bird colonises an island that no four-footed mammalian predators have been able to reach, development of a flightless and often larger form is favoured (Steadman, 2006). Feduccia (1996) suggests they can also change their morphology in adaptation to a ground dwelling existence over a short time span, probably within 1-2My. This has support from island faunas in New Zealand (moa, kakapo, adzebill, rails, weka, flightless goose, takahe etc.), where there was no strong continuous selection for flight maintenance (Trewick, 1996, 1997). The kiwi now has only small vestigial wings and the extinct moa had none. There is a current study by D M Lambert (Griffith University, Brisbane, Queensland.) to determine what mutation(s) led to loss of wing development.

By 70Mya the smaller dinosaurs appear to have been declining (Penny and Phillips, 2004). With no large predators left and with only one, as yet unidentified, mouse-size ground-dwelling mammal fossil from Miocene (16-19Mya) sediments (Worthy *et al.* 2006), large

flightless birds would have been in a dominant position. Moa were predominantly herbivorous and are classified into different genera (Worthy and Holdaway, 2002). Kiwi could have undergone speciation, taking over the role of a small nocturnal omnivorous / carnivorous mammal, following the availability of the densely forested floor niche. Interestingly, other New Zealand native birds are nocturnal; living examples include the kakapo and morepork. New Zealand's only known endemic land mammals were three species of bats (Worthy and Holdaway, 2002).

The New Zealand coastline has fluctuated widely due to sea level changes, plate-tectonic movement and volcanic activity. Around 80Mya Zealandia was low lying, lushly forested and half the size of present day Australia. During the late Oligocene drowning/marine transgression (36-24 Mya), possibly only a string of archipelagic and ephemeral islands were left above water (29-23Mya, a period of ~ 6My). These relatively small, low lying islands, with wide, shallow marine shelves, were extremely vulnerable to the many oscillations in the sea levels. This would have limited the range of niche space and diversity and led to a widespread genetic 'Oligocene bottleneck' for the two ratite orders (Cooper and Cooper, 1995). Extensive extinction of many other lineages or earlier diversifications of many endemic species would have occurred in the Oligocene. The present diversity probably evolved slowly from each of these isolated, small communities of surviving ancestors, after the increase in land area and new habitat availability. Worthy *et al.* (2006), Knapp *et al.* (2007) and Jones *et al.* (2009) acknowledge there is a biological case for some remaining emergent land during the Oligocene, however, some authors argue that the New Zealand sector may have been completely submerged for some or all of that time (Waters and Craw, 2006; Landis *et al.* 2008). Trewick *et al.* (2007) suggest that New Zealand needs to be reconsidered as a separate entity to Zealandia to answer the question 'When did New Zealand emerge?' There is evidence to suggest that a large island existed to the east of the Three Kings Ridge between 38 – 21Mya (Meffre *et al.* 2006). However, it is not known whether there was any movement of the biota to New Zealand from this large island.

In the late Miocene and Pliocene the stress regime between the Pacific and Australia tectonic plates changed from transcurrent to transpressional. This caused lateral elongation along the Alpine Fault that lead to the current New Zealand coastline (King, 1998) and to the rapid uplift and formation of New Zealand´s axial mountain ranges (Cooper and Cooper, 1995). The Southern Alps, that are still currently growing, created highland and lowland

environments and divided the climate in the South Island. On the west coast the wet rainforests developed with the increase in rainfall and on the sheltered east coast a warmer, drier climate developed. During the Pleistocene period the South Island was significantly affected by repeated glacial, interglacial cycles. Large areas of the South Island were covered by ice that left deforested areas bare or dominated with grasses and shrubs when the ice retreated, ~ 14,000ya. In the North Island volcanic activity was extensive and widespread during the Quaternary Period and in the Central Plateau it is still active today. These events resulted in the kiwi populations becoming fragmented, as kiwi are highly territorial and disperse relatively slowly. Breeding among themselves, in their isolated communities, the kiwi shared a small gene pool and this favoured natural selection adaptations to the local environment. It also left small populations vulnerable to extinction. The brown kiwi do not normally interbreed and have become separate species, but they have remained morphologically conservative. The majority of the morphological characters shared by moa and kiwi are ancestral for birds (Cracraft, 1974; Bledsoe, 1998) and phylogenetically uninformative, but they could have evolved convergently from ancestors that were not highly specialized (Cooper *et al.* 1992). The Pleistocene period, settlement of humans (~ 800 years ago), habitat destruction, hunting, and the recent introduction of terrestrial mammalian predators have decimated the kiwi populations In 1896 all kiwi were declared protected, however, kiwi numbers have continued to decline and this downward trend has continued until recently. Since 1991 a number of trusts and programs have been established to protect the remaining, endangered kiwi populations.

The three kiwi are congeneric and are traditionally divided into three species, the great spotted or Roroa (*A. haastii*), the LSK (*A. owenii*) and the brown (*A. australis*). Kiwi is a Maori word and is, therefore, both singular and plural. Apteryx is a Latin word meaning wing-less, without wings. Burbidge *et al.* (2003) recommended full species recognition for the North Island brown (*A. mantelli*), Okarito brown or Rowi (*A. rowii*), and suggested *A. australis* be restricted to the tokoeka, the southern South Island browns; *A. a.* 'Haast' and *A. a.* 'Fiordland / Stewart Island'. The Stewart Island tokoeka is also known as *A. a. lawryi*. The eastern brown kiwi, *Apteryx* 'East South Island', is listed as extinct (Holdaway *et al.* 2001), but they may not have been a separate species. An ancient DNA study by Shepherd and Lambert (2008) grouped the eastern brown kiwi with the tokoeka. All taxa are currently listed as threatened (Colbourne, 2005).

The basal southern South Island clade appears to be the remnant of the original ancestral population. Baker *et al.* (1995) suggested the brown kiwi colonised northward, diverging at Okarito before crossing the land bridge to the North Island. There is a concordant split in molecular phylogenies between the brown kiwi in the southern South Island and those elsewhere in New Zealand, confirming the Okarito rowi is the sister group to the North Island clade. In pre-human times the Okarito rowi and the North Island LSK (L.D.Shepherd, personal communication) were the only kiwi species present in the lower half of the North Island (Shepherd and Lambert, 2008). Burbidge *et al.* (2003) estimated the brown kiwi diverged from the spotted kiwi around 16Mya, the great spotted (GSK) and LSK 6Mya, the southern tokoeka (Fiordland and Haast) and Okarito rowi 8Mya, rowi and North Island brown 6Mya and the Stewart Island from the southern tokoeka 4Mya. The Fiordland and Stewart Island tokoeka show high levels of genetic similarity, suggesting gene exchange across the Foveaux Strait has not been an obstacle, but the geographically isolated Haast tokoeka is highly genetically distinct (Herbert and Daugherty, 2002). The diversification of the spotted sister group and the brown kiwi lineages occurred in a similar time frame. Brown kiwi were once widely distributed over all New Zealand until the populations became separated when the land bridges disappeared between the three islands, due to rising sea levels (~ 9,000ya), and natural landscape changes kept them isolated. Baker *et al.* (1995) found high population structuring in the matriarchal lineage of cytochrome b (Cytb), with virtually every population possessing private alleles, and similarly, nuclear allozyme allele frequencies showed levels of population subdivision that are equal to the highest seen in vertebrates. This confirms that gene flow between kiwi populations is limited. The GSK, found only in the northwest of the South Island, had a similar pre-human distribution. A study by Herbert and Daugherty (2002), found the GSK population to have moderate levels of allozyme allele diversity with no apparent geographic patterning, possibly indicating limited gene flow with genetic drift in small geographically isolated populations.

The LSK, also known as the little gray kiwi, is named after the 19th century anatomist Sir Richard Owen. The smallest living ratite, it did not survive the introduced mammalian predation. The distinct North Island subspecies of the LSK (*Apteryx owenii iredalei*) became extinct in the late 1800's. L.D. Shepherd (Massey University, New Zealand) has studied the ancient DNA of the North Island LSK. The last known LSK to inhabit the South Island was recovered in 1938. Despite differences in morphology, genetics and ecology, one hybrid (between a female Okarito rowi and probably a male LSK) was found near Franz Josef in

1993 (Peat, 1999), in an area where the LSK population had approached extinction. The hybrid parentage was confirmed with allozyme variation data (Herbert and Daugherty, 2002). Hybridization may have occurred under extreme mate shortage conditions or was a common occurrence in the area. There was a West Coast LSK that was recognised and named *A. occidentalis.* It was larger with more distinct markings than *A. owenii* and had dark coloured legs like *A. haastii* but with a paler plumage. The Franz Josef hybrid was firstly identified morphologically as a small GSK, so was *A. occidentalis* a recurring hybrid or a distinct species (that is now extinct) and the other possible male parent of the hybrid? Herbert and Daugherty (2002) discuss the kiwi naming issues mentioned in this paragraph. The hybrid bird was transferred to Mana Island in 1992 and has not bred (Colbourne, 2005).

Holdaway *et al.* (2001) in their biogeographic analysis have listed the LSK as extinct because there are no natural mainland populations. The present LSK, once widespread in the western side of the South Island, is now reduced to small isolated populations on several offshore predator-free island sanctuaries around New Zealand. Recently LSK were introduced into their first mainland wildlife sanctuary. These islands and the mainland sanctuary have been populated from LSK found on Kapiti Island, a population that had already experienced founder effects. When the original founder population of LSK arrived on Kapiti is not known as they could have been there when it was once connected to the mainland (~ 12,000ya) or been introduced later, either before or after European arrival. There is no record of an introduction of LSK to Kapiti Island and the results of an allozyme variation study is consistent with either a population that increased slowly from a few birds or with a long-isolated natural population reduced to a low level (Jolly and Daugherty, 2002). The problems facing the 'range restricted' LSK will be compounded by genetic drift (small populations), a possible founder effect (diversity is foundered on ~ 5 birds), and predator vulnerability. The conservation genetics and long-term population viability of the LSK is being monitored by Kristina Ramstad (personal communication, Victoria University, Wellington, New Zealand).

The only other naturally occurring population of LSK was on D'Urville Island but only two LSK have been found to have survived. The D'Urville Island LSK are genetically very similar to the LSK on Kapiti Island (Herbert and Daugherty, 2002; Jolly and Daugherty, 2002) and have the same partial (~ 400bp) 12S rRNA gene sequence (Cooper and Cooper, 1995). The two birds from D'Urville Island, a female and a male, were transferred to Long

Island, in the Marlborough Sounds, in the 1980's along with two LSK from Kapiti Island, and are breeding successfully (Jolly and Daugherty, 2002; Colbourne, 2005).

The ratite biogeographic history can help reconstruct the evolutionary pathways by providing specific time frames / intervals when migration between the continents was possible. There is nuclear and mitochondrial gene evidence for the divergence of extant avian (and mammalian) orders around 100Mya (Hedges *et al.* 1996). Härlid *et al.* (1997) suggested the evolutionary divergences among the ratites probably occurred not earlier than 90Mya and estimated the chicken / ostrich divergence to be about 80-90Mya. Janke and Arnason (1997) estimated the paleognath ostrich and neognath chicken divergence at ~ 92Mya.

There is debate over whether rhea or ostrich is basal within ratites. Phillips *et al.* (2010) agreed with Prager (1976), Sibley and Ahlquist (1981, 1990) and Härlid *et al.* (1998) that the ostrich is basal. van Tuinen *et al.* (1998) found rRNA gene trees supported a basal ostrich and protein-coding trees supported a basal rhea, but thought that the ostrich is most basal among the living ratites. Lee *et al.* (1997) placed the rhea basal. Cooper *et al.* (2001) estimated the divergence time of the basal rhea from the other ratite taxa at 84-94Mya, the moa 82Mya (Zealandia's separation date), the ostrich / kiwi divergence 72-78Mya and the kiwi 65-72Mya. Haddrath and Baker (2001) also examined the mtDNA from two extinct moa and found the moa to occupy the basal position. Additional taxon sampling is required to resolve this issue. Missing lineages would fill in the gaps of the evolutionary tree, their ancient origin, early history and help resolve their phylogenetic relationships. Improved methods using ancient DNA will provide additional sequences from extinct taxa to add to the evolutionary tree. This could reduce the long branches of the ostrich and rhea, decrease the effects of the long-branch attraction and stabilize the tree (Hendy and Penny, 1989; Lee *et al.* 1997).

Morphological studies, molecular and anatomical analysis have not been conclusive (Lee *et al.* 1997). The morphological tree of Cracraft (1974) found the kiwi is basal and related to the moa, whereas, Bledsoe (1998) placed the kiwi in a more derived position close to the emu and the cassowary. The combined morphological and molecular data tree of Lee *et al.* (1997) was identical to their morphological only data tree, where the kiwi and moa lineages were basal and the rhea and ostrich were a sister group to the cassowary and emu. Bourdon *et al.* (2009) also combined both data sets and placed the moa and kiwi as the basal sister group, with the elephant bird and ostrich as a sister group to the rhea, emu and cassowary clade. There are

also morphological discrepancies that could be due to the subjective interpretation of shared traits, as the majority are considered 'primitive' (Bledsoe, 1988). Härlid and Arnason (1999) suggested the morphological characteristics were probably acquired through neoteny (the persistence of juvenile features in the adult form) and they were descendants of flying ancestors. The DNA analysis also suggested there was no real difference between the paleognaths and other birds and Struthioniformes were closer to Galliformes. Almost all the anatomical character differences could be attributed to neoteny (Härlid and Arnason, 1999) and these features may also be associated with flightlessness; a flat breastbone, no keel, underdeveloped breast muscles and wings (with the exception of the moa).

The unrooted morphological and molecular trees of Lee *et al.* (1997), after excluding the moa, are identical and show that the early divergences within the ratites occurred during a relatively short period. Close divergent dates within the ratites has resulted in short internodal distances (branch lengths) in the tree and this makes resolving relationships difficult without additional taxa and DNA sequences. Phillips *et al.* (2010) results include the tinamou within the ratites and estimated the ratite divergences occurred between 83 – 60Mya and the emu / cassowary divergence at 21Mya. The phylogenetic resolutions using mtDNA sequences can provide a temporal framework for biogeographical events during the break-up of Gondwanaland and the possible flight loss(s) among ratites at the Cretaceous-Tertiary (K/T) boundary and during the resulting extinction event(s).

The ratite 12S rRNA (domain III) is one of the most conserved areas of the mitochondrial genome (Mindell and Honeycutt, 1990). Cooper *et al.* (1992) using only a 12S rRNA data set included four extinct moa species to the extant ratites and found the rhea was basal. Different tree building methods using different data sets from 12S rRNA, combined with addition data sets, have resulted in different molecular trees. Phylogenetic analyses of Cooper and Penny (1997) found the kiwi and ostrich reversing positions, Lee *et al.* (1997) placed the rhea as basal and van Tuinen *et al.* (1998) placed the ostrich as basal. Results using large data sets such as concatenated sequences of genes, gives more robust results than analyses performed on individual genes (Cao *et al.* 1994) that can support different phylogenetic hypotheses (see Cooper, 1997). Long sequences combining multiple nuclear genes and complete mtDNA genomes should permit a more accurate molecular estimation of the divergence dates of the various ratite taxa. It will also avoid the bias of using only one maternally inherited marker and give more confidence to molecular trees. When molecular and morphological data,

palaeontological and phylogenetic information are combined with more recent geological estimates for the separation of Gondwanan continents and when possible land availability is explored further, it should be possible to reach a more accurate conclusion about the ratite evolutionary tree.

Cracraft (1974, 2001) and van Tuinen and Hedges (2001) favoured exclusive land-based (vicariant) distributions of the ratites and this aspect has partial support (Bledsoe, 1988; Sibley and Ahlquist, 1990; Cooper *et al.*1992, 2001; Haddrath and Baker, 2001). With the molecular data providing more recent phylogenetic dates that post-date the break-up of Gondwanaland (Phillips *et al.* 2010) a biogeographical dispersal hypothesis is now more likely.

Ratites, including the kiwi, are important from a conservation viewpoint, and are scientifically interesting in understanding dispersal around the Southern continents. The mitochondrial genome of the LSK was successfully isolated and sequenced. A manuscript is presented in Appendix 2 in which the mitochondrial genome, sequenced here, is combined with other new ratite mitochondrial genomes and analysed to give a better understanding of ratite dispersal and evolution.

Current IUGS ratified (2009)

Era	Period	Epoch & Subepoch		Age	Age (Ma)	GSSP
Cenozoic	Quaternary	Holocene			0.012	◀
		Pleistocene	'L'	'Tarantian'	0.126	
			'M'	'Ionian'	0.781	
			'Early'	'Calabrian'	1.806	◀ Vrica, Calabria, Italy
				Gelasian	2.588	◀ Monte San Nicola, Sicily, Italy
	Neogene	Pliocene		Piacenzian	3.600	◀
				Zanclean	5.332	◀
		Miocene		Messinian	7.246	◀
				Tortonian	11.608	◀
				Serravalian	13.65	
				Langhian	15.97	
				Burdigalian	20.43	
				Aquitanian	23.03	◀
	Paleogene	Oligocene		Chattian	28.4	
				Rupelian	33.9	◀
		Eocene		Priabonian	37.2	
				Bartonian	40.4	
				Lutetian	48.6	
				Ypresian	55.8	◀
		Paleocene		Thanetian	58.7	
				Selandian	61.7	
				Danian	65.5	◀ El Kef, Tunisia

Figure 2: Geological time-scale for the Cenozoic, from Gibbard *et al.* (2010).

The cretaceous period extended from ~ 145Mya to ~ 65.5Mya.

The Tertiary includes the paleogene and the neogene.

Chapter 2 – **MATERIALS AND METHODS**

Introduction

Reconstruction of vertebrate phylogenies has been advanced using PCR and mtDNA. Mitochondrial genomes are small, there are a high number per cell and they are maternally inherited. Very rarely is there recombination and they exhibit an evolutionary rate 5-10 times higher than single-copy nuclear genes, resulting in a greater sensitivity to fluctuations in overall population size. Mitochondria are also an easily isolated, purified and directly sequenced source of genetic information outside of the nucleus. The nDNA has an independently inherited set of molecular characters and evolves slower than mtDNA. This complements the mtDNA, but nDNA has a low ratio of single-locus genes and this makes sequences difficult to amplify in ancient DNA (Cooper, 1997). The mtDNA is also haploid as it has a clonal inheritance and does not possess the gene families of the diploid nDNA.

Pure mtDNA was extracted from the fresh blood of two little spotted kiwi (LSK). These were identified at Otorohanga Kiwi House as kiwi $R^{021}/_{004}$ and $Y^{R35}/_{031}$ which I will call simply R1 and Y2 in this thesis. The complete mitochondrial genome of LSK Y2 was to be amplified by long-range PCR in six overlapping fragments. The four long fragments obtained were 3 - 4kb in length and covered the genome from Leucine (L3450) to Cytochrome b (H16065). Another fragment covering from Phenylalanine (L1278) to 16S rRNA (H3420) was obtained from the feather tips of LSK Y2. These fragments then became templates for shorter, overlapping PCR amplifications for sequencing. The D-loop did not amplify as a large fragment and was sequenced in small fragments.

Laboratory Protocols

To prevent cross-contamination from the other mammalian and avian mtDNA extractions and sequencing, extreme care had to be taken in the laboratory. The allocated bench area was continually cleaned with ethanol and all personal equipment and reagents were kept separate and DNA free. Every extraction procedure and all sets of PCR amplifications were run with negative controls to detect reagent and any other possible DNA contamination. Only sterilized milliQ H_2O was used and some reagents required autoclaving. Except where otherwise stated in the text all chemicals were BDH and Anala R grade, supplied by BioLab Scientific LTD.

Permits

The project objective is to sequence the complete mitochondrial genome of the LSK, *Apteryx oweni*. The procedure required mtDNA to be extracted from a fresh blood sample. Following a discussion with Dr Murray Potter, Kiwi advisor of Ecology in the Institute of Natural Resources, Massey University, it was agreed that the quantity of blood that would be taken from each kiwi would not exceed one millilitre (1ml). Two captive kiwi were bled in case a second sample was required.

The method for extracting mtDNA from blood was first tested using 2ml of fresh chicken blood extracted in heparin. Permission was obtained from the Massey University Poultry Research Unit and was arranged with Don Thomas, Poultry Research Officer and veterinarian Debbie Anthony of Veterinary Clinical Sciences, Massey University.

Mr Chris Smuts from Technical Services – Fauna at Waikato Conservancy, Department of Conservation was contacted to obtain authority to extract blood from absolutely protected wildlife (pursuant to Section 53 of the Wildlife Act 1953). This was granted by Gregory Martin of Waikato Conservation (dated 17-Nov-1997). The LSK were housed at the Otorohanga Kiwi House, Otorohanga. Permission was obtained from the following;

Mr Roy Dench, President of the Otorohanga Zoological Society,
Eric Fox, Curator of the Otorohanga Kiwi House,
Megan Clemance, Veterinary Surgeon of Otorohanga and
Dominic Fortis, Head Bird Keeper and kiwi handler.

1ml of blood was collected in EDTA (ethylenediamine-tetraacetate acid) from kiwi R1 and 0.9ml of blood was extracted in heparin from kiwi Y2. Eight feathers were also collected from Y2. Both EDTA and heparin were used as anticoagulants in case there was a downstream interference from heparin (Beutler *et al.* 1990). The blood was collected in 0.5ml sterile vials and remained on ice for the trip back to Palmerston North where the intact mitochondria were immediately extracted and stored at 4° C overnight.

Method 1: ISOLATING INTACT MITOCHONDRIA FROM RED BLOOD CELLS

The method used here was a modification of the purification of mtDNA from HeLa Cells (Higuchi and Linn, 1995).

The blood was removed from the vials by rinsing with PBS [140mM NaCl, 2.5mM KCl (Univar), 8.1mM NaH_2PO_4, and 1.5mM KH_2PO_4 (Univar), pH 7.5 (Micol *et al.* 1996)]. Red blood cells were isolated by centrifugation in a Heraeus microcentrifuge at 2,800 x g for 10 mins. The supernatant and interface of leukocytes (white layer) were removed (birds have nucleated red blood cells). The remaining fraction of red blood cells was washed twice with an equal volume of PBS by centrifugation at 2,600 x g for 10 mins and then resuspended in an equal volume of MgRSB Buffer [10mM NaCl, 1.5mM $MgCl_2$ (Sigma),10mM Tris-HCl, pH 7.5]. The hypotonic buffer causes the red blood cells to swell. The solution was allowed to stand at 4° C for 10 mins.

The cells were then disrupted with a Virtis 45 homogenizer in 5 x 1sec bursts to obtain a 60% - 70% cell breakage to ensure a high yield of mitochondria. Insufficient homogenization or swelling of the cells can lower the yield. The Virtis attachment was washed before, during and after each homogenization, firstly in 2% SDS (sodium dodecyl sulphate), milliQ H_2O 4-5 times and then in 2.5 x MSB buffer [0.525M D-mannitol, 175mM sucrose, 12.5mM EDTA, 12.5mM Tris-HCl, pH7.5]. 2.5 x MSB buffer was added to the sample to give a final concentration of 1 x MSB, and then centrifuged in a Heraeus microcentrifuge at 1,000 x g for 10 mins. This removed cell debris, nuclear and membrane fractions.

The supernatant (4.7ml) was further centrifuged in an Eppendorf standard laboratory bench top centrifuge at 13,000 rpm for 20 mins at 4° C. The opalescent and pink supernatant above the mitochondrial pellet was removed as this contained the microsomal fraction and lysosomes. The pellet was thoroughly suspended in 50µl EDTA-free MSB [10mM $MgCl_2$ (Sigma) and 0.5mM phenylmethyl-sulfonyl fluoride (Sigma)]. This was stored at 4° C overnight.

Digestion

The suspension was digested with 6,000 Kunitz units of DNase I, [Pancreatic DNase (Lyophilised) Amersham] in a heating block at 37° C for 30 mins. This gave a concentration of 1.2Ku DNase I / µl of mtDNA and removed all detectable contaminating nDNA. After digestion, the suspension was washed three times in 10 x volumes of 1 x MSB in an Eppendorf bench top centrifuge, twice at 13,000 rpm for 20 mins at 4° C and once at 11,100 rpm for 20 mins at 4° C. The washed pellet was resuspended in 50µl of 1 x MSB. Long term storage of the intact mitochondrial sample was at -80° C. The working stock solutions of 1:10 and 1:50 dilutions were stored at -20° C.

EVALUATION

The purity of the intact mitochondria and extracted mtDNA were evaluated in three ways; by staining with Janus Green B, staining by DAPI, and also by PCR. The chicken and kiwi intact mitochondria were visualized by Janus Green staining and DAPI staining and also PCR. This was done to confirm there was no other detectable DNA present. The extracted mtDNA of both the chicken and the kiwi were evaluated using DAPI staining and PCR.

Staining

1. Janus Green B Stain (JG-B).

JG-B was used on extracted intact mitochondria to confirm that intact mitochondria were present. Diethyl safranin-azo-dimethyl aniline - Janus Green B, (JG-B) is a redox (oxidation-reduction) dye, colourless in reduced form and green when oxidized, it colours cell organelles capable of performing redox reactions. Intact mitochondria, therefore, have a high specificity towards JG-B staining.

A 1:10,000 solution of Janus Green-B (Eastman Organic Chem, Rochester, N.Y.) in aqueous 0.15M NaCl was used (Packer, 1967). One drop of a 1:50 dilution of the intact mitochondria fraction was put on a slide, smeared and allowed to dry for 1 min. One drop of JG-B solution was then added to the unfixed slide. Very tiny round green dots were visible, by eye, on the slide and these were viewed microscopically at x 200 magnification. This was done with the assistance of Elizabeth Nickless, Head Technician, on the standard light microscope, Zeiss Axiphot compound microscope. This was a positive result for the presence of intact mitochondria.

2. DAPI (4′, 6-diamidino-2-phenylindole).

DAPI stain binds preferentially to A + T rich DNA. It can be used as a highly sensitive and specific fluorescent probe for DNA at the level of individual cells. It will permit detection of a single full-sized mitochondrial molecule, provided it is packaged in a compact form. DAPI staining was done on both the intact mitochondria and the extracted pure mtDNA (Packer, 1967).

A working solution of DAPI was made by dissolving 2mg of DAPI powder in 1ml of absolute methanol in a foiled covered bottle and this was stored at 4° C in the dark. A final concentration of 10% DAPI (v/v) was added to both the intact mitochondria and mtDNA samples. The DAPI stained sample was added to a slide, covered with a cover slip and sealed with a clear sealant. The dye takes 3 mins to develop, the DNA stains blue and the DAPI flakes are yellow. The slide was viewed under high-resolution on the fluorescent microscope, Zeiss epifluorescent (magnification range 630-1,000), with the help of the technician Michiko Kakubayashi and Elizabeth Nickless. The stained slide was then wrapped in tin foil and stored in the freezer ($-20\,^{o}$ C) for a few days and then re-observed.

The stain was used to determine the purity of the extracted intact mitochondria, using a sample of a 10µl aliquot of a 1:50 dilution of intact mitochondria and 1µl of DAPI stain. The stained slide was stored for two days to produce a stronger reaction. Only blue fluorescent dots were visible under the microscope. This showed the mitochondria were intact and no other DNA was present, confirming there was no detectable nDNA.

The same procedure was used for the mtDNA extracted from intact mitochondria. A solution containing a 3µl aliquot of a 1:50 dilution of mtDNA and 0.3µl of DAPI stain was used. Blue fluorescent spots, clumps and thick strands of mtDNA were observed, confirming mtDNA.

3. PCR

This was performed on both the chicken and kiwi intact mitochondria to ensure no other DNA was present. A simple and rapid method for total genomic DNA preparation suitable for PCR is by microwave irradiation (Ohhara *et al.* 1994). This is intended to destroy cell structures and release the DNA, avoiding the problem of entrapment of the target DNA by coagulated organic materials that seem to reduce the accessibility of it to the DNA polymerase.

Extracted intact mitochondria and refrigerated human blood (extracted in heparin from a volunteer and stored in a fridge for two days) were each made into 1 - 5µl aliquots and made up to a 10µl volume with milliQ H_2O in 1.5ml microcentrifuge tubes with tightly fitting lids. The tubes were microwaved, on high, for 2 x 1 min in a Sharp Carousel domestic microwave. After irradiation, the samples were chilled on ice and the PCR reaction mixture, Qiagen PCR Protocol using the Q-Solution, was added [1 **x** Qiagen PCR Reaction Buffer, 250 µM of each deoxy-nucleosidetriphosphate (dNTP, Boehringer Mannheim), 0.5 µM of each primer (forward and reverse), 1 - 5µl aliquots of blood and intact mitochondria, 1 **x** Qiagen PCR Q-Solution, 1 unit of Qiagen *Taq* DNA polymerase]. The samples were made up to a total volume of 20µl with milliQ H_2O. The 'universal' primers, forward primer L14841 and reverse primer H15149 (Kocher *et al.* 1989), for mitochondrial cytochrome b (Cytb) were used to produce a fragment size of ~ 307bp. One drop of paraffin (Sigma) was added to each tube and PCR was performed in a MJ Research PTC-150 Minicycler using the thermo-cycling program of; 1 cycle; 94° C for 2 mins, 35 cycles; 94° C for 1 min / 50° C for 1 min / 72° C for 1 min and finally for sequence extension, 5 mins at 72° C, then an indefinite hold at 4°C

After amplification a 10µl aliquot from each sample was mixed with 1µl of 10 x gel loading buffer, loaded into separate lanes on a 1.5% (w/v) Seakem LE agarose / 1 x TAE buffer gel in 1 x TAE running buffer. A 10µl aliquot of a 1Kb-Plus DNA ladder (Gibco BRL, Invitrogen) was run as a size standard. The gel was electrophoresed at 100V for 50-60 mins. The DNA samples were visualized by ethidium bromide fluorescence on a UV transilluminator and digitally photographed (see p. 25).

The gel showed bands (~ 300bp) in the microwaved human blood samples but no bands were present in the microwaved intact bird mitochondrial samples. The extraction method used to isolate the intact mitochondria inhibited the exposure and amplification of the mtDNA using the microwave method. The microwaved intact mitochondrial sample did not amplify a PCR sequence indicating there was no other detectable DNA present.

The 'universal' primers did amplify the chicken mitochondrial Cytb from the extracted mtDNA (Method 2) and does amplify LSK mtDNA Cytb from total DNA extracted from whole blood (Baker *et al.* 1995). Extracted mtDNA of the LSK was confirmed in Method 3.

<u>Method 2:</u> EXTRACTING THE mtDNA FROM THE INTACT MITOCHONDRIA

The 20µl aliquot of intact mitochondria were diluted 1:1 with milliQ H_2O water and digested in an autoclaved lysis buffer [10mM Tris-HCL (pH 8.0), 1 mM EDTA, 100 mM NaCl] with 0.1 mg/ ml proteinase K and 1% SDS (sodium dodecyl sulfate) (Cooper *et al.* 1992), in a heating block for two hours at 37° C.

Phenol / Chloroform extraction

An equal volume of Tris-equilibrated phenol (pH 8.0) was added to the sample, mixed 1-2 mins by gentle inversion and spun at 13,000 rpm for 3 mins in the bench top centrifuge. The top aqueous layer was removed and added to an equal volume of chloroform ($CHCl_3$). This was mixed gently by inverting for 1-2 mins and spun at 13,000 rpm for 3 mins. The top aqueous layer was removed and the volume measured.

Ethanol precipitation

Added to this was 0.1 volume NaOAc [3mM, pH 5.2] and 2.5 volume (-20° C cold) absolute ethanol and it was left overnight at -20° C, then spun at 14,000 rpm for 30 mins at 4° C. The ethanol was poured off and the pellet washed by overlaying with a small volume of 80% (-20° C cold) absolute ethanol and re-spun at 14,000rpm for 15 mins at 4° C. The ethanol was poured off, the pellet air dried and then thoroughly resuspended in milliQ H_2O. Long term storage of the undiluted stock mtDNA sample was at -80° C. The working stock solutions for PCR, 1:10 and 1:50 dilutions were stored at -20° C.

CONCENTRATION OF THE EXTRACTED mtDNA

An aliquot of 3µl of the stock solution was mixed with 2µl of 10 x gel loading buffer and with 100ng/µl of a 1Kb-Plus DNA ladder was electrophoresed on a 1.5% (w/v) Agarose gel in 1 x TAE Buffer for 15 mins at 100 V, visualized using ethidium bromide staining for 10-15 mins and photographed using UV light. Both feather and extracted mtDNA gave a final concentration of ~ 1.2ng/µl.

PCR AMPLIFICATION

PCR is the exponential, enzymatic amplification of a specific region of DNA. The process is rapid and reliable. After an initial denaturation step to ensure complete DNA strand

separation there is a three step cycle: the template separation, the primer annealing and the sequence replication. Each new copy is now a template and by repeating the three step cycle many times, millions of copies can be produced simply and efficiently.

All PCR amplifications used aliquots of a 1:10 or 1:50 dilution of the purified mtDNA or gel extracted mtDNA. As insurance against possible human contamination some primers were kiwi / avian specific. Also, as a precaution, every PCR amplification run included a blank (no DNA) as a negative control. These served to check that all the reagents and the primers were clean and there was no DNA contamination.

A list of the primers and their sequences that were used for short-range PCR amplifications and the complete LSK mitochondrial genome sequence are in Appendix 1.

Long Range PCR

The Expand^{TM} Long Template PCR System (Boehringer Mannheim) was used for long-range PCR, and their System 3 was chosen. The manufacturer's protocol was followed, except the primer concentration was increased to 10µM. An MJ Research PTC-200 Peltier Thermal Cycler was used with the thermo-cycling program of: 92° C for 90secs, followed by 30 cycles of; 92° C for 40secs / 55° C for 60secs / 72° C for 3 mins + 15secs per cycle, then 4° C for an indefinite hold.

Annealing temperatures (T_m), were based on the averaged melting temperatures of the individual primer pairs and their homogeneity to the template. T_m values were based on the GC method; 2 (A+T) ° C + 4(G+C) ° C. Higher annealing temperatures allow for more specific amplifications, but the pure mtDNA used lower temperatures. The amplification was performed with both paired primers using a concentration of 1.5µM.

After the PCR amplification, 10% of the PCR products were run on a 0.5% (w/v) / 1 x TAE buffered agarose gel. This was run for an hour at 84V, stained in ethidium bromide, photographed and fragment size determined. The gel bands were then cut from the gel, purified and stored, incase they were needed as a template for shorter PCR amplifications.

The error rate for long-range PCR fragments (Taq + Pwo) is 13 x 10^{-6} and short-range PCR fragments (Taq only) is 26 x 10^{-6} (Roche, PCR Application Manual). To limit the error rate

each long-range PCR fragment was amplified using one round of short-range PCR to generate fragments for sequencing.

Short-range PCR

PCR reaction mixture

Qiagen PCR Kits (BioLab Scientific Limited) were used for all short PCR amplifications. Each PCR reaction mixture of 20µl contained: 1 **x** Qiagen PCR 10 x Reaction Buffer, 250 µM of each deoxy-nucleosidetriphosphate (dNTP, Boehringer Mannheim), 0.5 µM of each primer (forward and reverse), 1 µl of DNA extract (≤ 0.12ng template / reaction), 1 **x** Qiagen PCR Q-Solution, 1 unit of Qiagen *Taq* DNA polymerase and was made up to volume with milliQ H_2O. To prevent possible cross-contamination the DNA template was always added last. Usually this was either 1 µl of a 1:10 dilution of the standard stock (0.12ng / µl) or 1 µl of the gel extracted mtDNA. Q-Solution changes the melting behaviour of DNA and can be used for PCR systems which do not work well under standard conditions. The system was chosen as it was unknown if the primers used for moa sequencing (Cooper *et al.* 2001), for the short PCR amplifications, would amplify the LSK sequence.

PCR amplifications cycle

All short PCR amplifications were done in a MJ Research PTC-200 Peltier or a MJ Research PTC-150 Minicycler, both with heated lids to prevent sample evaporation. The reaction tubes were placed in the thermal cycler when the initial temperature was reached, using the thermo-cycling program of; 1 cycle; 94° C for 120secs, 35 cycles; 94° C for 60secs / 50° C for 60secs / 72° C for 60secs and finally for sequence extension, 5 mins at 72° C, then an indefinite hold at 4° C until prepared for sequencing. The annealing temperature was based on an average melting temperature of all the individual primer pairs used, as this would increase the primer homogeneity to the template mtDNA and decrease possible mispriming.

Agarose gel electrophoresis

Aliquots of all the samples: extracted DNA and mtDNA, PCR amplification products, extraction and negative controls were visually checked for product. Ten percent of the reaction volume was mixed with 1 - 2µl of 10 x gel loading buffer [27.5% (w/v) Ficoll Ty 400 (Pharmarcia), 0.44% (w/v) bromophenol blue (Serva) and 0.44% (w/v) xylene cyanol (Sigma)]. This mixture was then loaded onto a 1% (w/v) Seakem LE

agarose (FMC BioProducts) / 1 x TAE (Tris acetate EDTA) buffer [40mM Tris acetate, 1Mm EDTA, (pH 8.0)] gel with 1 x TAE running buffer. To maintain sample purity only every second lane was loaded.

Gibco BRL (Invitrogen) 1Kb and 1Kb-plus DNA ladders were routinely included with each gel run as a standard to check product size and concentration. The gel was electrophoresed at 100V in 1 x TAE running buffer for 45 - 60 minutes. DNA visualization was achieved by staining the gel in ethidium bromide for 10-15 mins and the resulting fluorescence viewed on a UV transilluminator (wavelength 302 nm, UVP Incorporated). To avoid possible contamination the gel was placed on a clean, clear plastic film before viewing. The gel was digitally photographed on the UV light using a video camera (Panasonic) and ImagePC software (Scion).

Gel-purified DNA

The resulting visualized gel band for each sample was excised with a sterile scalpel blade, transferred to a micro-centrifuge tube, weighed, labelled and stored at -20° C. A Qiagen Gel Extraction Kit (BioLab Scientific Limited) was used and the manufacturer's protocol in the Qiagen QIAquick^{TM} handbook: QIAquick Gel Extraction Kit Protocol was followed. The purified samples were stored at -20° C. An extra step of gel purification is required if more than one band is present in the gel lane. However, no non-specific primer binding occurred in my samples and only one band / sample was present. Gel-purified DNA can be used as a template to prepare clean PCR products for direct sequencing.

PCR template purification

All substrates (free dNTPs, polymerases and salts) were required to be removed from the PCR product to prevent interference with the sequencing reaction. A Qiagen PCR Template Purification Kit (BioLab Scientific Limited) was used and the manufacturer's protocol in the Qiagen QIAquick^{TM} handbook: QIAquick PCR Purification Kit Protocol was followed. The PCR product that has a high concentration of DNA can be sequenced directly after the primers are removed.

Quantifying the purified PCR

A DNA template concentration of 30-35ng/μl is required for automatic sequencing reactions. To quantify the concentration of the purified PCR product usually a 3μl aliquot was run on a 1% (w/v) Agarose TAE buffer gel in 1 x TAE running buffer with a 1μl Low Mass Ladder (LML) and a 1Kb-Plus DNA ladder. This was run for 45 mins at 100 V, the gel was stained using ethidium bromide and photographed. The intensity of the bands was compared to the specific length and concentration of the LML bands to give an estimated concentration.

<u>Method 3:</u> DNA EXTRACTION FROM FEATHERS

This method was trialled for three reasons. Firstly, because it is a non-invasive method and the collection, transportation and storage is efficient. The feathers from the kiwi Y2 were stored in a clean, plain white envelope at room temperature for three months before the DNA was extracted. Secondly, DNA has been used in the identification of the present day kiwi, both for classification and parentage, and can also be used in future research to look at the classification and diversity of extinct kiwi species. Thirdly, this was a control for the extracted pure mtDNA.

Four feathers from the LSK Y2 were placed in a drop of milliQ H_2O and the section just below the feathers was removed. Using the previous lysis buffer (Method 2) with 0.2 mg/ ml proteinase K, 1% SDS and 6.5μl of 1M DTT (dithiothreitol) in a final volume of 800μl, the feather tips were incubated at 51.6° C for 24 hours in a hybridization cabinet. The slow revolving cylinder rack ensured constant, gentle inversion of the tubes. These were then stored overnight at -4° C. The supernatant was removed for cleaning and the sediment was stored at -20° C for possible re-extraction. The previous phenol/chloroform extraction was followed using two phenol steps and one chloroform step. Inversions were for 5 mins with centrifuging at 8000 rpm for 6 mins for the first centrifuge and 8000 rpm for 5 mins for the following two. The supernatant was loaded onto a Centricon -30 disposable filtration column and washed 3 x with milliQ H_20 and centrifuged in a Heraeus at 5500 x g for 20 mins each time. The Centricon column was inverted, a small volume of milliQ H_2O was added and then it was centrifuged in a Heraeus until 2000 x g was reached. Long term storage of the undiluted stock mtDNA sample was at -80° C. The working stock solutions for PCR, 1:10 and 1:50 dilutions were stored at -20° C.

Short fragments were amplified using a PCR reaction mixture containing; 0.25µl of Ampli *Taq* Gold polymerase (Perkin Elmer), 1 x Gene Amp 10 x PCR Buffer II, 2.5µl 20mM Mg Cl_2 , 250µM of each dNTP, 2.5µl of each 10µM primer (forward and reverse), 5.0µl of 10 mg/ml Bovine Serum Albumen (BSA) and 1µl of extracted mtDNA. The samples were made up to a total volume of 25µl with milliQ H_2O. The MJ Research machine was used with the thermo-cycling program of: 1 cycle; 94° C for 2 mins, 35 cycles; 94° C for 40secs / 55° C for 60secs / 72° C for 60secs, and a final sequence extension of 4 mins at 72° C, then an indefinite hold at 4° C. The primers, F1753 (12SA) and R2150 (12SB5), give a fragment size of ~ 400bp. Primers F1278 (tPhe) and R2150 (12SB5) gave a fragment size ~ 872bp. The primers F2035 (12SC2) and R3420 (16S) did not amplify a sequence using the feather extracted DNA (see p2). Long term storage of the undiluted stock mtDNA sample was at -80° C. The working stock solutions for PCR, 1:10 and 1:50 dilutions were stored at -20° C.

A 10µl aliquot, mixed with 2µl of 10 x gel loading buffer, was run on a 1.5% (w/v) Agarose gel in 1 x TAE Buffer for 1hour at 100 V, visualized using ethidium bromide staining for 10-15 mins and photographed using UV light. The appropriate sized bands were cut from the gel, purified, and stored at 4° C.

The above method was also used to extract mtDNA from the isolated intact Y2 LSK mitochondria using a 15µl aliquot of the stock solution (Method 1). PCR was performed using the above primers and also the additional primers of F2035 (12SC2) and R3420 (16S). A fragment size of ~ 1385bp was obtained.

Sequencing

Cycle sequencing reactions were performed on both strands using an ABI BigDye PRISM kit, using 3´ fluorescent dye-labelled dideoxynucleotide triphosphate terminators and read on an ABI automated DNA sequencer, Model 377 (Applied Biosystems).

MITOCHONDRIAL ALIGNMENT USED FOR PRIMERS

The complete mitochondrial genome was isolated and the majority was amplified by long-range PCR into overlapping fragments of 1.0 – 4.2 kb in length. These fragments were templates for a series of shorter, overlapping PCR fragments. The original 12 primers for the long PCR were based on consensus sequences from the complete mtDNA alignments of the

ostrich (Härlid *et al.* 1997), the rhea (Härlid *et al.* 1998), the chicken (Desjardins and Morais, 1990) and a standard human sequence. Partial alignments from the moa, *Anomalopteryx didiformis* (Andi) and *Dinornis giganteus* (Digi), were from tRNA-Met to a partial ND2, continuing to a partial COII, then from CBI to Cytb. The 12S rRNA (to 1852), 16S rRNA, COI (to 7754), COII (from 8428), tRNA-Lys and Cytb of two tinamous, *Nothura maculosa* (Noth) and *Tinamus major* (TinM), cassowary, *Casuarius bennetti* (CasB), *Dromaius novaehollandiae* (Emu) and a kiwi (*Apteryx australis*) were also included.

The manual alignment and the short primers used, usually 20bp, were from Alan Cooper's mtDNA sequencing of the two moa genera, *Emeus crassus* and *Dinornis giganteus* (Cooper *et al.* 2001).

PRIMERS

Primer Design

Primers were designed from the conserved regions of the mitochondrial alignment (see p. 28). Longer PCR products require longer primers (20 - 30 nucleotides) and higher annealing temperatures (T_m). This ensures binding and amplification of a specific mtDNA sequence and reduces the chances of amplifying nuclear copies or non-target sequences such as other mammalian and avian mitochondrial genomes that were being sequenced in the same laboratory. Using long fragments of mtDNA also avoided possible mispriming of the primers with similar sequences on the same genome. Degenerate sites were incorporated in some primers as required. A complete data set for all primers and primer pairs were obtained using Amplify, Version 1.2 (© 1992, 1993 Bill Engels).

Annealing Temperature (T_m)

The primer lengths were 20 - 28 bases. The ideal T_m range is 55° C – 70° C. When using the estimated T_m formula: 4 (G+C) °C +2 (A+T) °C, it is recommended to use an annealing temperature 5 °C below T_m.

Duplex formation, hairpins and upper-lower primer-dimers

These structures will decrease the amount of free primers available to bind to the specific sites. The rule is to prevent any secondary structure which has negative free energy and prevent 3′ end mispriming more than 2 – 4 base pairs.

Degenerate sites

A degenerate site is where a primer has more than one base incorporated into a particular site. This is useful if the aligned consensus sequence is ambiguous at a base, as this will increase the chance of a primer annealing to the specific sequence. Degenerate sites are limited to 2 or 3 in a 20 nucleotide primer, as they can decrease the primers T_m and specificity. Nucleotide T can be used at an ambiguous base site instead of a degenerate site as it can bind with any other base without significant effect (Kwok *et al.* 1990).

Degenerate sites:

Code		Nucleotide	Code		Nucleotide
Y	=	C + T	V	=	A + C + G
W	=	T + A	B	=	T + C + G
R	=	A + G	D	=	A + T + G
M	=	A + C	H	=	A + T + C
K	=	T + G	N	=	A + T + C + G
S	=	C + G			

Optimal Annealing Temperature

The annealing temperature difference between the upper and lower primers should be less than 3° C to optimize the PCR reaction.

The perfect primer has a G+C content > 50%; to increase the stability (backbone), increase the primer-template binding strength, and help increase T_m. Whereever possible use G or C at 3' end of primer as this helps annealing.

LONG-RANGE PCR PRIMERS

The long primers were used to divide the mitochondrial genome into six long overlapping PCR fragments. The 12 primers required were ordered from Gibco BRL, Custom Primers from Life Technologies, New Zealand. (http://order.lifetech.com/lti_store/index.icl).

The L and H refer to the light and heavy strands respectively, and the primer numbering refers to the nucleotide position of the 3' base of the primer in the published chicken mitochondrial sequence (Desjardins and Morais, 1990). The L-strand is the forward primer and the H-strand is the reverse primer.

The 12 long PCR primers and the sequences:

L1278 5' GGC ACT GAA GAT GCC AAG ATG GTA 3'

H4015 5' GGA GAG GAT TTG AAC CTC TG 3'

L3450	5' GAA GAC CCT GTG GAA CTT GAA 3'
H7662	5' AGG AAG ATG AAG CCY AGA GCT CA 3'
L7318	5' ACA TTC TTT GAY CCW GCR GGA GG 3'
H10884	5' GGG TCR AAG CCR CAT TCG TAG GG 3'
L10647	5' TTT GAA GCA GCA GCC TGA TAY TG 3'
H13734	5' AGG CCA AAT TGR GCT GAT TTT CC 3'
L12976	5' CAA GAA CTG CTA ATT CCT GCA TCT G 3'
H16065	5' GYG RTC TTC YGT CTT TGG TTT ACA AGA C 3'
L15656	5' AAC CTG TTA GGR GAY CCA GA 3'
H1861	5' TCG ATT ATA GAA CAG GCT CCT C 3'

Primers: H4015, L3450, H7662, L7318, H16065, L15656, and H1861 had been used by Lee *et al.* (1997) to obtain mtDNA sequences from the ratites, including a brown kiwi *Apteryx australis mantelli.* Some of the above primers were modified for sequencing the LSK in this project.

Division of the LSK mitochondrial genome into six overlapping fragments

FRAGMENT 1: From t-Phe – t-Leu (~ 2737bp) used primers L1278 and H4015.
FRAGMENT 2: From 16SrRNA – COI (~ 4212bp) used primers L3450 and H7662.
FRAGMENT 3: From COI – ND3 (~ 3566) used primers L7318 and H10884.
FRAGMENT 4: From COIII – ND5 (~ 3087bp) used primers L10647 and H13734.
FRAGMENT 5: From t-Ser (L) – t- Thr (~ 3089bp) used primers L12976 and H16065.
FRAGMENT 6: From Cytb – 12S rRNA (D-Loop) used primers L15656 and H1861.

Fragments 1 and 6 of the genome did not amplify. DNA extracted from feathers and the extracted mtDNA was used to obtain Fragment 1 and the extracted mtDNA was later used to obtain Fragment 6.

Combinations of the primers used for PCR amplification

For each of the long-range fragments 1-5, the primers used are given, and then the short-range fragments are listed for each long-range fragment. The sequences for the short-range primers are in Appendix 1.

FRAGMENT 1: From t-Phe – t-Leu (~ 2737bp). Extracted mtDNA and DNA extracted from feathers were used to obtain this fragment.

1278F	-	1805R	=	527bp	
1753F	-	2150R	=	397bp	
2035F	-	2577R	=	542bp	
2471F	-	2919R	=	448bp	
2907F	-	3420R	=	513bp	2142bp
1278F	-	2150R	=	872bp (feathers and mtDNA)	
1753F	-	2150R	=	397bp (feathers and mtDNA)	
2035F	-	3420R	=	1385bp (mtDNA extracted using Method 3)	

FRAGMENT 2: From 16SrRNA – COI (~ 4212bp) used primers L3450 and H7662.

3450F	-	3797R	=	347bp	
3787F	-	4311R	=	524bp	
3787F	-	4791R	=	1004bp	
4747F	-	5201R	=	454bp	
5197F	-	5583R	=	386bp	
5574F	-	6120R	=	546bp	
5976F	-	6367R	=	391bp	
6335F	-	6642R	=	307bp	
6634F	-	7062R	=	428bp	
7056F	-	7553R	=	497bp	
7459F	-	7662R	=	203bp	4212bp

FRAGMENT 3: From COI – ND3 (~ 3566) used primers L7318 and H10884.

7318F	-	7553R	=	235bp	
7546F	-	7816R	=	270bp	
7807F	-	8325R	=	518bp	
8320F	-	8807R	=	487bp	
8320F	-	9284R	=	964bp	
9246F	-	9755R	=	509bp	
9246F	-	10229R	=	983bp	
9712F	-	10229R	=	517bp	
9712F	-	10743R	=	1031bp	
9874F	-	10229R	=	355bp	
10716F	-	10884R	=	168bp	3566bp

FRAGMENT 4: From COIII – ND5 (~ 3087bp) used primers L10647 and H13734.

10647F	-	11152R	=	505bp	
11904F	-	12800R	=	896bp	
12312F	-	12800R	=	488bp	
12788F	-	13200R	=	412bp	
13052F	-	13592R	=	540bp	
13562F	-	13734R	=	171bp	3087bp

FRAGMENT 5: From t-Ser (L) – t- Thr (~ 3089bp) used primers L12976 and H16065.

12976F -	13592R	=	616bp	
14002F -	14539R	=	537bp	
14002F -	14946R	=	944bp	
14452F -	14946R	=	494bp	
14452F -	15345R	=	893bp	
15212F -	15649R	=	437bp	
15303F -	16065R	=	762bp	
15671F -	16065R	=	394bp	3089bp

FRAGMENT 6: From Cytb – 12S rRNA (D-Loop).

This long fragment did not PCR. The D-loop and remaining gaps in the genome were later sequenced by Gillian Gibb using standard primers and a few designed ones. Standard protocols were used and the fragments were analysed on an Applied Biosystems DNA sequencer Model ABI 3730 (see Phillips *et al.* 2010 in Appendix 2).

The final alignment of all DNA sequences is shown in Figure 3, p 43.

Sequence conformation and Alignment

PCR sequences were confirmed through a BLAST search (The National Centre for Biotechnology Information database) which aligns a sequence with the closest matching species. These sequences were also checked visually against known kiwi sequences available at the time in the mitochondrial alignment.

Both the H and L strand sequences were aligned, manually edited and all strands assembled in Sequencher 4.1 (Gene Codes Corporation®., Ann. Arbour, MI 48108). The L-strand sequence is numbered from tRNA-Phe. All protein-coding regions are transcribed from the H-strand except NADH6, which is coded on the L-strand. The alignment and a list of all the primers used in this thesis are tabled in Appendix 1 of this thesis, and in Supplementary Information (http://sysbio.oxfordjournals.org/cgi/content/full/59/1/90, Phillips *et al.* 2010).

Phylogenetic Analysis

See Appendix 2.

Chapter 3 - **RESULTS**

Genome length and control region

The complete mitochondrial genome of the LSK (*Apteryx owenii*) has been sequenced and submitted to GenBank; acquisition number GU071052. The genome is 17,020bp long, though the length may not be absolute due to heteroplasmy in the control region. This is due to the occurrence of different numbers of a repeated motif, of different lengths, in different mitochondrial genomes/organelles of the same individual, as well as between individuals. Heteroplasmy is also found in the rhea but was not found in the chicken and ostrich (Härlid *et al.* 1998). The total length of the genome, minus intergenic spacers, is 16,828bp, similar in size to the other sequenced ratites and tinamous, and the control region is ~ 1169bp long (Table 2, p46). The control region is a non-coding sequence characterized by the occurrence of the repetitive motifs (TTTT...A). These are not straight repeats in that the number of T nucleotides varies, but they all end with an A nucleotide. The multiple repeats occur one after the other and appear as a block from around 16644bp – 16747bp, at the 3' end of the control region (see Appendix 1). This 104bp sequence is larger than the similar (TTTT…A) 90bp sequence of the GSK found between 16614bp – 16703bp (Haddrath and Baker, 2001).

One conserved sequence block (CSB) found close to the 5′ end of the control region (15696bp – 15712bp) of the LSK contained a sequence similar to the CSB-II of the ostrich (15589bp – 15605bp) and rhea (15549bp – 15565bp). All three conserved structural blocks, F, D and C boxes, were present in domain II and control region sequence data has been applied to population structure and systematics of birds (Baker and Marshall, 1997).

Genome organization

The mitochondrial genome codes for 22 tRNAs, 13 protein-coding genes and 2 rRNAs. The H-strand is a guanine rich strand encoding 14 tRNAs, 12 protein-coding genes and 2 rRNAs. The L-strand is a cytosine rich strand encoding 8 tRNAs and 1 protein-coding gene and is indicated by (L). Birds have a different mitochondrial gene order compared with other vertebrates. The LSK has the same avian gene order first reported for the chicken mitochondrial genome (Desjardins and Morais, 1990). The organization of the LSK genome is shown in Fig. 4 and this arrangement is currently the standard gene order for ratites (except for the elegant-crested tinamou). It is also the only known arrangement existing for galliformes and anseriformes (ducks and geese) and these two groups join on a common

branch of the avian evolutionary tree to form the group galloanseres (Slack *et al.* 2003). van Tuinen *et al.* (2000) phylogenetic analyses indicated that galloanseres, which are adjacent to paleognaths, form a basal lineage of neognathous birds. This finding supports the traditional rooting position on the avian tree to be between paleognaths and neognaths.

The standard ratite gene arrangement of [NADH5/Cytb/ tRNA-Thr/tRNA-Pro (L)/**NADH6 (L)/tRNA-Glu (L)**/control region] is only one rearrangement (translocation) away from the presumably ancestral gene order found in many reptiles [NADH5/**NADH6 (L)/tRNA-Glu (L)/**Cytb/ tRNA-Thr/tRNA-Pro (L)/control region] (Desjardins and Morais, 1990). NADH5 and NADH6 refer to nicotinamide adenine dinucleotide dehydrogenase, subunits 5 and 6. One exception to the paleognath mitochondrial gene arrangement is the elegant-crested tinamou, where NADH6 had moved from in front of the control region to behind it [NADH5/Cytb/ tRNA-Thr/tRNA-Pro (L)/control region/**NADH6 (L)/tRNA-Glu (L)**] (Haddrath and Baker, 2001). This appears to be a unique, independent translocation of NADH6 (L) within the tinamou and this has been found to have occurred in other avian taxa (Mindell *et al.* 1998a). The rearrangement would also suggest the tinamou is not basal in the evolutionary tree.

The double stranded, circular mitochondrial genome operates a physical/mechanistic constraint on gene order changes and these changes are thought to occur due to replication error near origin of replication sites (Mindell *et al.* 1998a), followed by possible gene loss or reduction of any duplicated genes. There are at least four distinct avian mitochondrial gene orders and the area between duplicate control regions and gene order changes have been useful markers for phylogenic resolution as they are stable and can retain information over millions of years (Gibb *et al.* 2007). The identification of a repeat 1 (CR 1) retrotransposon insert in the control region of the chicken, a rare DNA change, has also been used to determine relationships among closely related birds (see Gibb *et al.* 2007).

Origins of replication

The H-strand origin of mitochondrial replication (ori-H) is within the control region and replication is thought to involve the CSB(s). Parallel gene order evolution among non-monophyletic birds involving genes in close proximity to the control region suggests a constraint in gene rearrangement at the ori-H site (Mindell *et al.* 1998a).

In most vertebrate mtDNA, including tetrapods, the L-strand origin of replication (ori-L) is a stable hairpin structure found between tRNA-Asn (L) and tRNA-Cys (L) genes located within the WANCY region, a cluster of 5 tRNAs. There is an absence of an ori-L in the LSK. This appears to be a common characteristic of both birds and crocodilians, and these lineages diverged from each other ~ 254Mya (Janke and Arnason, 1997). In the chicken the tRNA-Asn (L) and tRNA-Cys (L) genes are separated by one nucleotide and in the LSK there is a gap of only two nucleotides.

In the chicken both Ori-L and Ori-H, separated by only 1bp, are located in the control region close to the tRNA-Phe gene. Transcription of both mtDNA strands is initiated from one major bidirectional promoter (L'Abbe *et al.* 1991). The same process(s) could be operating in the ratites.

Contiguous genes

Eleven genes are butt-joined: tRNA-Glu (L) / control region/tRNA-Phe / 12S rRNA / tRNA-Val / 16S rRNA / t RNA- Leu; tRNA-Met / NADH2; tRNA-Gly / NADH3; tRNA-Arg / NADH4L; tRNA-His / tRNA-Ser (AGY) and tRNA-Leu (CUN) / NADH5.

Intergenic spacers

The mitochondrial genome of the LSK is arranged in a very compact manner. Most of the non-coding regions occur between tRNA genes and protein-coding genes, ranging from 0 – 30bp (Table 2, p. 46). The 9bp intergenic spacer sequence found between tRNA-Pro (L) / NADH6 (L) genes in the chicken (Desjardins and Morais, 1990) is smaller than the ratite mitochondrial intragenic region. The tRNA-Pro (L) / NADH6 (L) intragenic regions of the ostrich, moa, tinamou and rhea were all similar (12-18bp) with the kiwi, emu and cassowary having the largest insertions, ranging from 23-30bp. The 30bp insert region of the LSK is strongly biased toward G and against C in the H-strand sequence: G = 53%: C = 7%; T = 20%; A = 20%. In the brown kiwi (*Apteryx australis*) there is an extra 3bp insertion relative to the 30bp of the LSK and GSK and it is thought that this region may still be expanding in the kiwi species (Cooper, 1997). Both tRNA-Pro (L) and NADH6 (L) are encoded on the L-strand.

The non-coding regions between tRNA genes vary greatly in the ratites. The tRNA-Thr / tRNA-Pro (L) intragenic regions are similar in the elegant-crested tinamou and rhea (5bp;

9bp), the moa (26-27bp), the ostrich (55bp), the tinamou, GSK, emu and cassowary (82-83bp), but the LSK is the largest with 111bp. This is 28bp larger than the LSK's sister group, the GSK, and also appears to be expanding in the kiwi species.

The genes having the largest intergenic spacers are usually related to gene rearrangement (translocation) regions in proximity of the control region. The ancestral gene order has been maintained in the paleognaths, but the non-coding tRNA-Thr/tRNA-Pro (L)/NADH6 intragenic region appears to be increasing (Table 2, p. 46).

If changes in the gene order are a result of duplication followed over time by deletion / reduction events, the large non-coding sequences not yet eliminated would indicate recent changes in gene order (Mindell *et al.* 1998a; Gibb *et al.* 2007). The present gene order has been conserved among the ratites sequenced to date. The Haddrath and Baker (2001) genetic analyses placed the moa as the basal ratite. Slack *et al.* (2003) using various combinations of 19 avian taxa and six reptilian outgroups found the positions of the kiwi, moa and ostrich were not completely resolved. An ((emu, cassowary), (rhea)) kiwi grouping occurred and in the unrooted tree there was an ostrich / kiwi interchange. This interchange was also found to occur by the quartet method of Cooper and Penny (1997).

Overlapping regions

There are three 1bp overlaps involving tRNAs. The overlap between tRNA-Cys (L) / tRNA-Tyr (L) is on the L-strand and the overlap between tRNA-Ser (AGY) / tRNA-Leu (CUN) is on the H-strand. Endonucleolytic cleavage by a single precursor molecule cannot produce two mature tRNAs. It is thought the mature form of the tRNAs may be completed by a post-transcriptional addition of a single residue at either the 3′ end of one tRNA or the 5′ end of the other tRNA. L'Abbe *et al.* (1990) found, in the chicken mitochondria, a post-transcriptional addition of an extra guanylate residue at the 5′ end of tRNA-His occurred independently of the 3′ end maturation. The extra G base at the 5′ terminus and the added CCA sequence at the 3′ terminus are not encoded in the chicken tRNA-His gene. L'Abbe *et al.* (1990) suggested that the same process(s) that are operating in the addition of the CCA terminus to the 3′ end of tRNA-His could be involved in the addition of the extra G to the 3′ terminus to produce the mature tRNA.

The third 1bp overlap occurs between tRNA-Gln (L) / tRNA-Met and there is one protein-coding / tRNA overlap of 6bp between COI / tRNA-Ser (UCN) (L). These two overlaps, with genes encoded on opposite strands (H and L), are not true overlaps as different RNA transcripts are involved.

In two cases the protein-coding genes overlapped in different reading frames:
The first example is ATPase 8 / ATPase 6: a 7bp overlap (ATP synthase subunits 6 and 8).

```
 ATPase 8 (7835… 8002) and ATPase 6 (7993…8676)
                w   p  w   t   *
7981 atgatcctga ccatgaacct aagcttcttc gaccaatttg ccagccccca actactagga
                     m   n    l    s….
```

The second example is NADH4L / NADH4: a 4bp overlap.
NADH4L (9950…10246) and NADH4 (10240…11607)

```
                                                  l   q    c    *
10201 acatggctcc gatcacctac ataacctaaa cctcctacaa tgctaaaaat tattcttccc
                                                      m    l   k    i….
```

The overlapping regions and base pair number are highly conserved among the paleognaths sequenced to date (Table 2, p. 46).

Stop codons

The stop codon for the mtDNA protein-coding genes COII and COIII is a T (Table 1, p45). The stop codon is completed by the post-transcriptional polyadenylation addition of two 3' A residues to the primary transcript mRNA to give TAA (Ojala *et al.* 1981). This is consistent with other Paleognaths (Table 2, p. 46).

```
   COII (7074…7761)    I    tRNA-Lys (7762…7833)
7741 cattactttc cacatcctca tcattaagaa gctatgataa cagcactagc cttttaagct
                          (AA added after t)

                              COIII (8676…9459)    I    tRNA-Gly (9460…9528)
          9421 gattattcct ctacataacc atctactgat gaggatcatg ctcttctagt atatcaatta
                                                         (AA added after t)
```

An alternative stop codon is TAG and the 3' residues AG can constitute a part of the adjacent tRNA gene as in NADH2 and tRNA-Trp.

```
                         NADH2 (3986…5024)        I        tRNA-Trp (5025…5101)
  4981 cattactctc ctacctctca cacccataat atttgctacc ctctagaaac ttaggataac
```

(The underlined ag is part of the stop codon, and the start of the next tRNA)

The ATPase6 stop codon TAA is completed with the 3' residue A that constitutes the start of the codon ATG (Met) of the protein-coding gene COIII.

ATPase 6 (7993…8676) I COIII (8676…9459)
8641 ctcttattaa gcttatactt acaagagaat atttaatgac acaccaagca cattcctacc
(The underlined a is part of the stop codon, and the start of the next protein-coding codon)

It is also possible an endonucleolytic cleavage after the T nucleotide, position 5024 in NADH2 and the A nucleotide, position 8676 in ATPase6 / COIII (highlighted above), produces an incomplete stop codon (Ojala *et al.* 1981) and the stop codons do not overlap. The same mechanism can not apply for the 1 nucleotide overlap involving the stop codon AGG, between NADH1 and tRNA-Ile.

Protein-coding genes

The length of protein-coding genes: NADH2 (346 amino acids), COI (516 aa), ATPase8 (55 aa), ATPase6 (227 aa), COII (261 aa), NADH3 (116 aa) and NADH4L (98 aa) are conserved among the sequenced paleognaths. The difference in the lengths of the remaining protein-coding genes is only 1 – 3 aa.

The start codon for COI in the LSK is Met (ATA) and this differs from the other paleognaths that start with Val (GTG). The length of NADH4 (455 aa) and stop codon (AGA) in both the LSK and GSK is different to the other paleognaths that have a conserved length of 458 aa and a stop codon of TAA. The remaining protein-coding genes have a Met (ATG / ATA) start codon. The protein genes were translated by the same variant genetic code as the chicken.

Most of the protein-coding genes and both rRNA genes are separated by one or more tRNA genes. This separation may act as a recognition signal(s) for mitochondrial RNA processing, where the secondary structure of the primary transcript is the recognition signal rather than the nucleotide sequence (Ojala *et al.* 1981). Protein genes not separated by a tRNA are: ATPase8 / ATPase6 that overlap in different reading frames and is followed directly by Cytb, with only a 2 nucleotide gap, and NADP4 / NADP4L that also overlap in different reading frames (see p. 38).

tRNA and rRNA genes

The length of tRNAs (see Table 3, p47): Leu (URR; 74bp), Gln (L; 71bp), Ala (L; 69bp), Asn (L; 73bp), Cys (L; 67), Gly (69), His (70bp) and Ser (AGY; 66bp) are conserved among the paleognathes sequenced to date. The length of the remaining tRNAs and also the two rRNAs in the paleognaths show small variations and are similar to that of the chicken.

NADH3 and NADH6 (L)

Phylogenetic analyses (Appendix 2) are based on concatenated sequences using only 11 of the protein-coding genes. NADH3 and NADH6 (L) are excluded for the following reasons. The length of NADH3 in the ratites appears to deviate significantly from the length of the same gene in other vertebrates. This is due to an AGA termination codon after position 207 nucleotide, (sequence number 9732). RNA-editing is thought to correct the sequence before translation, either by deleting position 174 nucleotide (sequence number 9702 'C') or by a translational frameshift, producing a change in the third codon position (Mindell *et al.* 1998b). Without this alteration the frame shift generates a stop codon, AGA, at position 207 nucleotide (sequence numbers 9733-9735) and this produces a truncated protein product. Both the amino acid sequence after the correction and the final length of the protein is conserved among the paleognaths. The position 174 nucleotide 'C' is present in all the ratites sequenced to date (Härlid *et al.* 1997, 1998; Cooper *et al.* 2001; Haddrath and Baker, 2001) and appears to have been maintained instead of being selected against.

NADH3 (9529…9701, 9703…9880) tRNA-Arg (9882…9949)

9661 ctaggatcag cccgactccc cttctcaatc cgattcttcc tcagtagcca tcctattcct

9721 actcttcgac ttagaaatcg ctcttcttct ccccctcccc tgagcaatcc aactcccaca

9841 ttatgaatga gcacagggag gcctagaatg ggcagagtaa

NADH6 (L) is encoded on the L-strand, the opposite strand to the other mtDNA protein-coding genes and differs distinctly in nucleotide and amino acid composition and is therefore excluded from all analyses.

A revised annotation of paleognathous mitochondrial genomes

Slack *et al.* (2003) have re-examined and revised all published avian and reptilian mitochondrial genome annotation to obtain a consistent labelling for gene start and stop positions. A summary of the complete mitochondrial features are presented in table form and

are being maintained and updated. The three tables are available as Supplementary Information at http://awcmee.massey.ac.nz/downloads.htm (Slack *et al.* 2003). The changes will eliminate discrepancies and improve the future accuracy of mitochondrial sequencing, annotation and alignment. Using the new annotation the LSK data has been added to the three summary tables and presented here together with a partial data set (Slack *et al.* 2003), containing published paleognaths and a chicken, as an out group/comparison.

Table 1 Protein-coding genes (p. 45). Start and stop codons, and gene length (aa). The most likely start and stop positions were defined by taxa that had only one viable start or stop codon or by following previously accepted annotation. The stop codons have been included in the length of the intergenic spacers and this clearly identified overlaps between coding sequences (Table 2). The extra nucleotide found in NADH3 is not included. It is assumed there are no intergenic spacers or overlaps between the two rRNAs, the control region and the flanking tRNAs in the region from tRNA-Glu to tRNA-Leu (UUR).

Table 2 Control regions, intergenic spacers and total genomes (p. 46). Length (in nucleotides). The control region and rRNA genes are defined under the assumption that there are no intergenic spacers or overlaps between these regions and the flanking tRNAs. Stop codons of preceding protein-coding genes have been included in the intergenic spacers. There are no overlaps between tRNAs that precede protein-coding genes. The total length of the genome = ∑ protein-coding gene lengths x three + ∑ RNA lengths + control region + ∑ intergenic spacers – ∑ overlaps + 1 nucleotide (only for taxa that have an extra nucleotide in NADH3).

Table 3 tRNA and rRNA genes (p. 47). Length (in nucleotides).
The tRNA start and stop positions were identified using both secondary structure and multiple sequence alignments. Amino acid abbreviation: Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.

Colour coding for mitochondrial summary Tables 1, 2 & 3.

Light gray shading: This feature is conserved amongst paleognaths.

Dark purple shading: Changes to a feature conserved in most of the taxa are highlighted in the tables and the range columns.

Abbreviations used in Tables 1-3.

T-- and TA-: Incomplete stop codons; >: Incomplete region/genome, full length unknown; N/A: Not applicable to the taxa; –: Zero (no spacer); Overlap: The number of nucleotides that the two genes overlap by (stop codons are not invoved in the overlaps); Unknown: The mitochondrial genome of the taxa is incomplete and does not include this gene/region. R and Y in the range column: R (purine, A or G), Y (pyrimidine, C or T).

Avian taxa used in the table data set and Genbank accession numbers

Emu (*Dromaius novaehollandiae*; AF338711), double-wattled cassowary (*Casuarius casuarius*; AF338713), little spotted kiwi (*Apteryx owenii*; GU071052), great spotted kiwi, (*Apteryx haastii*; AF338708), greater rhea 1 (*Rhea americana*; Y16884), greater rhea 2 (*Rhea americana*; AF090339), lesser rhea (*Pterocnemia pennata*; AF338709), ostrich 1 (*Struthio camelus*; Y12025), ostrich 2 (*Struthio camelus*; AF338715), eastern moa 1 (*Emeus crassus*; AY016015), eastern moa 2 (*Emeus crassus*; AF338712), little bush moa (*Anomalopteryx didiformes;* AF338714), giant moa (*Dinornis giganteus*; AY016013), great tinamou (*Tinamus major*; AF338707), elegant-crested tinamou (*Eudromia elegans*; AF338710) and chicken (*Gallus gallus*; X52392).

Avian taxa references

Emu, double-wattled cassowary, great spotted kiwi, lesser rhea, ostrich 2, eastern moa 2, little bush moa, great tinamou and the elegant-crested tinamou (Haddrath and Baker, 2001); greater rhea 1 (Härlid *et al.* 1998); greater rhea 2 (Mindell *et al.* 1999); ostrich 1 (Härlid *et al.* 1997); eastern moa 1 and giant moa (Cooper *et al.* 2001) and chicken (Desjardin and Morais, 1990).

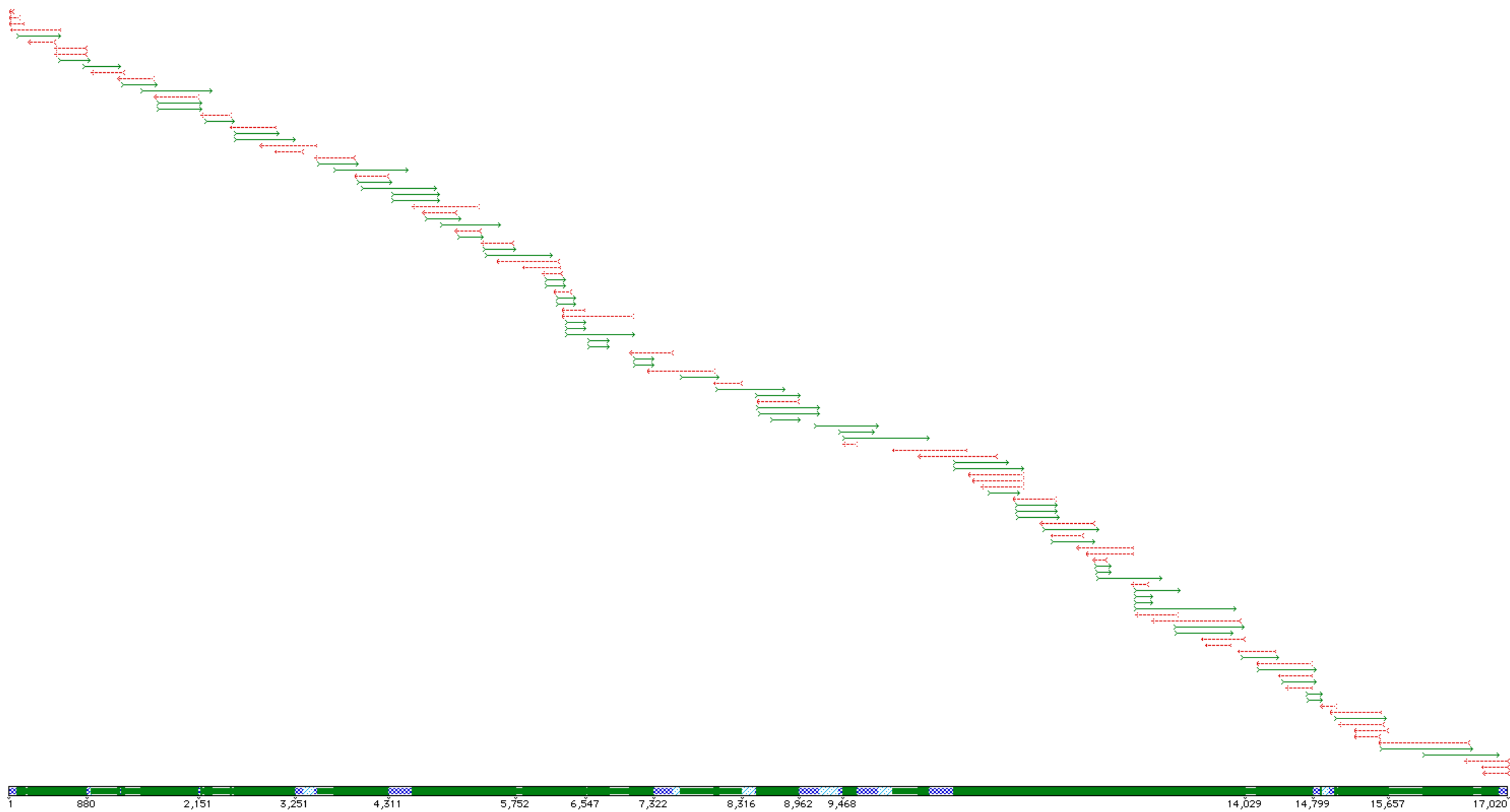

Figure 3: Sequencer output of overlapping DNA sequences for the LSK mitochondrial genome.

Apteryx owenii mitochondrion, complete genome

gi|288900678|ref|NC_013806.1|

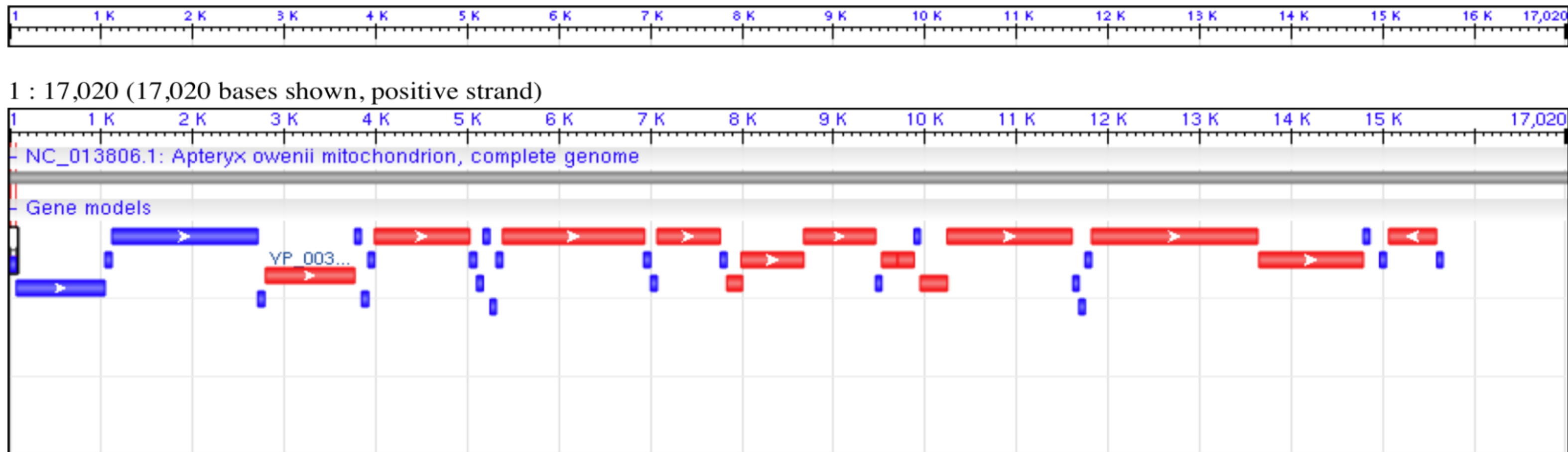

Figure 4: Graphical representation of the LSK mitochondrial genome

Table 1. Length (in amino acids). Start and Stop Codons of Paleognath Mitochondrial Protein-Coding Genes.

Genes (13)		Emu	Cassowary	LSKiwi	GSKiwi	Greater Rhea1	Greater Rhea2	Lesser Rhea	Ostrich1	Ostrich2	Eastern Moa1	Eastern Moa2	LBMoa	Giant Moa	Tinamou	CrTinamou	Range (paleognaths)	Chicken
NADH1	Length	323	323	324	324	324	324	324	324	324	323	323	323	323	322	322	322 - 324	324
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAA	AGG	AGG	AGG	AGG	AGG	AGG	AGG	TAA	TAA	TAA	TAA	AGA	AGA	AGR, TAA	TAA
NADH2	Length	346	346	346	346	346	346	346	346	346	346	346	346	346	346	346	346	346
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATA (Met)	ATG (Met); ATA (Met)	ATG (Met)
	Stop codon	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG
COI	Length	516	516	516	516	516	516	516	516	516	516	516	516	516	516	516	516	515
	Start codon	GTG (Val)	GTG (Val)	ATG(Met)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val); ATG (Met)	GTG (Val)
	Stop codon	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG	AGG
COII	Length	229	229	229	229	229	229	229	229	229	228	228	228	228	228	228	228; 229	227
	Start codon	GTG (Val)	GTG (Val)	GTG(Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	GTG (Val)	ATG (Met)
	Stop codon	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	TAA
ATPase8	Length	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	54
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA
ATPase6	Length	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227	227
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA
COIII	Length	261	261	261	261	261	261	261	261	261	261	261	261	261	261	261	261	261
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--	T--
NADH3	Length	116	116	116	116	116	116	116	116	116	116	116	116	116	116	116	116	116
	Start codon	ATA (Met)	ATA (Met)	ATA(Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATA (Met)	ATG (Met)
	Stop codon	TAA	TAA	TAA	TAA	TAG	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA; TAG	TAA
NADH4L	Length	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA
NADH4	Length	458	458	455	455	458	458	458	458	458	458	458	458	458	458	458	455; 458	459
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAA	AGA	AGA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA; AGA	T--
NADH5	Length	605	605	604	604	605	605	605	605	605	605	605	605	605	603	606	603 - 606	605
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	GTG (Val)	GTG (Val)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATA (Met)	ATG (Met)	ATG (Met); GTG, ATA	ATG (Met)
	Stop codon	AGA	AGA	TAA	TAA	TAA	TAA	TAA	AGA	AGA	AGA	AGA	AGA	AGA	TAA	TAA	TAA, AGA	TAA
Cyt*b*	Length	379	379	379	379	379	379	379	379	379	380	380	380	380	379	379	379; 380	380
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)
	Stop codon	TAA	TAG	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA	TAA; TAG	TAA
NADH6 *(L)*	Length	173	173	173	173	174	174	174	173	173	173	173	173	173	173	unknown	173; 174	173
	Start codon	ATG (Met)	ATG (Met)	ATG(Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	ATG (Met)	unknown	ATG (Met)	ATG (Met)
	Stop codon	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	TAG	unknown	TAG	TAA

Table 2. Length (in nucleotides) of Paleognath Mitochondrial Control Regions, Intergenic Spacers and Complete Genomes

	Emu	Cassowary	LSKiwi	GSKiwi	Greater Rhea1	Greater Rhea2	Lesser Rhea	Ostrich1	Ostrich2	Eastern Moa1	Eastern Moa2	LBMoa	Giant Moa	Tinamou	CrTinamou	Range (paleognaths)	Chicken
Control Region (CR)	1093	>1136	1169	1187	1171	1161	1202	1031	1034	1501	>1100	>1150	1508	1102	>383	1031 - 1508	1227
Additional Non-Coding Region	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	unknown	N/A	N/A
tRNA-Phe/12S rRNA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
12S rRNA/tRNA-Val	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
tRNA-Val/16S rRNA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
16S rRNA/tRNA-Leu (UUR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
tRNA-Leu (UUR)/NADH1	8	10	9	9	11	11	10	9	9	11	11	11	11	9	9	8 -11	9
NADH1/tRNA-Ile	3	3	1	1	11	11	11	1	1	4	4	4	4	19	21	1 - 21	3
tRNA-Ile/tRNA-Gln *(L)*	6	5	9	9	14	14	13	11	11	7	7	8	8	7	7	5 - 14	5
tRNA-Gln *(L)* /tRNA-Met	1 overlap	1 overlap	1overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap
tRNA-Met/NADH2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
NADH2/tRNA-Trp	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
tRNA-Trp/tRNA-Ala *(L)*	1	1	1	1	1	1	1	-	-	1	1	1	1	1	1	0; 1	6
tRNA-Ala *(L)* /tRNA-Asn *(L)*	3	3	1	1	2	2	2	4	4	1	1	1	1	1	-	0 - 4	3
tRNA-Asn *(L)* /tRNA-Cys *(L)*	3	3	2	2	3	3	3	1	1	2	2	2	2	2	2	1 - 3	1
tRNA-Cys *(L)* /tRNA-Tyr *(L)*	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap
tRNA-Tyr *(L)* /COI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
COI/tRNA-Ser (UCN) *(L)*	6 overlap	6 overlap	6overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap	6 overlap
tRNA-Ser (UCN) *(L)* /tRNA-Asp	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2; 3	2
tRNA-Asp/COII	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
COII/tRNA-Lys	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4
tRNA-Lys/ATPase8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ATPase8/ATPase6	7 overlap	7 overlap	7overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap	7 overlap
ATPase6/COIII	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
COIII/tRNA-Gly	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
tRNA-Gly/NADH3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
NADH3/tRNA-Arg	4	4	4	4	5	5	5	4	4	4	4	4	4	3	4	4; 3, 5	4
tRNA-Arg/NADH4L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
NADH4L/NADH4	4 overlap	4 overlap	4overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap	4 overlap
NADH4/tRNA-His	4	4	10	10	4	4	4	4	4	4	4	4	4	5	5	4; 5, 10	1
tRNA-His/tRNA-Ser (AGY)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
tRNA-Ser (AGY)/tRNA-Leu (CUN)	1 overlap	1 overlap	1overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	1 overlap	-
tRNA-Leu (CUN)/NADH5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-
NADH5/Cytb	12	12	15	15	15	15	15	12	12	13	13	13	13	2	2	2 - 15	7
Cytb /tRNA-Thr	7	7	4	4	8	8	8	5	5	4	4	4	4	4	4	4 - 8	6
tRNA-Thr/tRNA-Pro *(L)*	82	82	111	83	9	9	9	55	55	26	26	27	27	85	5	5 - 111	-
tRNA-Thr/CR	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CR/tRNA-Pro *(L)* +	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	-	N/A; 0	N/A
tRNA-Pro *(L)* /NADH6 *(L)*	23	24	30	30	18	18	17	13	13	14	14	14	14	12	N/A	12 - 30; N/A	9
NADH6 *(L)* /tRNA-Glu *(L)*	3	3	3	3	3	3	3	3	3	-	-	-	-	-	unknown	0; 3	2
tRNA-Glu *(L)* /CR #	-	-	-	-	-	-	-	-	-	-	-	-	-	-	unknown	0	-
CR/tRNA-Phe #	-	unknown	-	-	-	-	-	-	-	-	unknown	unknown	-	-	unknown	0	-
Total Length of Genome	16711	>16757	16828	16816	16714	16704	16747	16591	16595	17061	>16662	>16716	17070	16703	>15302	16591 - 17070	16775

Table 3. Length (in nucleotides) of Paleognath Mitochondrial tRNA and rRNA Genes

tRNAs (22)	Emu	Cassowary	LSKiwi	GSKiwi	Greater Rhea1	Greater Rhea2	Lesser Rhea	Ostrich1	Ostrich2	Eastern Moa1	Eastern Moa2	LBMoa	Giant Moa	Tinamou	CrTinamou	Range (paleognaths)	Chicken
F, Phe	70	70	72	72	68	68	68	69	69	69	69	71	69	66	66	66 - 72	70
V, Val	73	71	73	73	71	71	72	68	68	74	74	74	75	71	70	68 - 75	73
L, Leu (UUR)	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74
I, Ile	71	72	71	71	71	71	71	71	71	71	71	71	71	71	72	71; 72	72
Q, Gln *(L)*	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
M, Met	71	71	69	69	69	69	69	69	69	69	69	69	69	69	69	69; 71	69
W, Trp	75	75	77	77	71	71	71	74	74	77	77	77	76	72	72	71 - 77	76
A, Ala *(L)*	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69
N, Asn *(L)*	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
C, Cys *(L)*	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	66
Y, Tyr *(L)*	71	71	71	71	71	71	71	71	71	72	72	72	72	71	71	71; 72	71
S, Ser (UCN) *(L)*	74	74	74	74	73	73	73	74	74	73	73	73	73	73	73	73 - 74	75
D, Asp	69	69	69	69	69	69	70	69	69	69	69	69	69	69	69	69; 70	69
K, Lys	70	70	72	71	68	68	68	68	68	70	70	70	70	70	70	68-72	68
G, Gly	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69
R, Arg	68	68	68	68	68	68	68	69	69	68	68	68	68	70	69	68 - 70	68
H, His	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	69
S, Ser (AGY)	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	67
L, Leu (CUN)	71	71	71	71	71	71	71	71	71	73	73	73	73	71	72	71; 72, 73	71
T, Thr	70	70	70	70	70	70	70	70	70	70	70	70	70	67	70	67; 70	69
P, Pro *(L)*	70	70	70	70	70	70	70	70	70	70	70	70	70	70	73	70; 73	70
E, Glu *(L)*	68	69	68	68	69	69	69	68	68	67	67	67	67	68	unknown	67 - 69	68
rRNAs (2)																	
12S	964	967	970	969	963	963	964	966	966	969	971	969	968	989	972	963 - 989	975
16S	1596	1592	1595	1595	1583	1583	1585	1579	1580	1600	1600	1602	1601	1588	1585	1579 - 1602	1621

SUMMARY OF RESULTS

The complete mitochondrial genome sequence of the LSK (*Apteryx owenii*) has been submitted to Genbank; acquisition number GU071052. The gene content and organization is the same standard avian gene order reported in the chicken mitochondrial genome (Desjardins and Morais, 1990). The NADH6 (L) / Cytb arrangement is found in all the Paleognaths and Galloanseres with one exception, the elegant creasted tinamou. The length of the LSK genome is 17,020bp and the control region is ~ 1169bp and both are similar in size to other ratite genomes (Table 2). The lengths are not absolute due to heteroplasmy caused by differences in the number of repeated motifs at the 3′ end of the control region. The block of the repetitive motifs (TTTT...A) are between 16,644bp – 16747bp, the number of T nucleotides in the motifs varies but all end with an A nucleotide.

There is an absence of an ori-L in the WANCY region of the mitochondria, only 2bp are present between tRNA-Asn (L) and tRNA-Cys (L), and this is a common characteristic of both birds and crocodilians. The Ori-H is within the control region. Present in the LSK, and all sequenced ratites, is an insertion of a ‘C’ nucleotide in position 174 of the NADH3 gene that appears to have been maintained instead of being selected against. The ‘C’ nucleotide is thought to be deleted by RNA-editing or by a translational frameshift. Both the amino acid sequence after the correction and the final length of the protein is conserved among the paleognaths. A conserved sequence block (CSB) found in the LSK control region (position 15696-15712) was similar to the ostrich and rhea CBII and the chicken CBI.

The intergenic spacer sequences in the LSK range from 0 – 111bp, slightly larger than the GSK and those found in other sequenced ratite mitochondrial genomes (Table 2). The non-coding regions between NADH6 / tRNA-Pro (L) and tRNA-Thr / tRNA-Pro (L) appear to be expanding in the kiwi species. Eleven of the 37 genes are butt-joined and there are three 1bp overlaps between the tRNAs. The first overlap, tRNA-Gln (L) / tRNA-Met, involves both the L and H-strands. The remaining two overlaps occur on the same strands, between tRNA-Cys (L) / tRNA-Tyr (L) on the L-strand and tRNA-Ser (AGY) / tRNA-Leu (CUN) on the H-strand. A post-transcriptional addition of a single residue is thought to occur to one of the same strand overlapping tRNAs to produce the mature tRNA. The one protein-coding / tRNA overlap of 6bp is between COI / tRNA-Ser (UCN) (L). There is post-transcriptional

polyadenylation addition of two 3' A residues to the mRNA to complete the stop codon (TAA) in the protein-coding genes COII and COIII.

There is minor variation in the length of tRNAs and protein-coding genes (Table 3) with two of the protein-coding genes overlapping in different reading frames: ATPase 6 / ATPase 8 and NADH4L / NADH4. The length of seven protein-coding genes and eight tRNAs are conserved among the paleognaths. Most of the protein-coding genes and both rRNA genes are separated by one or more tRNA genes.

The start codon for protein-coding genes in the LSK is Met (ATG/ATA) and COII is Val (GTG). The LSK differs from the other paleognaths that have a COI start codon of Val (GTG) (Table 1).

There is post-transcriptional polyadenylation addition of two 3' A residues to the mRNA to complete the stop codon (TAA) in the protein-coding genes COII and COIII that have an incomplete stop codon T. The length of NADH4 (455 aa) and stop codon (AGA) in both the LSK and GSK is different to other paleognaths that have a conserved length of 458 aa and a stop codon of TAA.

Overall, the mitochondrial genome of the little spotted kiwi was successfully sequenced. It is a typical ratite/tinamou sequence (see Phillips *et al*. 2010, Appendix 2).

Chapter 4 - DISCUSSION

The earliest fossil records or basal lineages of many avian taxa are Gondwanan: ratites, galliforms, anseriforms, parrots, pigeons, passerines, loons and penguins (Cooper and Penny, 1997). A greater diversity of birds occurred in the southern hemisphere (Cracraft, 2001). The ratites originated in the late Cretaceous (Cooper *et al.* 2001; Cracraft, 2001) probably in South America (van Tuinen and Hedges, 2001), the area which has the highest diversity. The oldest unquestionable ratite fossil, *Diogenornis*, was found in Brazil and dated from the early Paleocene (Alvarenga, 1983). Antarctica was one of the major dispersal routes from South America to the other southern continents (Ridley, 1993). There are 46 species of tinamou remaining in South America and the tinamous, also a Gondwanan group, are nested within the ratites (Harshman *et al.* 2008; Phillips *et al.* 2010). The tinamou and ratites (paleognaths) have a monophyletic origin (Prager *et al.* 1976; Casper *et al.* 1994) and the traditional rooting position on the avian tree between paleognaths and neognaths is supported with morphological, nuclear and mitochondrial data (Groth and Barrowclough, 1999; Cracraft *et al.* 2004; Slack *et al.* 2007).

A moa-tinamou sister group (Phillips *et al.* 2010) implies that their ancestors and those of the ostrich were once capable of flight and may have been smaller and slightly stronger fliers than the present tinamou. South America and Africa had separated around 90Mya. It is possible the ratite ancestor(s), aided by a proto-Antillean land connection (van Tuinen *et al.* 1998), flew (or were blown) from South America to North America in the late Cretaceous and then reached Europe during the low sea-levels. At this time the Bering Strait was above water and as a filter bridge it allowed a 'selective' connection between North America and Asia (Ridley, 1993). Early tertiary small flying paleognathous bird fossils (lithornithids) found in the northern hemisphere (Houde, 1986, 1988), indicate that these ratites were once fairly widespread, although no fossils have been found that are older than the Tertiary. The palaeontological research by Olson (1985) and Houde (1988) revealed a series of grades in ratite evolution had occurred in North America and Europe from the late Paleocene to the middle Eocene. These grades, represented by eight species in three genera, are unlikely to be the ancestors of the modern ratites although they appear to be older than the Paleogene ratite fossils of Tambussi *et al.* (1994). The lack of very old ratite fossils in the southern

hemisphere is compounded by the Antarctic ice cap, the 'drowning' of New Zealand (Cooper and Cooper, 1995) and the Australian desert interior (Boles, 1997).

One scenario for ostrich evolution is as follows. The ancestral 'flying' ostrich in South America migrated south, dispersing naturally. A ratite fossil from Seymour Island indicates a migration route from South America through to Antarctica was available (Tambussi *et al.* 1994). Late Cretaceous subaerial connections between Antarctica and Madagascar-Seychelles-India-Sri Lanka through the Kerguelen plateau (see Cracraft, 2001) would have allowed the ostrich to reach Eurasia, when India moved north and collided with Asia (Cooper *et al.* 2001). From there they appear to have eventually migrated around the coast and into Europe and across the land bridge into Africa (~ 23Mya, Phillips *et al.* 2010). An early Miocene ostrich fossil has been found in Africa (Mourer-Chauviré *et al.* 1996) but the earliest ostrich-like fossils are from early - mid Tertiary deposits in Europe, India and Mongolia (Cooper *et al,* 2001). The fossil evidence implies the ostrich evolved in Eurasia before entering Africa. The present day ostrich has lost the ability to fly and the species has grown larger.

The elephant bird could be a descendant of the ostrich (van Tuinen *et al.* 1998; Cooper *et al.* 2001) isolated on Madagascar as India separated and continued north, or it may have been isolated and flightless from an earlier ratite migration. More sequences (~ 1000bp, Cooper *et al.* 2001) are needed from the elephant bird to resolve its position in the evolutionary tree. Indo-Madagascar began separating from Africa 160Mya and then from Australia / Antarctica ~ 130Mya and India split from Madagascar 84Mya, reaching Asia 39Mya (Cracraft, 2001). Phillips *et al.* (2010) suggest the ostrich ceased flying ~ 65Mya. Researching other possible land-based taxa that may have used this pathway to reach Eurasia and Africa would confirm this hypothesis.

Rapid and extensive speciation would have occurred following the availability of different niches as the rhea expanded out into Antarctica and onto the remaining separating landmasses. The timing suggests and also agrees with a rhea-Australasian connection (Prager *et al.* 1976; Sibley and Ahlquist, 1990; van Tuinen *et al.* 1998; Phillips *et al.* 2010). Antarctica and Australia had separated ~ 65Mya (Smith *et al.* 1994), but Tasmania was still linked to both continents until 33-34Mya. The most recent common ancestor (MRCA) of the Australasian clade was on Australia when Australia separated from Antarctica, the emu and

cassowary lineages separating before 25Mya (Boles, 1992, 1997, 2001). The emu and the cassowary are a monophyletic group, with the cassowary migrating as far as New Guinea. The kiwi and the moa diverged from their respective MRCA's before migrating to New Zealand and becoming isolated (Phillips *et al.* 2010). The fossil record for the moa indicates a Pliocene radiation (Baker *et al.* 1995), but there is no good fossil data for the timing of the early kiwi (Worthy and Holdaway 2002). The moa and the kiwi are not a monophyletic group (Cooper *et al.* 1992), however, both show some evidence of a genetic 'Oligocene' bottleneck (Cooper and Cooper, 1995). A tinamou / moa sister group could imply the tinamou flew back from the Antarctica to South America.

From the probable weakly flying ancestor the ratites appear to have had a late Cretaceous speciation following the break-up of Gondwanaland. The migration routes of the ratites follow those of flora / fauna and mammalian dispersion around Gondwana (Ridley, 1993) and the northern continents. There appears to be two subsequent dispersals to New Zealand, the moa and the kiwi. The moa lost flight (~ 56Mya) after arriving in New Zealand and the kiwi may have lost flight before it reached New Zealand shores (Phillips, *et al*. 2010). One theory is the kiwi dispersed from Australia, to New Zealand but no ratite fossils have yet been found on New Caledonia, the main land-based, limited dispersal route (Herzer, 1998). There is no evidence as to the size of the ancestral kiwi before the Pleistocene and had the kiwi been a small weakly flying bird, it is more likely to have arrived in New Zealand, assisted by the strong westerly winds, from Antartica / Australia. There was an island off the northwest coast of New Zealand that existed during the Oligocene (Meffre, *et al.* 2006), and islands like this could have aided dispersal and produced an 'Oligocene' bottleneck for the kiwi. The kiwi could have lost flight during this time on such an island and rafted to New Zealand when the island submerged. Kiwi divergence and radiation is young, only 16Mya (Burbridge *et al.* 2003). The oldest kiwi populations are found in the lower half of the west coast of the South Island and the kiwi appear to have colonised northward from there to the North Island (Baker *et al.* 1995; Burbridge *et al.* 2003).

Since arriving in New Zealand the kiwi, with low dispersal power, has survived multiple significant environmental changes; from subtropical temperatures to glaciations, lush forests to grasslands and near extinction by recent human occupation and introduced terrestrial mammalian predators. After separating from Antarctica, New Zealand was (tectonically) relatively stable until the late Eocene. The Oligocene drowning resulted in the Oligocene

bottleneck for the moa and kiwi. The formation of the Antarctic circumpolar marine currents (ACC) in the Oligocene brought cooler temperatures in the Pliocene and lead to the ice ages in the Pleistocene. Plate tectonic convergence in the Pliocene has produced the uplifting of the Southern Alps, creating highlands and lowlands, and dividing the climate in the South Island. The plates rotation and separation also lead to the opening of Cook Strait (~ 5Mya), which flows between the North and South Islands. The New Zealand coastline has fluctuated widely as plate-tectonic movement has resulted in a lateral displacement of ~ 500kms along the Alpine Fault. The west coast is now relatively more northward and the east coast is conversely more southward. The extensive Pleistocene volcanic activity, mostly in the North Island, has subdued but some activity is still present. Present day sea levels have fragmented the New Zealand sector into three main islands, North, South and Stewart Island.

These geological events have had a major impact on the kiwi evolution. Burbidge *et al.* (2003) estimated the brown kiwi diverged from the spotted kiwi around 16Mya and an early Miocene radiation date is compatible with the re-emergence of land after the Oligocene drowing. Lineage splitting within the sister groups, the brown and spotted kiwi, occurred ~ 4 - 9Mya (Burbridge *et al.* 2003) during the time of mountain-building, major plate tectonics in New Zealand and a cooling climate. The Pleistocene ice age has also aided the geographic isolation and ecological speciation of kiwi. The kiwi colonized northward, dividing into three main North Island populations: Northland, King Country/Taranaki/ Whanganui and Coromandel/Bay of Plenty/East Coast/Hawkes Bay.

Although slowly increasing in number through monitoring and protection programs, the kiwi’s small, geographically isolated habitats, nocturnal behaviour and predation, leaves them vulnerable and prone to local extinctions. Their range-restricted habitats extremely limit the gene flow among kiwi groups and this reduces their diversity. Kiwi have shown adaptation to a wide range of habitats including sub-alpine grassland (Peat, 1999), non-native forests, no forest cover / scrub areas and even farmland (Colbourne, 2005).

All kiwi taxa are currently listed as threatened (Colbourne, 2005) and more New Zealand predator-free areas will be required to maintain the different populations. DoC and kiwi conservation groups are closely monitoring kiwi populations (Holzapfel *et al.* 2008) but more genetic information is required on all kiwi taxa to help identify and prioritize conservation schemes of these species.

Although the allozyme data of LSK on Kapiti and D'Urville Islands is consistent with a South Island origin, it cannot exclude a North Island origin or confirm that the birds are indigenous or were introduced (Herbert and Daugherty, 2002). Future molecular data from the extinct North Island LSK could help determine their phylogenetic relationships and help establish the conservation status of both extant LSK populations. On Long Island two management options are now available; either to allow the LSK to continue to interbreed freely with the Kapiti Island LSK or attempt to breed back to the purer D'Urville Island LSK stock and establish a new founder population (Jolly and Daugherty, 2002). The latter option will conserve any genetic diversity not yet identified in the populations. More evidence is also required to confirm if there are still LSK in the Franz Joseph area or if *A. occidentalis* is a separate species or a recurring hybrid.

It was previously thought the ratites shared a common flightless ancestor and were an example of a vicariant origin. Recent molecular data (Phillips *et al.* 2010) now shows that ratites dispersed and diverged from ancestors that could fly, and each species appears to have individually lost flight in different environments, and evolved in geographical isolation. This makes the modern ratites a product of parallel evolution. Exceptions to a vicariant origin are the tinamou, kiwi and ostrich. The tinamou possibly flew back to South America and both the kiwi and ostrich lost flight ~ 65Mya (Phillips *et al.* 2010) and continued to disperse. The kiwi may have lost flight before dispersing to New Zealand and the ostrich lost flight before dispersing to Eurasia and Southern Africa. Dispersal appears to be a more likely explanation of the ratite distribution and biogeographical history than vicariance.

Ratite fossils have been found on all the southern continents that made up Gondwanaland, and possibly in the northern hemisphere (Houde, 1986, 1988), but there are still gaps in the evolutionary tree (especially the elephant birds). Only the rhea and ostrich are known from the early Tertiary (Cooper and Penny, 1997; Houde, 1986). More taxon sampling is required to resolve the basal ratite phylogeny and obtain more reliable divergence dates and associated conclusions on the timing of dispersal. Estimating divergence times using fossils requires the dated fossils to be identified to the present-day groups. The most recent geological calibration point is the divergent date of the emu and cassowary, calculated from the Australian fossils *Emuarius gidju* (Boles, 1992; 1997) and *Emuarius guljaruba* (Boles, 2001). Molecular clocks have been calibrated using fossils from within (Cooper and Penny, 1997) or outside (Hedges *et al.* 1996; Härlid *et al.* 1997, 1998; Härlid and Arnason, 1999) avian groups. van

Tuinen and Hedges (2001) applied a sequential calibration, using an external calibration to estimate an internal calibration, to estimate the divergence times of the major avian clades. Divergent dates from molecular clock studies are much earlier than indicated by the fossil record.

Complete mitochondrial genomes of the moa have not yet fully resolved early ratite evolution and phylogeny. Haddrath and Baker (2001) placed the moa as the basal taxon, Cooper *et al.* (2001) placed the rhea basal with the moa branching off next and Phillips *et al.* (2010) place the moa and tinamou as a sister group within the paleognaths. The complete elephant bird mitochondrial genome, using ancient DNA methods (and possibly next generation sequencing), may help resolve the earlier ratite lineages of the rhea, moa, ostrich and tinamou, and stablilize the relationships within the ratite evolutionary tree.

REFERENCES

Alvarenga, H.M.F., (1983). Uma ave Ratitae do paleoceno brasileiro: Bacia Calcária de Itaborai, estado do Rio de Janeiro, Brazil. *Boletim do Museu Nacional (Rio de Janeiro), Geologia.* **41,** 1-11. (not sighted)

Baker, A.J., Daugherty, C.H., Colbourne, R., & McLennan, J.L., (1995). Flightless brown kiwis of New Zealand possess extremely subdivided population structure and cryptic species like small mammals. *Proc. Natl. Acad. Sci. USA.* **92,** 8254-8258.

Baker, A.J., Huynen, L.J., Haddrath, O., Millar, C.D., & Lambert, D.M., (2005). Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand. *Proc. Natl. Acad. Sci. USA.* **102,** 8257-8262.

Baker, A.J., & Marshall, H.D., (1997). Mitochondrial Control Region Sequences as Tools for Understanding Evolution. In: Mindell, D.P., (ed.). Avian molecular evolution and systematics. Pp 51–82. Academic Press, San Diego, Calif.

Beutler, E., Gelbart, T., & Kuhl, W., (1990). Interference of Heparin with the Polymerase Chain Reaction. *BioTechniques.* **9,** 166.

Bledsoe, A.H., (1988). A phylogenetic analysis of postcranial skeletal characters of the ratite birds. *Ann. Carnegie Mus.* **57,** 73-90.

Boles, W.E., (1992). Revision of *Dromaius gidju* Patterson and Rich (1987) from Riversleigh, northwestern Queensland, Australia, with a reassessment of its generic position. *Nat. Hist. Mus.* LA Country Sci Ser. **36,** 195-208.

Boles, W. E., (1997). Hindlimb proportions and locomotion of *Emuarius gidju* (Patterson and Rich, 1987) (Aves: Casuariidae). *Memoirs of the Queensland Museum.* **41,** 235-240.

Boles, W.E., (2001). A new emu (Dromaiinae) from the Late Etadunna Formation. *Emu.* **101,** 317-321.

Bourdon, E., De Ricqles, A., & Cubo, J., (2009). A new Transantarctic relationship: morphological evidence for a Rheidae-Dromaiidae-Casuariidae clade. *Zool. J. Linn. Soc.* **156,** 641-663.

Burbidge, M.L., Colbourne, R.M., Robertson, H.A., & Baker, A.J., (2003). Molecular and other biological evidence supports the recognition of at least three species of brown kiwi. *Conserv. Genet.* **4,** 167-177.

Cao, Y., Adachi, J., Janke, A., Pääbo, S., & Hasegawa, M., (1994). Phylogenetic relationships among eutherian orders estimated from inferred protein sequences of

mitochondrial proteins; instability of a tree based on a single gene. *J. Mol. Evol.* **39,** 167-180.

Caspers, G–J., Wattel, J., & de Jong, W.W., (1994). αA-crystallin sequences group tinamou with ratites. *Mol. Biol. Evol.* **11,** 711–713.

Colbourne, R., (2005). Kiwi (*Apteryx* spp.) on offshore New Zealand islands. Populations, translocations and identification of potential release sites. DOC Research and Development Series 208. Department of Conservation, Wellington, New Zealand.

Cooper, A., (1997). Studies of avian ancient DNA: from Jurassic Park to modern island extinctions. In: Mindell, D.P., (ed.), Avian molecular evolution and systematics. Pp 345 – 373. Academic Press, San Diego, Calif.

Cooper, A., & Cooper, R.A., (1995). The Oligocene bottleneck and New Zealand biota: genetic record of a past environmental crisis. *Proc. R. Soc. Lond. B Biol. Sci.* **261,** 293-302.

Cooper, A., Lalueza-Fox, C., Anderson, S., Rambaut, A., Austin, J., & Ward, R., (2001). Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature.* **409,** 704-707.

Cooper, A., Mourer-Chauviré, C., Chambers, G.K., van Haeseler, A., Wilson, A.C., & Pääbo, S., (1992). Independent origins of New Zealand moas and kiwis. *Proc. Natl. Acad. Sci. USA.* **89,** 8741-8744.

Cooper, A., & Penny, D., (1997). Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. *Science.* **275,** 1109-1113.

Cracraft, J., (1974). Phylogeny and evolution of the ratite birds. *Ibis.* **116,** 494–521.

Cracraft, J., (2001). Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. *Proc. R. Soc. Lond. B Biol. Sci.* **268,** 459-469.

Cracraft, J., Baker, F.K., Braun, M., and 14 other authors (2004). Phylogenetic relationships among modern birds (Neornithes): to-ward an avian tree of life. In: Cracraft, J., & Donoghue, M.J. (eds.), Assembling the tree of life. Pp 468-489. New York: Oxford Univ. Press.

de Queiroz, A., (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* **20,** 68-73.

Desjardins, P., & Morais, R., (1990). Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. *J. Mol. Biol.* **212,** 599-634.

Diamond, J.M., (1987). How do flightless mammals colonize oceanic islands? *Nature*. **327,** 374.

Exon, N.F., Kennett, J.P., Malone, M.J., (2009). Antarctic-Australia Separation. http://www.odplegacy.org/PDF/Outreach/Brochures/Greatest_Hits2/exon.pdf (retrieved, 2/12/09)

Feduccia, A., (1996). The origin and evolution of birds. Yale Univ. Press, New Haven, Connecticut.

Felsenstein, J., (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17,** 368-376.

Gibb, G.C., Kardailsky, O., Kimball, R.T., Braun E.L., & Penny, D., (2007). Mitochondrial Genomes and Avian Phylogeny: Complex Characters and Resolvability without Explosive Radiations. *Mol. Biol. Evol.* **24,** 269–280.

Gibbard P.L., Head M.J., Walker, M.J.C. & the subcommission on quaternary stratigraphy. (2010) Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. Journal of Quaternary Science 25: 96-102

Goldberg, J., Trewick, S.A., & Paterson, A.M., (2008). Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. *Phil. Trans. R. Soc. B.* **363,** 3319-3334.

Groth, J.G., Barrowclough, G.F., (1999). Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol. Phylogenet. Evol.* **12,** 115-123.

Haddrath, O., Baker, A.J., (2001). Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. *Proc. R. Soc. Lond. B Biol. Sci.* **268,** 939–945.

Härlid, A., & Arnason, U., (1999). Analyses of mitochondrial DNA nest ratite birds within the Neognathae: Supporting a neotonous origin of ratite morphological characters. *Proc. R. Soc. Lond. B Biol. Sci.* **266,** 305–309.

Härlid, A., Janke, A., & Arnason, U., (1997). The mtDNA sequence of the Ostrich and the divergence between Paleognathous and Neognathous birds. *Mol. Biol. Evol.* **14,** 754-761.

Härlid, A., Janke, A., & Arnason, U., (1998). The complete mitochondrial genome of *Rhea americana* and early avian divergences. *J. Mol. Evol.* **46,** 669–679.

Harshman, J., Braun, E.L., Braun, M.J., Huddleston, C.J., and 17 other authors. (2008). Phylogenomic evidence for multiple losses of flight in ratite birds. *Proc. Natl. Acad. Sci. USA*. **105,** 13462–13467.

Hedges, S.B., (1996). Historical biogeography of West Indian vertebrates. *Annu. Rev. Ecol. Syst.* **27,** 163–196.

Hedges, S.B., Parker, P.H., Sibley, C.G., & Kumar, S., (1996). Continental breakup and the ordinal diversification of birds and animals. *Nature.* **381,** 226-229.

Hendy, M.D., & Penny, D., (1989). A framework for the quantitative study of evolutionary trees. *Syst. Zool.* **38,** 29-309.

Herbert, J., & Daugherty, C.H., (2002). Genetic variation, systematics and management of kiwi (*Apteryx* spp.). In: Overmars, F., (ed.). Some early 1900s studies in kiwi (*Apteryx*) genetics and management. Science and Research Internal Report 191. Pp 11-34. Department of Conservation, Wellington, New Zealand.

Herzer, R.H., (1998). Tectonic control of terrestrial species migration to New Zealand in the early to middle Miocene. In: Cooper, R.A., & Jones, C.M., (eds.). Geology and Genes. Pp 35-38. *Geol. Soc. N.Z. Misc. Publ.* **97.** ISBN 0-908678-69-X, Wellington, New Zealand.

Higuchi. Y., Linn, S., (1995). Purification of all forms of HeLa cell mitochondrial DNA and assessment of damage to it caused by hydrogen peroxide treatment of mitochondria or cells. *J. Biol. Chem.* **270,** 7950-7956.

Holdaway, R. N., Worthy, T. H., & Tennyson, A. J. D., (2001). A working list of breeding bird species of the New Zealand region at first human contact. *N Z. J. Zool.* **28,** 119-187.

Holzapfel, S., Robertson, H.A., McLennan, J.A., Sporle, W., Hackwell, K., & Impey, M., (2008). Kiwi (*Apteryx* spp.) recovery plan. Threatened Species Recovery Plan 60. Department of Conservation, Wellington, New Zealand.

Houde, P.W., (1986). Ostrich ancestors found in the Northern Hemisphere suggest new hypothesis of ratite origins. *Nature.* **324,** 563-565.

Houde, P.W., (1988). Paleognathous birds from the early Tertiary of the Northern Hemisphere. *Publ. Nuttall. Ornithol. Club.* **22,** 1-148.

Janke, A., & Arnason, U., (1997). The Complete Mitochondrial Genome of *Alligator mississippiensis* and the Separation between Recent Archosauria (Birds and Crocodiles). *Mol. Biol. Evol.* **14,** 1266-1272.

Jolly, J.N., & Daugherty, C.H., (2002). Comparison of little spotted kiwi (*Apteryx owenii*) from Kapiti and D'Urville Islands. In: Overmars, F., (ed.). Some early 1900s studies in kiwi (*Apteryx*) genetics and management. Science and Research Internal Report 191. Pp 57-65. Department of Conservation, Wellington, New Zealand.

Jones. M.E.H., Tennyson, A.J.D., Worthy, J.P., Evans, S.E., Worthy, T.H., (2009). A sphenodontine (Rhynchocephalia) from the Mioceneof New Zealand and

palaeobiogeography of the tuatara (Sphenodon). *Proc. R. Soc. Lond. B Biol. Sci.* **276,** 1385-1390.

Jordon, G.J., (2001). An investigation of long-distance dispersal based on species native to both Tasmania and New Zealand. *Aust. J. Bot.* **49,** 333-340.

King, P.R., (1998). Paleogeographic reconstructions of New Zealand. In: Cooper, R.A., & Jones, C.M., (eds.). Geology and Genes. Pp 45-48. *Geol. Soc. N.Z. Misc. Publ.* **97.** ISBN 0-908678-69-X, Wellington, New Zealand.

Knapp, M., Mudaliar, R., Havell, D., Wagstaff, S.J., Lockhart, P.J., (2007). The drowning of New Zealand and the problem of Agathis. *Syst. Biol*. **56,** 862-870.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., & Wilson, A.C., (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA.* **86,** 6196-6200.

Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C., & Sninsky, J.J., (1990). Effects of primer-template mismatches on the polymerase chain reaction; human immunodeficiency virus type 1 model studies. *Nucleic Acids Research*. **18,** 999-1005.

L'Abbé, D., Duhaime, J-F., Lang, B.F., & Morais, R., (1991). The Transcription of DNA in Chicken Mitochondria Initiates from One Major Bidirectional Promotor. *J. Biol. Chem.* **266,** 10844-10850.

L'Abbé, D., Lang, B.F., Desjardins, P., & Morais, R., (1990). Histidine tRNA from Chicken Mitochondria has an Uncoded 5′-Terminal Guanylate Residue. *J. Biol. Chem.* **265,** 2988-2992.

Landis, C.A., Campbell, H.J., Begg, J.G., Mildenhall, D.C., Paterson, A.M., & Trewick, S.A., (2008). The Waipounamu Erosion Surface: questioning the antiquity of the New Zealand land surface and terrestrial fauna and flora. *Geol. Mag.* **145,** 173-197.

Lee, K., Feinstein, J., & Cracraft, J., (1997). The phylogeny of ratite birds: resolving conflicts between molecular and morphological data sets. In: Mindell, D.P., (ed.). Avian molecular evolution and systematics. Pp 173–211. Academic Press, San Diego, Calif.

Mayes, C.L., Lawver, L.A., & Sandwell, D.T., (1990). Tectonic history and new isochron chart of the South Pacific. *J. Geophys. Res*. **95b,** 8543–8567.

Meffre, S., Crawford, A.J., & Quilty, P.J., (2006). Arc-continent collision forming a large island between New Caledonia and New Zealand in the Oligocene. *ASEG Extended Abstracts.***1**. doi:10.1071/ASEG2006ab111.

Micol, V., Fernández-Silva, P., Attardi, G., (1996). *In Vivo* Footprinting of Human Mitochondrial DNA in Cultured Cell Systems. *Meth. Enzymol.* **264,** 3-11.

Mindell, D. P., & Honeycutt, R.L., (1990). Ribosomal RNA in vertebrates: Evolution and phylogenetic applications. *Ann. Rev. Ecol. Syst.* **21,** 541-566.

Mindell, D.P., Sorenson, M.D., & Dimcheff, D.E., (1998a). Multiple independent origins of mitochondrial gene order in birds. *Proc. Natl. Acad. Sci. USA*. **95,** 10693–10697.

Mindell, D.P., Sorenson, M.D., & Dimcheff, D.E., (1998b). An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* **15,** 1568-1571.

Mindell, D.P., Sorenson, M.D., Dimcheff, D.E., Hasegawa, M., Ast, J.C., & Yuri, T., (1999). Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst. Biol.* **48,** 138-152.

Mourer-Chauviré, C., Senut, B., Pickford, M., & Mein, P., (1996). The oldest representative of the genus Struthio (Ave: Struthionidae), *Struthio coppensi* n. sp., from the lower Miocene of Namibia. *C. R. Acad. Sci. Paris.* **322,** 325-332.

Olson, S.L., (1985). The fossil record of birds. In: Farner, D.S., King, J.R., & Parkes, K.C., (eds.). Avian biology. Vol. VIII. Pp. 79 – 238. Academic Press, Orlando, Florida.

Ojala, D., Montoya, J., & Attardi, G., (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature.* **290,** 470–474.

Packer, L., (1967). Experiments in Cell Physiology. Academic Press, New York.

Peat, N., (1999). Kiwi New Zealand's remarkable bird. Random House, Auckland, New Zealand.

Penny, D., & Phillips, M.J., (2004). The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution. *Trends Ecol. Evol.* **19,** 516–522.

Phillips, M.J., Gibb, G.C., Crimp, E.A., & Penny, D., (2010). Tinamous and Moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. *Syst, Biol.* **59,** 90-107.

Prager, E.M., Wilson, A.C., Osuga, D.T., & Feeney, R.E., (1976). Evolution of flightless land birds on southern continents: Transferrin comparison shows monophyletic origin of ratites. *J. Mol. Evol*. **8,** 283-294.

Ridley, M., (1993). *Evolution*. Blackwells, Cambridge, MA.

Shepherd, L.D., & Lambert, D.M., (2008). Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. *Mol. Ecol.* **17,** 2174-2184.

Sibley, C.G., & Ahlquist, J.E., (1981). The phylogeny and relationships of the ratite birds as indicated by DNA-DNA hybridization. In: Scudder, G.G.E., & Reveal, J.L. (eds.), Evolution Today. **2,** 301-335. *Proc. Int. Congr. Syst. Evol. Biol.* Hunt Inst. Botanic Document, Pittsburgh.

Sibley, C.G., & Ahlquist, J.E., (1990). Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale Univ. Press, New Haven, Connecticut.

Sibley, C.G., & Frelin, C., (1972). The egg white protein evidence of ratite affinities. *Ibis.* **114,** 377–387.

Slack, K.E., Delsuc, F., McLenachan, P.A., Arnason, U., & Penny, D., (2007). Resolving the root of the avian mitogenic tree by breaking up long branches. *Mol. Phylogenet. Evol.* **42,** 1-13.

Slack, K.E., Janke, A., Penny, D., & Arnason, U., (2003). Two new avian mitochondrial genomes (penquin and goose) and a summary of bird and reptile mitogenic features. *Gene.* **302,** 43-52.

Smith, A.G., Smith, D.G., & Funnell, B.M., (1994). Altas of Mesozoic and Cenozoic Coastlines. Cambridge University Press, Cambridge, England.

Stapel, S.O., Leunissen, J.A.M., Versteeg, M., Wattel. J., & de Jong, W.W., (1984). Ratites as the oldest offshoot of avian stem - evidence from α-crystallin A sequences. *Nature.* **311,** 257–259.

Steadman, D.W., (2006). Extinction and biogeography of tropical Pacific birds. Chicago, University of Chicago Press.

Tambussi, C.P., Noriega, J.I., Gazdzicki, A., Tatur, A., Reguero, M.A., & Vizcaino, S.F., (1994). Ratite bird from the Paleogene La Meseta Formation, Seymour Island, Antarctica. *Pol. Polar Res.* **15,** 15-20.

Trewick, S.A., (1996). Morphology and evolution of two takahe: flightless rails of New Zealand. *J. Zool.* Lond. **238**, 221–237.

Trewick, S.A., (1997). Flightlessness and phylogeny amongst endemic rails (Aves: Rallidae) of the New Zealand region. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **352,** 429-446.

Trewick, S.A., Paterson, A.M., Campbell, H.J. (2007) Hello New Zealand. *J. Biogeog*. **34,** 1-6.

van Tuinen, M., & Hedges, S.B., (2001). Calibration of Avian Molecular Clocks. *Mol. Biol. Evol.* **18,** 206-213.

van Tuinen, M., Sibley, C.G., & Hedges, S.B., (1998). Phylogeny and Biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. *Mol .Biol. Evol*. **15,** 370-376.

van Tuinen, M., Sibley, C.G., & Hedges, S.B., (2000). The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Mol. Biol. Evol.* **17,** 451-457.

Veevers, J.J., (1991). Phanerozoic Australia in the changing configuration of proto-Pangaea through Gondwanaland and Pangaea to the present dispersed continents. *Aust. Syst. Bot.* **4,** 1-11.

Waters, J. M., & Craw, D., (2006). Goodbye Gondwana? New Zealand Biogeography, Geology, and the Problem of Circularity. *Syst. Biol.* **55,** 351-356.

Worthy, T.H., & Holdaway, R.N., (2002). The lost world of the moa: prehistoric life of New Zealand. Pp 8. Canterbury University Press, Christchurch, New Zealand.

Worthy, T.H., Tennyson, A.J.D., Archer, M., Musser, A.M., Hand, S.J., Jones, C., Douglas, B.J., McNamara, J.A., & Beck, R.M.D., (2006). Miocene mammal reveals a Mesozoic ghost lineage on insular New Zealand, southwest Pacific. *Proc. Natl. Acad. Sci. USA*. **103,** 19419-19423.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K., (2001). Trends, rhythms, and aberrations in global climate 65Ma to Present. *Science.* **292,** 686-693.

APPENDIX 1

Sequences of short-range PCR primers used to amplify and sequence the LSK mitochondrial genome, plus GenBank submission of final sequence.

FORWARD PRIMERS

L1753	5'AAA CTG GGA TTA GAT ACC CCA CTA T 3'	25 bp	12SA	FH
L2035	5' GAG ATG GAA GAA ATG GGC TAC 3'	21 bp	12SC2	Fb
L2471	5' AAA AGT TAC TTG ACG CAA TAG AG 3'	23 bp	16S	Fb
L2907	5' CAT GTT GGC CTT CAA GCA GC 3'	20 bp	16S	Fb
L3787	5' CGA TTA ACA GTC CTA CGT G 3'	19 bp	16S	Fb
L4747	5' CCA TTC GCC CTA TTC TTC CTA GC 3'	23 bp	ND1	FH
L5197	5' GGT CAG CTA ATT AAG CTA TCG 3'	21 bp	tMet	FH
L5574	5' CTA GGT CTA GTC CCC TTC C 3'	19 bp	ND2	FH
L5976	5' CCA ATA CTA AAT GCC ACC CTA ATA C 3'	25 bp	ND2	Fb
L6335	5' GGC CTT CAA AGC CTT AAA TAA GAG 3'	24 bp	tTrp	Fb
L6634	5' CCT AAC GCT TGA ACA CTC AGC 3'	21 bp	tTyr (L)	Fb
L7056	5' CCC TTT AGC AGG TAA CCT CG 3'	20 bp	COI	Fb
L7459	5' GAA CCA TTC GGA TAT ATA GG 3'	20 bp	COI	FH
L7546	5' GTC GGA ATA GAC GTA GAC AC 3'	20 bp	COI	FH
L7807	5' CTC TCT ATA GGT GCC GTA TTT GC 3'	23 bp	COI	FH
L8320	5' TCA CAG GTC AAA ACC CTG 3'	18 bp	tAsp	Fb
L9246	5' CTG AAC CTG ACC ATG AAC C 3'	19 bp	ATP8	Fb
L9712	5' TTC GAC CAT TAG CCC TAG G 3'	19 bp	ATP6	Fb
L9874	5' GCA GTA GCA ATA ATC CAA GC 3'	20 bp	ATP6	FH
L10716	5' GAG GTT CCT GCT CTT CTA G 3'	19 bp	tGly	FH
L11904	5´CCC GTT GAG GTA ATC AAC C 3'	19 bp	ND4	FH
L12312	5´GCA CTA ATA ACC AGC TCA ATC TGC 3'	24 bp	ND4	FH
L12788	5´CTC AAA CAC ACG AGA ACA CC 3'	20 bp	ND4	FH
L13052	5´GGT CTT AGG AAC CAC CCA TCT TGG TG 3'	26 bp	ND5	FH
L13562	5´GAT GAC ATG GAC GAG CAG AA 3'	20 bp	ND5	Fb
L14002	5´CCA CAT CAA GCC AAC TAG G 3'	19 bp	ND5	FH
L14452	5´GCA TCA TAG CAG GCC TCC TA 3'	20 bp	ND5	Fb
L15212	5´GGC CGA GGA CTC TAC TAC GGC TC 3'	23 bp	Cytb	Fb
L15303	5' GGT TAT GTC CTC CCA TGA GGA 3'	21 bp	Cytb	Fb
L15671	5' CCC AGA AAA CTT CAC ACC AGC 3'	21 bp	Cytb	FH

REVERSE PRIMERS.

H1805	5' TTA AGC GTT TGT GCT CGT AG 3'	20 bp	12SP2	Rb
H2150	5' GAG GGT GAC GGG CGG TGT GTA C 3'	22 bp	12SB5	RH
H2577	5' GCT TAA ATT CAT TTT GCT TGG 3'	21 bp	16S	Rb
H2919	5' GTT GAG CTT TGA CGC ACT C 3'	19 bp	16S	Rb
H3420	5' CAG GGT CTT CTC GTC TTA TG 3'	20 bp	16S	Rb
H3797	5' CGA CCT GGA TTT CTC CGG TCT G 3'	22 bp	16S	RH
H4311	5' GGG ATT CAA ATG GTG ATG GCT AG 3'	23 bp	ND1	Rb
H4791	5' GAA CTT GGG TTG AGG AAT AG 3'	20 bp	ND1	Rb
H5201	5' CCA TCA TTT TCG GGG TAT GG 3'	20 bp	tMet	RH
H5583	5' CCT TGG AGG ACT TCT GGG AA 3'	20 bp	ND2	Rb
H6120	5' GCA TAG TAT GCG AGG CGT AG 3'	20 bp	ND2	RH
H6367	5' GAT ACA GGT TAA TGT CCT GTT GG 3'	23 bp	tAla	Rb
H6642	5' CGG GTA ATG AAG GTC ACA GG 3'	20 bp	COI	Rb
H7062	5' GCA AGG TCT ACA GAG GCT CCA GC 3'	23 bp	COI	Rb
H7553	5' GTG GCT GAT GTG AAG TAG 3'	18 bp	COI	Rb
H7816	5' GGG AAT CAG TGT GTA AAT CCT G 3'	22 bp	COI	Rb
H8325	5' GGT TGG CCA CAT GAG ATG 3'	18 bp	COII	Rb
H8807	5' GGA ACT GCT CAT GAG TGG 3'	18 bp	COII	Rb
H9284	5' GTA GGA TTA GTG GGA TTC C 3'	19 bp	ATP8	Rb
H9755	5' GTG GCA GTA GAG ATG AGT TGG 3'	21 bp	ATP6	Rb
H10229	5' TTG GGG CTA GGC TTG AGT G 3'	19 bp	COIII	Rb
H10743	5´GGG GTT ACA CCA GAT TTT TAG AG 3'	23 bp	t-Gly	Rb
H11152	5´ATT TGC TGA GCC GAA ATC AG 3'	20 bp	t-Arg	Rb
H12800	5´GTA TGG GGA GAA TGT GGA G 3'	19 bp	ND4	Rb
H13200	5´GGG ATG AGG CTG AGA AAG AAG CG 3'	23 bp	ND5	Rb
H13592	5´GCC TAC GTC TCC AAT ACG GTT GTA G 3'	25 bp	ND5	Rb
H14539	5' CCG AGG ATT GTG ACG ACG ATG G 3'	22 bp	ND5	RH
H14946	5'AGG GGG AGG GGA GGT CAA TTA GG 3'	23 bp	Cytb	Rb
H15345	5'CCG ATG TAG GGG ATG GCT GAG 3'	21 bp	Cytb	Rb
H15649	5'TTG CTG GTG TGA AGT TTT CTG GGT C 3'	25 bp	Cytb	Rb

The L - strand is the forward (F) primer and the H - strand is the reverse (R) primer. The primer numbering refers to the position of the 3’ base in the chicken sequence (Desjardins and Morais, 1990). (H and b) refer to a primer compatible with human (H) mtDNA and a bird (b) specific primer respectively (see Cooper *et al.* 2001). Bird specific primers were used to remove the risk of human contamination.

```
LOCUS       GU071052        17020 bp    DNA     circular VRT 29-JAN-2010
DEFINITION  Apteryx owenii mitochondrion, complete genome.
ACCESSION   GU071052
VERSION     GU071052   GI:262179844
SOURCE      mitochondrion Apteryx owenii (little spotted kiwi)
  ORGANISM  Apteryx owenii
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria;
            Aves; Palaeognathae; Apterygiformes; Apterygidae; Apteryx.
REFERENCE   1  (bases 1 to 17020)
  AUTHORS   Phillips, M.J., Gibb,G.C., Crimp,E.A. and Penny,D.
  TITLE     Tinamous and Moa Flock Together: Mitochondrial Genome Sequence
            Analysis Reveals Independent Losses of Flight among Ratites
  JOURNAL   Syst. Biol. 59 (1), 90-107 (2010)
REFERENCE   2  (bases 1 to 17020)
  AUTHORS   Gibb,G.C.
  TITLE     Direct Submission
JOURNAL     Submitted (08-OCT-2009) Allan Wilson Centre for Molecular Ecology
            and Evolution, Institute of Molecular BioSciences, Massey University,
            Private Bag 11-222, Palmerston North, New Zealand
FEATURES     Location/Qualifiers
   source        1..17020
           /organism="Apteryx owenii"
           /organelle="mitochondrion"
           /mol_type="genomic DNA"
           /db_xref="taxon:8824"
           /tissue_type="blood"
           /country="New Zealand: Otorohanga Kiwi House"
   tRNA          1..72
           /product="tRNA-Phe"
   rRNA          73..1042
           /product="12S ribosomal RNA"
   tRNA          1043..1115
           /product="tRNA-Val"
   rRNA          1116..2710
           /product="16S ribosomal RNA"
   tRNA          2711..2784
           /product="tRNA-Leu"
           /codon_recognized="UUR"
   CDS          2794..3768
           /codon_start=1
           /transl_table=2
           /product="NADH dehydrogenase subunit 1"
           /protein_id="ACY28892"
           /db_xref="GI:262179845"

/translation="MWPNTLTYLTMSLSYVIPILIAVAFLTLVERKILSYMQARKGPNVVGP
FGLLQPVADGVKLFIKEPIRPTISSPLLFTATPVLALLLAITIWIPLPLPFPLDLNLGILF
LLSMSSLAVYSILWSGWASNSKYALIGALRAVAQTISYEVTLAIILLSIIMLSGNYTL
NTLAVTQEPLYLIFSTWPLMMMWYISTLAETNRAPFDLTEGESELVSGFNVEYAAG
```

PFALFFLAEYANIMLMNALTTILFLNPSSLNLSSELFPLVLATKTLLLSSGFLWVRAS
YPRFRYDQLMHLLWKNFLPLTLALCLWHTSMPISYAGLPPYL"
tRNA 3767..3837
/product="tRNA-Ile"
tRNA complement(3847..3917)
/product="tRNA-Gln"
tRNA 3917..3985
/product="tRNA-Met"
CDS 3986..5024
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:5024,aa:TERM)
/transl_table=2
/product="NADH dehydrogenase subunit 2"
/protein_id="ACY28893"
/db_xref="GI:262179846"

/translation="MNPHAKLISLLSLFLGTTITISSNHWVMAWTGLEINTLAIIPLISKSHHP
RAVEASIKYFLVQATASALVLFSSISNAWVTGQWDITQMTHPTSCLLLTTAIAMKL
GLAPFHFWFPEVLQGTSLTTALLLSTIMKFPPMTILLMTSHSLNHTMLVSLALISAA
LGGWMGLNQTQTRKILAFSSISHLGWMAIIITYSPKLTILTFYLYILMTTAVFLSLNT
TKTLKLSTMMTSWTKTPVLNASLMLTLLSLAGLPPLAGFMPKWLILQELTKQEMT
VTATIISMLSLLGLFFYLRLAYYATITLPPNSANHMKQWHINKTTSTLTAILTSLSITL
LPLTPMMFATL"
tRNA 5025..5101
/product="tRNA-Trp"
tRNA complement(5103..5171)
/product="tRNA-Ala"
tRNA complement(5173..5245)
/product="tRNA-Asn"
tRNA complement(5248..5314)
/product="tRNA-Cys"
tRNA complement(5314..5384)
/product="tRNA-Tyr"
CDS 5386..6936
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit I"
/protein_id="ACY28894"
/db_xref="GI:262179847"

/translation="MTFTTRWLFSTNHKDIGTLYLIFGAWAGMVGTALSLLIRAELGQPGT
LLGDDQIYNVIVTAHAFVMIFFMVMPVMIGGFGNWLVPLMIGAPDMAFPRMNNM
SFWLLPPSFLLLLASSTIEAGAGTGWTVYPPLAGNLAHAGASVDLAIFSLHLAGISSI
LGAINFITTAINMKPPALSQYQTPLFVWSVLITAILLLLSLPVLAAGITMLLTDRNLN
TTFFDPAGGGDPVLYQHLFWFFGHPEVYILILPGFGMISHVVTYYAGKKEPFGYMG
MVWAMLSIGFLGFIVWAHHMFTVGMDVDTRAYFTSATMIIAIPTGIKVFSWLATL
HGGTIKWDPPILWALGFIFLFTIGGLTGIVLANSSLDIALHDTYYVVAHFHYVLSMG
AVFAILAGFTHWFPLFTGYTLHPTWAKAHFGVMFTGVNLTFFPQHFLGLAGMPRR

YSDYPDAYTLWNTMSSIGSLISMTAVIMLMFIIWEAFSSKRKMMQPELTTTNIEWIH
GCPPPHHTFEEPAYVQIQE"
tRNA complement(6928..7001)
/product="tRNA-Ser"
/codon_recognized="UCN"
tRNA 7004..7072
/product="tRNA-Asp"
CDS 7074..7761
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:7761,aa:TERM)
/transl_table=2
/product="cytochrome c oxidase subunit II"
/protein_id="ACY28895"
/db_xref="GI:262179848"

/translation="MANPSQFGFQDASSPIMEELVEFHDHALMVALAICSLVLYLLTLMLM
EKLSSNSVDAQEVELIWTILPAIVLILLALPSLQILYMMDEIDEPDLTLKAIGHQWY
WTYEYTDFKDLSFDSYMIPTSELPSGHFRLLEVDHRVVVPMESPIRVIITAGDVLHS
WAVPTLGVKTDAIPGRLNQTSFITTRPGIFYGQCSEICGANHSYMPIVVESTPLPHFE
NWSSLLSTSS"
tRNA 7762..7833
/product="tRNA-Lys"
CDS 7835..8002
/codon_start=1
/transl_table=2
/product="ATP synthase subunit 8"
/protein_id="ACY28896"
/db_xref="GI:262179849"

/translation="MPQLNPNPWFIIMLMSWLTFSFLMQPKLLSFLPTNLPTSKTKTI
LNPTPWSWPWT"
CDS 7993..8676
/codon_start=1
/transl_table=2
/product="ATP synthase subunit 6"
/protein_id="ACY28897"
/db_xref="GI:262179850"

/translation="MNLSFFDQFASPQLLGIPLILLSMLFPALLLPSPTNRWVTNRLS
TLQSWFLQLITKQLMIPLNKMGHKWALILSSLMTFLLLINLLGLLPYTFTPTTQLSM
NMALAFPLWLATLLTGLRNQPSISLGHLLPEGTPTILIPALILIETTSLLIRPLALGVRL
TANLTAGHLLIQLISTATMALLPIMPTISILTAFILLLLTILELAVAMIQAYVFVLLLSL
YLQENI"
CDS 8676..9459
/note="TAA stop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:9459,aa:TERM)

/transl_table=2
/product="cytochrome c oxidase subunit III"
/protein_id="ACY28898"
/db_xref="GI:262179851"

/translation="MTHQAHSYHLVDPSPWPIFGATTALLTTSGLIMWFHYNSMYLLTLGL
LSMLLVMLQWWRDIVRESTFQGHHTPTVQKGLRYGMILFITSEAFFFLGFFWAFFH
SSLAPTPELGGLWPPTGIKPLNPLEVPLLNTAILLASGVTVTWAHHSITEGNRKQAT
HALLLTILLGFYFTALQAMEYHEAPFSIADGVYGSTFFVATGFHGLHVIIGSSFLTIC
LLRLIKYHFTSDHHFGFEAAAWYWHFVDVIWLFLYMTIYWWGS"
tRNA 9460..9528
/product="tRNA-Gly"
CDS join(9529..9702,9704..9880)
/note="frameshift mechanism unknown (Mindell et al.,
1998,Mol. Biol. Evol., 15:1568-1571)"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 3"
/protein_id="ACY28899"
/db_xref="GI:262179852"

/translation="MNMITFMLLLSLLLSIILITLNFWLAQMNPDTEKLSPYECGFDP
LGSARLPFSIRFFLVAILFLLFDLEIALLLPLPWAIQLPHPLLTLLWTSTILILLTLGLV
YEWAQGGLEWAE"
tRNA 9882..9949
/product="tRNA-Arg"
CDS 9950..10246
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 4L"
/protein_id="ACY28900"
/db_xref="GI:262179853"

/translation="MSLLHLSFYSAFTLSGLGLAFHRTHFISALLCLESMMLSLYISLSIWPIQ
TQTPAFSLIPVLMLAFSACEAGAGLAMLVASTRTHGSDHLHNLNLLQC"
CDS 10240..11607
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 4"
/protein_id="ACY28901"
/db_xref="GI:262179854"

/translation="MLKIILPTAMLLPMALLSPKNLLWTNITTHSLLIATISLQWLHP
TYFPYKYLSQWTGIDQTSAPLLVLSCWLLPLMLMASQNHLRQEPLMRKRIFITTLI
MIQPFIILAFSATELILFYISFEATLIPTLILITRWGNQPERLSAGIYLLFYTLISSLPLLV
TLLHLHMQTGTLHLPTLELTHTSLTLSWTGTLSGLALLMAFMVKAPLYGLHLWLP
KAHVEAPIAGSMLLAALLLKLGGYGIMRVTLLMGPLMNYLCYPFLTLALWGALM
TSSICLRQTDLKSLIAYSSVSHMSLVIAAGMIQTNWSFSGAMILMISHGLTSSMLFCL
ANTNYERTHSRILLLTRGLQPLLPLMSTWWLLANLTNMALPPTTNLMAELTIMIAL

FNWSMFTIILTGLATLLTASYTLFMLLMTQRGTLPAHITSTPNTTTREHLLMTLHILP
MFLLMLK PELIL"
tRNA 11615..11684
/product="tRNA-His"
tRNA 11685..11750
/product="tRNA-Ser"
/codon_recognized="AGY"
tRNA 11750..11820
/product="tRNA-Leu"
/codon_recognized="CUN"
CDS 11821..13635
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 5"
/protein_id="ACY28902"
/db_xref="GI:262179855"

/translation="METTLLLNTSILLTLTILFTPTILPLLSSKLQATPTTITSTVKTAFLTSLIP
MTIFIYSGTETIISHWYWNFTPNLKVPLSFKMDQYSMLFLPIALFVTWSILQFAIWY
MASEPHINKFFTYLLMFLIAMLTLTIANNMFMLFIGWEGVGIMSFLLIGWWHGRAE
ANTAALQAVIYNRIGDVGLILSMAWLASTLSTWEIQQTSYENQIPMLPLLGLILAAA
GKSAQFGLHPWLPAAMEGPTPVSALLHSSTMVVAGIFLLIRTHPLLTNNQTALTLC
LCLGALSTLFAATCALTQNDIKKIIAFSTSSQLGLMMVTIGLNLPQLAFLHISTHAFF
KAMLFLCSGSIIHSLAGEQDIRKMGGLQKLLPTTTSCLTIGNFALMGTPFLAGFYSK
DLIIESLNNSYLNTWALLLTLLATSFTATYSLRMTLMVQTGFSRIPPITPTNENHPMV
TNPITRLAMGSIMAGLLITSNILPTKTPPMTMPMITKTAAIIVTIVGILLALELSNMTH
NLMLPKQSTYSNFTTTLGYFNPILHRLNPAKLLNNGQKIASHLIDFSWYKKVGPEG
LADLQLMATKTSTSLHTGLIKTYLGSFALSIFVLLLLI"
CDS 13648..14787
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="ACY28903"
/db_xref="GI:262179856"

/translation="MAPIIRKSHPLLKIINNSLIDLPSPSNISAWWNFGSLLGICLITQILTGLLL
AMHYTADTSLAFSSIAHICRNVQYGWLIRNLHANGASFFFICIYFHIGRGFYYGSYL
YKETWNTGIILLLTLMATAFVGYVLPWGQMSFWGATVITNLFSAIPYIGQTLVEWA
WGGFSVDNPTLTRFFALHFLLPFLIAGTTLIHLTFLHESGSNNPLGIISHCDKIPFHPY
FSLKDTLGFTLMLIPLLTLTFFSPNLLGDPENFTPANPLVTPPHIKPEWYFLFAYAILR
SIPNKLGGVLALAASVLILFLTPLLHKSKQRSMAFRPFSQALFWLLVANLLILTWVG
SQPVEHPFIIIGQLASFTYFLILLVLFPTIGALENKMIY"
tRNA 14789..14858
/product="tRNA-Thr"
tRNA complement(14970..15039)
/product="tRNA-Pro"
CDS complement(15067..15588)
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 6"

/protein_id="ACY28904"
/db_xref="GI:262179857"

/translation="MSYLVFFLGLCFVLGVLAVASNPSPYYGILGLVVGSVAGCGCLL
SLGMSFVSLVLLLVYLGGMLVVFVYSVSLAADPFPEAWGNWRVWGYGMGF
LLVVFVGVVVGGLMELEMGGVRTVDCGGPVSVRMDFSGVAMFYSWGVGMFLV
AGWGLLLTLFVVLE LVRGLSRGTIRAV"
tRNA complement(15592..15659)
/product="tRNA-Glu"
misc_feature 15660..17020
/note="control region"
BASE COUNT 5256 a 4994 c 2289 g 4479 t 2 others
ORIGIN
1 gtcctcatag cttacaaacc tcaaagcgcg gcactgaaga tgccaagacg gtagccacca
61 tacctgagga caaaagactt agtcctaacc ttactgttaa ttctcgctaa acgtatacat
121 gcaagtatct gcattccagt gtaaatgccc ttaacgtctt ctttccccag acaaaaggag
181 caggtatcag gcacacgcct atcgtagccc aaaacacctt gctcagccac acccccacgg
241 gtactcagca gtaactaaca ttaagcaata agtgcaaact tgacttagtt atagcgacca
301 tagggttggt aaatcttgtg ccagccaccg cggtcacaca agaaacccaa actaacagta
361 accggcgtaa agtgtggtaa catgttatcg aaataactaa gaccaaaaca caactgagct
421 gtcataagcc caagatgtac ttaaaacctt cataaaaatg atcttagtac ctacgaccaa
481 tttaaatcca cgaaagctaa gacacaaact gggattagat accccactat gcttagccct
541 aaatcctggt gcttacatta cctaagtacc cgcccgagaa ctacgagcac aaacgcttaa
601 aactctaagg acttggcggt gccctaaacc cacctagagg agcctgttct ataatcgata
661 acccacgata cacccaacca tctcttgcca acacagccta tataccgccg tcgccagctc
721 gcctaaaatg agagaacaac agcgaacaca acagctatcc cagctaacaa gacaggtcaa
781 ggtatagcct atgagatgga agaaatgggc tacattttct aaaatagaat actacgaaaa
841 agggtgtgaa atccccctta gaaggcggat ttagcagtaa aacagaataa gagagtctat
901 tttaagctgg ccctagggca cgtacatacc gcccgtcacc ctcctcacaa gctacattac
961 actttataca taataaatca ctaagctaaa gatgaggtaa gtcgtaacaa ggtaagtgta
1021 ccggaaggtg cacttagcac accaagacgt agctacaaac ccaaagcatt cagcttacac
1081 ctgaaagatg tctactagat catagaccgt cttgaggccc tccctctagc tcaattatcc
1141 tctaacataa gccaaacaat ctactataac tcctaaacca aaacattttc tacacttagt
1201 ataggcgata gaaaagctac ttgacgcaat agagacattt ccgtaccgta agggaaagat
1261 gaaatagcaa tgaaacacat aaagcaaaaa atagcaaaga tcaacccttg taccttttgc
1321 atcatgattt agcaagaaca tccaagcaca atgaatttaa gcttgctatc ccgaaaccca
1381 agcgagctac tcacaagcag ctaccattga gcgaacccgt ctctgtagca aaagagtggg
1441 aagacttgtt agtagaggtg aaaagccaat cgagctgggt gatagctggt tgcttgtaaa
1501 acgaatctca gttcacctct aacttccccc accagacacc cgccttaacc ctaatgtgga
1561 aagttagaag taatttaaag gaggtacagc tcctttaaaa aagaagacaa tctctaccag
1621 agggtaattt tatccattac accgatacgt aggccctaaa gcagccatca acaaagaatg
1681 cgtcaaagct caatcagctc taaaaatcca aaaacattac aatcccctca ccattaacaa
1741 gccaacctat aatcatagga gaattaatgc tagaatgagt aactagaagc cccttctctt
1801 aagcgcaaac ttacatcacc catcattaac agactaactt atactaccac ctccacaaga
1861 accaatatta aactaccctg ttaccccaac tcaggagcgc ctattagaaa gattaaaatc
1921 tgtaaaagga actaggcaaa cttagggctc gactgtttac caaaaacata gccttcagcc
1981 aacaacaagt attgaaggtg atgcctgccc agtgacccat gttcaacggc cgcggtatcc
2041 taaccgtgca aaggtagcgc aatcaattgt cccataaatc gagacttgta tgaatggcta
2101 aacgaggtcc taactgtctc ttacggataa tcagtgaaat tgatcttctt gtgcaaaagc
2161 aagaatataa acataagacg agaagaccct gtggaacttg aaaatcaagg accaacgtat

2221 ctaaattcca agcctataag gccaaccacc cttacaacaa ctggtcccta tttttcggct
2281 ggggcggcct tggagaaaaa aagagcctcc aaaaataaga ccatacacct tgactaagaa
2341 caactcctca aagtactaat agtaaccaga cccaatacta ttgattaatg aacaaagcta
2401 ccccagggat aacagcgcaa tctccttcaa gagcccatat cgacaaggag gtttacgacc
2461 tcgatgttgg atcaggacat cctaatggtg cagccgctat taagggttcg tttgttcaac
2521 gattaacagt cctacgtgat ctgagttcag accggagcaa tccaggtcgg tttctatcta
2581 tgcaacactc ttcccagtac gaaaggaccg gaaaagtaag gcccatacca aacagcacgc
2641 cttctttcta attagtgaaa ccaacttaac taacaagaac ctcttcccat cacccctaga
2701 aaagggctaa gctagtgtgg cagagctcgg caagtgcaaa aggcttaagc cctttaccca
2761 gaggttcaaa tcctctccct agctctactt cacatgtgac ccaacaccct aacctacctc
2821 actatatcac tatcatatgt catcccaatt ctgattgcag tagcttttct tacactagta
2881 gaacgaaaaa tcctaagcta catacaagcc cgaaaaggcc caaacgtcgt aggtccattt
2941 ggactcctgc aacccgtagc agacggagta aaactgttca ttaaagaacc catccgccct
3001 acaatctcat ccccgcttct attcactgca acccctgtct tagccctcct actagctatc
3061 accatctgaa tccccctccc cctccccttc cccttacag atctaaactt aggcatccta
3121 ttcctcctat caatatcaag cttagccgtg tattctattc tatgatcagg atgggcctca
3181 aactcaaaat acgctctaat tggagcccta cgagcagttg cacagactat ttcttatgaa
3241 gtaacattag caattatcct cctatccatt attatactaa gtggaaacta caccttaaac
3301 accctcgccg tcactcaaga gcccctctac ctaatcttct ccacctgacc cctaataata
3361 atatgataca tttctacact tgcggagaca aaccgagcac cattcgattt aaccgaagga
3421 gaatcagaac tcgtctcagg atttaacgta gaatatgctg caggcccttt tgccctattc
3481 tttctagccg aatacgccaa catcatacta ataaatgcat taactaccat tctattcctc
3541 aacccaagct cacttaacct ctcatccgaa ctctttccat tagtcctagc cacaaagacc
3601 ctattactct cttcgggctt tttatgagta cgagcatcct accctcgatt ccgttacgac
3661 cagctaatac acctcctctg aaaaaacttc ctacccctga cattagcact atgtttatga
3721 cacacaagca taccaatctc atatgcaggc cttccacctt acttaaggaa atgtgcctga
3781 accaaagggt cactatgata aagtgaacat agaggtttac aaaccctctc atttcctgaa
3841 aaggatttag aaaagtagga atcgaaccta cacagaggag atcaaaaccc ctcatacttc
3901 ctctatatta ttttctagta aggtcagcta attaagctat cgggcccata ccccgaaaat
3961 gatggtttaa tcccttccct tactaatgaa ccctcacgct aaactaatct ccctcctcag
4021 tcttttccta ggcacaacaa ttaccatctc aagcaaccat tgagtaatag cttgaactgg
4081 attagaaatc aacaccctcg caattattcc actcatctct aaatcccacc accctcgggc
4141 cgtagaagca tcaatcaaat attttctcgt acaagcaact gcctcagccc tcgtcctatt
4201 ctcaagtatt tctaatgctt gagttactgg tcaatgagac atcactcaaa taacacatcc
4261 aacctcctgt ctactattaa cgacagcaat cgccatgaaa cttggcctag caccccttcca
4321 tttctgattc ccagaagtct tacaaggaac atccctaact actgccctct tactctccac
4381 aatcataaaa ttcccccccta taaccattct cctcataacc tcccactcac taaaccatac
4441 aatacttgtc tcactggccc ttatttcagc tgcactggga ggatgaatag gactaaacca
4501 aacgcaaact cgtaaaatcc tagccttttc atccatctct catctaggat gaatagctat
4561 catcatcacc tacagcccca aactcacaat tctaaccttc tacctctata tcttaataac
4621 aacagcagta tttttatctc tcaacacaac aaaaacccta aagctatcta caataataac
4681 ctcttgaaca aaaactcctg tactaaacgc ctccctaata ttaacactct tatccctagc
4741 cggtctacct ccactagcag gattcatacc caaatgacta atcctccaag aactaaccaa
4801 acaggaaatg acagtcacag ccacaatcat ttccatgcta tccctactag gcttattctt
4861 ctacctccgc ctcgcatact atgcaacaat caccctcccc ccaaactctg ccaaccacat
4921 aaaacaatga cacattaata aaaccacaag caccttaact gccatcctca catccctctc
4981 cattactctc ctacctctca cacccataat atttgctacc ctctagaaac ttaggataac
5041 cataaaccaa aaccaaaggc cttcaaagcc ttaaataaga gttaaagcct cttagtttct
5101 gctaagacca acaggatact aacctgtatc tcctgaatgc aaatcagaca ctttaattaa
5161 gctaaggcct tcctagacag gtgggccttg atcccacaat aatctagtta acagctaaat

5221 gcctaatcct gcaggcttct gtctaacaaa ctccggtgca cctttagtac acatcaatga
5281 gcttgcaact caacatgaat ttcactacgg agttgataag aagaggaatc aaacctctat
5341 aaaaaggact acagcctaac gcttaatcat tcagccatct tacccatgac ctttactact
5401 cgatgactct tctcaactaa ccacaaagat atcggcacac tatatctcat tttcggtgca
5461 tgagcaggta tagtgggcac cgccctcagc ctgctcatcc gagccgaact cggtcaacca
5521 ggaaccttac taggagatga ccaaatctac aacgttatcg tcaccgccca tgcctttgta
5581 ataatcttct tcatagtgat gcccgtaata atcggaggct tcgggaactg acttgtcccc
5641 ctcataattg gcgctccaga catagcattt ccacgaataa ataacataag cttctgactt
5701 ctacccccct cattcctact cctgctagca tcatccacca ttgaagctgg agcagggacc
5761 ggatgaacag tataccccc attagccgga aaccttgccc atgcaggcgc atctgtagac
5821 cttgctatct tttcactcca cctagcaggt atctcttcaa tcttaggggc aattaacttc
5881 atcactacag caatcaacat aaaaccccca gccctctcac aataccaaac acccctattc
5941 gtatgatccg tactaatcac tgctattctt ctcctactgt ccctcccagt actcgctgct
6001 ggcattacaa tacttctcac agatcgaaac cttaatacta cattctttga ccccgctgga
6061 ggaggagacc ccgtcctata ccagcaccta ttctgattct tcggccaccc tgaagtctat
6121 attttaattc tacccggctt tggaataatc tcacacgtag taacctatta tgcaggtaaa
6181 aaagaaccct tcggatatat aggaatagtc tgagctatac tatctatcgg attcctagga
6241 ttcatcgtat gggcccacca catatttaca gtaggaatgg acgtagacac tcgagcctac
6301 tttacatccg ccaccataat catcgctatc ccaacgggta ttaaagtatt cagctgacta
6361 gctacactcc acggagggac aatcaaatga gatcctccta tcctatgagc cctggggttc
6421 attttcctat ttactattgg tggtttaacc ggtatcgtac tagcaaattc atcactcgac
6481 attgccctac acgatacata ctacgtagta gctcatttcc actatgtact ctccataggt
6541 gctgtctttg ctattctagc tggctttaca cactgattcc cactatttac cggatacacc
6601 ctccacccaa catgagcgaa agcacacttc ggggttatat tcacaggagt caacctaacc
6661 ttctttcccc agcacttcct aggcttagct ggaataccac gacggtactc agactacccg
6721 gatgcttata ccctatgaaa taccatatcc tcaattggct cactaatctc tataacagcc
6781 gtaatcatac ttatatttat tatctgagaa gcattctcct caaaacgaaa aataatacaa
6841 ccagaattaa ctaccaccaa catcgaatga atccacggct gcccaccccc acaccacacc
6901 tttgaagaac cagcctatgt tcagatccaa gaaaggaagg aatcgaacct ccacacgcta
6961 gtttcaagcc agctgcacta aaccacttat gcttctttct tatgagaagt tagtaaatac
7021 attacatagc cttgtcaaga ctaaatcaca ggttaaaacc ctgtacatct catgtggcca
7081 acccctctca atttggattt caagacgcat catcccctat tatagaagaa ttagtcgaat
7141 ttcatgacca tgccctaata gttgcactag caatctgtag cctagtactt tatttactca
7201 ctctaatatt aatagaaaaa ctatcttcca actcagtaga cgcgcaagaa gtagaactaa
7261 tctgaacaat cctgcccgct atcgtactaa tcttacttgc tcttccctcc ctccaaatcc
7321 tgtacataat agatgaaatc gatgagcctg acctaaccct aaaggctatc ggccatcaat
7381 gatactgaac ctatgaatac acagacttta aagacctgtc atttgactcc tacatgattc
7441 caacatcaga actaccctca ggccacttcc gactattaga agtagaccat cgagttgttg
7501 ttccaatgga atcacccatt cgagtaatca tcactgctgg agacgtactt cactcatgag
7561 cagttccaac cctaggagta aaaactgatg caatcccagg acgactgaac caaacatcct
7621 ttattactac tcgacctgga atcttctatg gtcaatgttc agaaatttgt ggagccaacc
7681 acagctatat gcccattgta gtagaatcaa ctccacttcc acactttgag aattgatcat
7741 cattactttc cacatcctca tcattaagaa gctatgataa cagcactagc cttttaagct
7801 agagaaagag gaatacttcc ccctccttaa tgatatgccc caactcaacc caaacccatg
7861 atttattatc atacttatat cgtgattaac cttttcattc cttatgcaac ctaaactttt
7921 atcctttcta cctacaaacc ttccaaccag caagactaaa acaatcctta atcccacacc
7981 atgatcctga ccatgaacct aagcttcttc gaccaatttg ccagccccca actactagga
8041 atcccactaa tcttactctc tatgctattc cctgctttat tgctcccatc cccccaccaac
8101 cgctgagtaa ccaatcgcct ttcaactctt caatcatgat ttcttcaact aattaccaaa
8161 caactaataa tcccattaaa caagataggc cataaatgag ctcttatcct ctcatcacta

8221 ataacctttc ttttactaat caacctccta ggactactac catacacatt tacaccaact
8281 acccaattat ccataaacat agcactagcc tttcccctct gactagcaac tctccttact
8341 ggactacgaa accaaccatc tatctctcta ggtcatctcc taccagaagg gacccctaca
8401 atcctcatcc ccgccctcat ccttattgaa acaacaagcc tacttatccg accattagcc
8461 ctaggtgttc gactcacagc taaccttaca gctggccacc tcctaattca acttatctct
8521 accgccacta tagccctact tccaatcata cccaccattt ccatcctaac agcattcatc
8581 ctactcctac taacaatcct agaactagca gtagctataa tccaagctta cgtctttgta
8641 ctcttattaa gcttatactt acaagagaat atttaatgac acaccaagca cattcctacc
8701 acctagtaga ccccagccca tgacccatct ttggtgccac cactgctcta ttaacaacct
8761 ccggactgat catatgattc cactataact caatgtatct attaacccta ggcttactct
8821 caatactcct agttatacta caatgatgac gagatatcgt acgagagagc acctttcaag
8881 gccatcacac ccctaccgtc caaaaaggcc tacgatacgg aataattcta tttatcacat
8941 cagaagcatt cttctttcta ggattcttct gagcattctt ccattcaagc ctagctccta
9001 caccagaatt aggaggacta tgacccccca caggaatcaa gcccctcaac cccctagaag
9061 tccccctact caacacagcc atcctcctag catcaggcgt caccgtaaca tgagcgcacc
9121 acagtattac agagggtaat cgaaaacagg caacacatgc cctactccta acaatcctcc
9181 taggatttta ctttaccgcc ttacaagcca tagaatacca tgaagccccc ttctcaattg
9241 ccgatggggt ctatggctct accttttttg tagccacagg attccacgga ctacacgtaa
9301 ttattggatc atcctttctt accatctgtc tcttacgact aatcaagtat cactttacat
9361 cagaccatca cttcggattc gaagctgcag cttgatattg gcacttcgta gacgttatct
9421 gattattcct ctacataacc atctactgat gaggatcatg ctcttctagt atatcaatta
9481 caattgactt ccaatctcta aaatctggta ttaccccaga gaagagcaat aaacataatc
9541 acatttatac tccttctctc cctactatta agcattatcc taattacact aaacttttga
9601 ttagctcaaa taaacccaga tacagaaaaa ctatccccat atgaatgcgg atttgatccc
9661 ctaggatcag cccgactccc cttctcaatc cgattcttcc tcagtagcca tcctattcct
9721 actcttcgac ttagaaatcg ctcttcttct ccccctcccc tgagcaatcc aactcccaca
9781 ccctctctta accctattat gaacttccac tatcctcatc ctcctcacac taggactagt
9841 ttatgaatga gcacagggag gcctagaatg ggcagagtaa cagaaagtta gtctaactaa
9901 gacagctgat ttcggctcag cagattatag tcacccctat aactttctta tgtccttact
9961 acacctaagc ttctactcag ctttcactct aagtggccta ggcctagcct ttcaccgaac
10021 ccactttatc tctgccctac tatgcctaga aagtatgata ctctctctat acatctccct
10081 ctcaatctga cctatccaaa cccaaacccc cgcattctcc ctgattccag tactcatact
10141 agccttttcc gcatgtgaag caggagcagg actagccata ttagtagcct ctactcgcac
10201 acatggctcc gatcacctac ataacctaaa cctcctacaa tgctaaaaat tattcttccc
10261 acagccatgc ttttacccat agccctttta tcccctaaaa accttctctg aacgaacatt
10321 accacacata gcctcctaat tgctaccatc agcttacagt ggcttcaccc aacctacttc
10381 ccctataaat acctctccca atgaacaggc attgatcaaa cctcagcccc actactagtc
10441 ctatcatgct gactcctccc tcttatacta atagcaagcc aaaaccatct tcgacaagaa
10501 cccctaatac gaaaacgtat cttcattaca accctcatca taatccaacc cttcattatc
10561 ttagccttct cagccacaga attaatccta ttctacattt catttgaagc caccctaatt
10621 cctacactaa tcctaattac tcgatgagga aaccaacctg aacggcttag tgctggtatc
10681 tacctactat tttacaccct tatcagctcc cttccactcc tagtcactct ccttcaccta
10741 cacatacaaa ccggtacttt acatctacca accctagaac taacccatac atcattaaca
10801 ctttcatgaa ctggaaccct atcaggatta gccctcttaa tagcattcat agtaaaagcc
10861 ccgctatatg gacttcacct gtgattacct aaagctcatg tagaagcccc cattgcaggc
10921 tccatactac tagctgcact tttactaaaa ctcggaggct atggtattat acgagtaacc
10981 ctactcatag gacctctcat aaactactta tgctacccat tcctaaccct cgccttatga
11041 ggagcactaa taacaagctc aatttgtctc cgtcaaactg atctaaagtc tctaattgca
11101 tactcatctg taagccatat aagcctagtc atcgccgctg gcataatcca aaccaactga
11161 tcattctcag gagcaataat cctcataatc tcacatggtc tgacatcctc aatattattc

11221 tgcttagcta acacaaacta cgaacgtaca cacagccgga tccttctact aactcgaggc
11281 ctacagcccc tcctacccct tataagtaca tggtgactcc tggctaacct aacaaacata
11341 gcactacccc ccacaaccaa cctaatagca gaactcacca tcataattgc actcttcaac
11401 tgatccatat tcactattat cctaaccggt cttgcaacac tactaaccgc ctcatatact
11461 ctcttcatac tcttaataac ccaacgagga acactacctg cccacattac atccactcct
11521 aacacaacca cacgagaaca tctcctaata actctccaca ttctccccat attccttctt
11581 atactaaagc ccgaactcat cttaagaacc ctatgcaaat atagtttcaa cccaaacatt
11641 agactgtgat tctaaaaata gaagttaaaa tcttcttatt cgccaagggg tggtttaacc
11701 agcaagaatt gctaactcct gcatctgagt ctaaaccctc agcccccttg cttttaaagg
11761 ataatagtaa tccactggcc ttaggagcca ctcatcttgg tgcaactcca agtgaaagca
11821 atggagacca cactactcct taacacctct atccttctca cactgacaat tttattcaca
11881 ccaactatcc tacccctctt atcatccaaa cttcaagcaa ctcccacaac catcacatca
11941 actgtcaaaa cagccttcct cacaagtctc attcctataa ctatcttcat ttactccggg
12001 acagaaacca tcatctctca ttggtactga aacttcaccc ccaacctcaa agtcccactc
12061 agcttcaaaa tagatcagta ttcaatatta ttcctcccca tcgccttatt tgtaacctga
12121 tccatcctac aattcgcaat ctgatacata gcatcagaac ccccacatcaa caaattcttc
12181 acctacctat taatattctt aatcgctata ctaaccctga caatcgctaa caacatattc
12241 atattattca ttggctgaga aggcgtaggc attatatcat tcctactcat tggctgatgg
12301 catggacgag cagaagccaa cacagccgcc cttcaagcag taatctacaa ccgaattgga
12361 gatgtaggcc tcatcctcag tatggcttga cttgcttcca ctctaagtac atgagaaatc
12421 caacaaacct cctacgaaaa ccaaatcccc atactacctc tcctaggcct tattctagct
12481 gccgcaggaa aatcagccca atttggtcta cacccctgac tcccagcagc aatagaaggt
12541 cctacacccg tttcagccct actacactcc agcaccatag tagtagcagg aattttcctt
12601 cttatccgca cccatccact actaactaac aaccaaacag ctttaaccct atgcctatgc
12661 ctaggggccc tatccacact atttgctgcc acatgtgctc tcacccaaaa cgatatcaaa
12721 aaaatcattg ccttctctac atcaagccaa ctaggattaa taatagtcac cattggactt
12781 aacctacccc aactagcctt cctccacatc tcaactcatg ctttctttaa agcaatatta
12841 ttcttatgct ccggatcaat catccacagc ctagcaggtg aacaagacat tcgaaaaata
12901 ggaggactac aaaaactcct cccaacaact acatcctgcc taacaattgg taattttgcc
12961 ctaataggaa ctccatttct tgctggattt tactcaaaag acctcattat cgaaagcctc
13021 aacaactcct atctaaacac ctgagctctc ttactcacac tcctagccac ttccttcact
13081 gccacttaca gtcttcgaat aacccttata gtccaaacag gatttagccg aatccctcca
13141 atcaccccaa ctaacgaaaa ccaccccata gtcaccaacc caatcactcg actcgcaata
13201 ggtagcatca tagctggcct tcttatcaca tccaatatcc tcccaacaaa aacacctcca
13261 ataaccatac ccataatcac aaaaacagcc gctatcatcg tcactatcgt aggcattctc
13321 ttagccctag aactatcaaa cataacccac aacctcatac tccccaaaca aagcacctac
13381 tcaaacttca ctaccactct aggctatttc aaccccatct tacatcgact aaaccctgca
13441 aaactactca acaacggaca aaaaatcgcc tcacacctaa tcgacttctc ttgatacaag
13501 aaagtagggc cagagggatt agctgatctt caactcatag caaccaaaac ctcaacctcc
13561 ctacataccg gactaatcaa gacgtaccta ggatcattcg ctctatctat tttcgtacta
13621 cttttactca tctaaaaaac aaaaccaatg gccccaatca tccgaaaatc ccaccccctta
13681 cttaaaatca tcaacaactc tctgatcgac cttccctccc cctcaaatat ctctgcgtga
13741 tgaaacttcg gatccttact aggaatctgc ctgattaccc aaatcctaac aggactgcta
13801 ctagccatac actatacagc tgacacatcc ctagcctttt catccatcgc ccatatctgt
13861 cgaaacgtcc aatatggatg actaattcgt aacctacatg caaacggagc atcattcttc
13921 tttatctgta tttacttcca catcggccga ggattttact atggctcata cttatacaaa
13981 gaaacctgaa acactggcat tatcctctta ctgacactaa tagcaactgc ctttgtaggc
14041 tatgtccttc catgaggaca aatatcattt tggggggcta cagttatcac gaacctattc
14101 tcagctatcc catacatcgg acaaacccta gtagaatgag cttgaggagg attctcagta
14161 gataacccca cacttacacg attcttcgcc ctacacttcc tccttccatt cctaatcgct

14221 ggtactaccc tcatccacct cacctttcta catgaatcag gatcaaacaa ccccctagga
14281 attatctcac actgtgacaa aatccccttc cacccatact tctccctaaa agatactcta
14341 ggattcacac tcatactcat ccccctacta accctaactt tcttctcccc caaccttcta
14401 ggagacccag agaactttac tccagcaaac cctctagtca caccccccaca tattaagcca
14461 gaatgatact tcctatttgc atacgccatc ctacgttcta tccccaacaa attaggaggg
14521 gtacttgctt tagctgcatc cgtattaatc ctatttttaa ctccactcct acacaaatcc
14581 aaacaacgct caatagcctt ccggcccttc tctcaagcac tattctgact actcgtagcc
14641 aacttgctta ttctgacatg agtaggtagc cagccagtag aacacccatt catcatcatc
14701 ggccaattag cctcattcac ctacttcctc atcctcctag tcttatttcc tacaattgga
14761 gctctggaga ataaaataat ttactaatac tctaatagtt tatagaaaac attggtcttg
14821 taaaccaaag actgaagaat acacctcttc ttagagtaca ctcctctcag taactggggg
14881 gggggggrr gggggggga aagaaggaga aaattaaaac aatagaaaaa aatgagaaaa
14941 taacaaacca taaagctata tccaacatct cagaaaaaaa ggaatcaaac ctttacctcc
15001 agctcccaaa gctggtattc tccattaaac tatcttctga tagctacaac tccctccctc
15061 ctccccctac actgcccgaa tcgtcccccg agataacccc cgcacaagct ccaaaaccac
15121 aaacaaagtg agcaataacc cccaccctgc aactaaaaac atccccaccc cccaagaata
15181 aaacatagct accccactaa aatctattcg aacagaaact ggacctccac aatccacagt
15241 acgtactccc cccatctcaa gctccattaa acctcccact accaccccca caaaaacaac
15301 caacaaaaac cccataccat aacctcagac tcgccaattc cctcaagcct caggaaacgg
15361 atcagcagcc aacgacacag aatagacaaa taccacaagc attcccccca aatataccaa
15421 aagtaacact aaagatacaa acgacattcc caaactcaac aaacagccac atccagcaac
15481 cgaccccaca accaatccaa gaatcccata ataaggcgaa ggattagaag ccactgctag
15541 tacccccaaa acaaagcaca accctaaaaa aaatacaagg taagacatag tagtttctgc
15601 ctagctacta cccagggcct acggtctgaa aaaccatcgt tgtcctcaac tacagaaaca
15661 gttcatattc ctctcccaac ataaaattct ttccccccct acccccccat gtataattgt
15721 acattcgttt accaacccca tacaaatctt tgctcgatgt atcacccccg aacacggtct
15781 aaccgcaccc cttgtaggca aatacagtat ggtcaccgaa cacatttctt ggagccgggg
15841 ttaaattcaa tagactaggc taaacgcatt aatagctgct ctgtacatgc gcaatgatct
15901 accccaacag attacccgta agatcctcgt tcaatggtta atggacaaac cctacctaac
15961 tatgctcgtc aatctgactg tgattcctgg cgccagatgg atttcctgat tcaacacctc
16021 acgtgaaatc agcaacccct gtaaaacgta tccaatatga ctagcttcaa gaccattcat
16081 tccccgcgca ctaccgggtc ataaatcgat tggggcgcct ctggttccta ggtcaaggcc
16141 attaattgat tgctccattt acgttcccct ttaagagaca tctggtattg tgagtttaca
16201 ccatctcacc cgtaatcgcg gcatccatat acgctttagt tcgcctggtt ccctttttttt
16261 ttctctggaa ccttcaacgt acccccatac tatgcgtgtc agaaggctca cactttaggt
16321 ctgagcatcc atggacctcg ggtttcgtct ttcggcaggg accgattaat gagacggttg
16381 gcgtatatgt ggaatcattc gggctctgat gcacttgcta ttccatttgg ttatggagta
16441 tccacccgga gaatctgcgt gtacaattca gttaatggtt gcaggacata actctcttac
16501 tttctacctc tttttacttc cgctaacttt ctaaacaaca ctagagcaat ttcatctatt
16561 actatcacca tcatttttgc atcatcatca cacattttat cttcaccacc cctccatcac
16621 cagcactaaa atttcattaa taatttttttt tttatttttt tattttttat tttttttatt
16681 tttttattt ttttatttt ttttattttt tttattttttt ttatttttttt tattttttta
16741 tttttattc tcttcattat ttatccctat cactagtact gaaattacat taataatttc
16801 tctttctttt cttttctttt tttctcttca ttatttatcc ctatcactag tactgaaatt
16861 acattaataa ttttttccct tttttcccct ctcccttcct ttcccctatt tattttatgc
16921 tggaattcta ttaataaccc cgcacccaca ttatacacac ttattaaaaa aatcctactc
16981 actataacga cccccttgca aaaaacaaac aaggcaaaaa