The Evolution of the Mitochondrial DNA in Tuatara (*Sphenodon punctatus*)

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Evolutionary Genetics

Massey University
Auckland, New Zealand

Elmira Mohandesan

October 2010

Science grows like a weed every year

- Kary Mullis

to Eli,
Best of Luck

Abstract

The enigmatic tuatara from New Zealand, occupies a central place in the evolution of vertebrates and tuatara have changed little morphologically since the Cretaceous period approximately 200 millions of years ago. A central aim of this thesis was to examine rates of molecular evolution in tuatara using entire mitochondrial genomes of both ancient and modern samples. A total of 51 complete mitochondrial genomes from 42 modern (from eight island groups) and 9 ancient samples (from eight locations on mainland) were sequenced using Sanger method.

These complete genomes were used to investigate the population genetic structure of tuatara. Diverse phylogenetic analyses suggest that *Sphenodon* is a monotypic genus. This is in contrast to the suggestion made by Daugherty et al. (1990 b) that there are two species of tuatara. This two species model was subsequently almost universally accepted. The current result illustrates that when species are incorrectly identified scarce conservation resources are inappropriately used to ensure their conservation.

Using these complete mitochondrial genomes and by employing three very different analytical methods, I have also estimated molecular evolutionary rates for tuatara. Using modern and ancient complete mitochondrial genomes, I showed that the rates of molecular evolution in tuatara are higher than other vertebrates. This result and the stable morphology of tuatara over tens of millions of years suggest a disconnect between molecular and anatomical evolution, as originally suggested by Allan Wilson in the 1970s. From a biological perspective perhaps this is not surprising, since morphological and molecular evolution are governed by very different biological processes.

I then explored the possibility that tuatara might be characterised by high mutation rates. Using Roche 454 next generation DNA sequencing, I recovered seven complete mitochondrial genomes in tuatara. A total of 28 potential heteroplasmies were detected among these genomes. These sites were also shown to be polymorphic among the 42 modern aligned genomes suggesting that they are characterised by high mutation rates. This result suggests that a high level of heteroplasmic sites in tuatara mitochondrial genome contributes to the high molecular rates estimated when comparing modern and ancient genomes.

Acknowledgement

Pursuing a Ph.D. project is a challenging experience that is not possible without the personal and practical support of numerous people. Although, it will not be enough to express my gratitude in words, I would still like to thank all of them for their love, support, and patience over the last few years.

First of all, it has been an honor for me to be Professor David Lambert's PhD student whose research record lists over 100 publications in such diverse areas as Ancient DNA research, molecular evolution, conservation genetics, evolutionary biology and evolutionary theory. I owe my deepest gratitude to him for all he has done for me from the first day of my arrival in New Zealand to all the way through my PhD. He always offered me valuable advices, patiently commented on manuscripts and with his cheerful enthusiasm and ever-friendly nature accompanied me through all hardships and frustrations.

Special thanks are also given to my ex-supervisors, Dr Javad Mowla and Dr Michael Hofreiter, who were instrumental in shaping up my academic career. Dr Mowla encouraged me to work hard and be brave to be a first student in Iran who worked on ancient DNA and molecular evolutionary field. He had confidence in me when I doubted myself. I continued my research with Dr Michael Hofreiter who invited me, as a visiting scholar, to Max Planck Institute where I learned a lot from him and his research group. He helped me to find a supervisor when I expressed my desire to pursue my PhD. During these years, my supervisors have been my friends and mentor and helped me to follow my dream "ancient DNA and evolution". Without their support and encouragement a newcomer like me would not be able to perform a PhD project in this field.

My thanks go out to Dr Shankar Subramanian for providing guidance, resources, and intellectual discussions. He has patiently taught me how to analyse my data using various analytical methods. He helped me to interpret my data with his valuable suggestions and constant guidance.

Thesis Structure, Financial Support and Regulatory Compliance

The first three chapters of this thesis give a broad overview of mitochondrial DNA evolutionary rates, its structure and function and the species classification and geographical distribution of the unique New Zealand reptile, tuatara (*Sphenodon*). These provide the background and intellectual framework for this thesis. Chapter four presents the materials and methods used in order to perform the research, including the samples, laboratory and analytical method. Chapter five presents the empirical data and phylogenetic analyses of 42 complete mitochondrial genomes. These were used to assess the genetic diversity and taxonomy of tuatara. Important consequences in relation to conservation priorities and management decisions are discussed. Chapter six provides a description of various analytical methods used to estimate molecular evolutionary rates for tuatara mitochondrial genome. Here, I present the analytical data related to molecular evolutionary rates estimated for tuatara complete mtDNA and specific rates for tRNAs, tRNAs, synonymous and non-synonymous regions. In chapter seven I use 454 sequencing data of seven complete mitochondrial genomes to investigate mutation rates in tuatara. I detected 28 potentially heteroplasmic sites among these genomes. An analysis of the 42 aligned genomes showed that these sites showed a polymorphic pattern among these genomes. This result further suggests that the high evolutionary rate characteristic of tuatara is driven by a high mutation rate. In chapter eight (discussion and conclusion), the correlation between molecular and morphological evolution were discussed.

The appendices also presented in this thesis derive from a number of studies. During the course of my PhD I have contributed to three published papers in collaborating. The first was titled "*Rapid Molecular Evolution in living fossils*" and was published in Trends in Genetics, 2008. This paper was featured on the cover of the issue and has been widely publicised around the world. For this paper I conducted laboratory analyses. The second one was a review article entitled "*New developments in ancient genomics*", published in Trends in Evolution and Ecology in 2008 and again made the cover of the issue. I reviewed the research articles and wrote sections relating to ancient DNA. In addition, I contributed to another paper entitled "*Molecular and morphological*

evolution in tuatara are decoupled ", published in Trends in Genetics in 2008. My contribution involved writing sections of the manuscript. I also present, in the appendix, a paper entitled "*Ancient DNA from Human and Animal Remains from North-West Iran.*" This paper was based on the results of research conducted at Max Planck Institute in Germany. I was the senior author of this paper that was published in the Journal of Sciences in 2008. In collaboration with others, I designed the study and performed the laboratory work at Max Planck Institute in Germany. The data analysis and the entire writing of this paper were performed during the course of my PhD. I took the major role in writing but had contributions from the other authors.

Financial Support

Table of Contents

List of Figures

List of Tables

Chapter One

Mitochondrial DNA Evolutionary Rates

1.1 Introduction

Historically evolution was understood as changes in the morphology and the species composition of animals and plants over long geological time spans. Implicit in this view was the idea, held by the early evolutionists like Lamarck and Darwin, that evolution was typically if not exclusively a gradual and slow process. More recently, evolution has come to be understood largely at a genetic and molecular level and the expectation is that rates and modes of the molecular evolution in different branches of the Tree of Life are likely to be positively correlated with rates of diversification and speciation. For this reason, rates of molecular evolution continue to be fundamental in our understanding of evolution itself.

Increased rates of molecular evolution will influence genetic variation within species and can ultimately lead to speciation as a result of evolving reproductive barriers and reducing gene flow. Moreover, rates of molecular evolution help us to time the evolution of particular characters or states and more generally to place a time scale on phylogenesis.

The rates of diversification and molecular evolution vary dramatically across phylogenetic lineages (Bromham & Penny 2003; Kumar 2005). For example, a comparison in mammalian lineages shows that rodents are evolving at a faster rate compared to apes and humans, which evolve more slowly (Kumar 2005). The evolutionary rates in avian species are relatively slower than most mammals (Mindell et al. 1996) and, relative to the other reptiles, snakes and lizards are evolving rapidly compared with turtles (Avise et al. 1992; Hughes & Mouchiroud 2001; Eo & DeWoody 2010). Although, there are valuable data available in this field most studies have sampled only a relatively small number of species and/or a small number of genes. Therefore, in order to make robust conclusions about the molecular evolutionary rates in any species a large number of genes should be evaluated in order to accurately estimate molecular rates. The complete mitochondrial genome sequence is an excellent model system for achieving this goal as it consists of reasonable number of genes, conservative gene order, manageable genome size (~16 kb), and heterogeneity in substitution rates. This work aims to sequence and analyse the complete mtDNA of modern and ancient tuatara populations in order to: (i) estimate molecular evolutionary rates for the complete mitochondrial genome of tuatara; (ii) partition this rate variation into different genes; and (iii) to examine the data for patterns of molecular evolution across the genome more generally.

1.2 The Molecular Evolutionary Rates

The threads of different types of molecular evolutionary rates can be found in the literature. These include mutation rates, pedigree rates and substitution rates (Ho & Larson 2006). Mutation rates referred to the within generation changes in nucleotide sequences in the genome, while the pedigree rates are estimated based on the calculation of the number of the nucleotide changes over several reproductive events within a small number of known generation times (Ho & Larson 2006). Substitution rates estimate the frequency at which the mutations are fixed in populations typically over a larger number of generations (Figure 1.1). The majority of the nucleotide changes in the genome are eventually removed from the gene pool of the population as a result of purifying selection or drift, therefore the substitution rate is always less than mutation rate, unless the perfect neutral condition applies. Failure to distinguish between these rates has given rise tosui some disputes in evolutionary biology. The inappropriate application of substitution rates to intra-specific

evolutionary questions has occurred in numerous studies such as dating of human migration, origins (Ward et al. 1991; Armour et al. 1996; Mellars 2005; Hovers et al. 2003) and animal domestication (Kim et al. 2002; Jansen et al. 2002; Guo et al. 2005). Although the reliable estimation of divergence dates for recent evolutionary events is difficult, this can be achieved using molecular data. Moreover, there are several options that can improve the accuracy of the molecular rate estimates. Using the calibration points relevant to the particular study (e.g. intra-specific calibration points for population-level studies) (Stoneking et al. 1992), reliable ancient DNA data from radiocarbon dated ancient materials (Shapiro et al. 2004; Drummond et al. 2003) and incorporating a relaxed molecular clock framework (Thorne et al. 1998) are examples.

According to molecular clock hypothesis the speed of nucleotide changes over a period of time is relatively constant and most of the changes in amino acid and nucleotide sequences are neutral rather than beneficial or deleterious (Kimura 1983). In contrast to this, it has been argued that molecular data show that rates can differ across organisms, genes and different timescale of the evolutionary events (Ho & Larson 2006; Pulquerio & Nichols 2007; Kumar & Subramanian 2002; Howell et al. 2003; Denver et al. 2000; Emerson 2007). In order to establish an absolute time scale, one approach is to incorporate some form of calibration information such as a known substitution rate (Shields & Wilson 1987; Randi 1996; Fleischer et al. 1998), heterochronous sequences with known ages (Lambert et al. 2002; Paxinos et al. 2002), or inferred ages of lineage splitting event into analysis. These three methods have been used in phylogenetic studies of birds, however their suitability of their application is not satisfactory (Ho 2007; Ho et al. 2005 a).
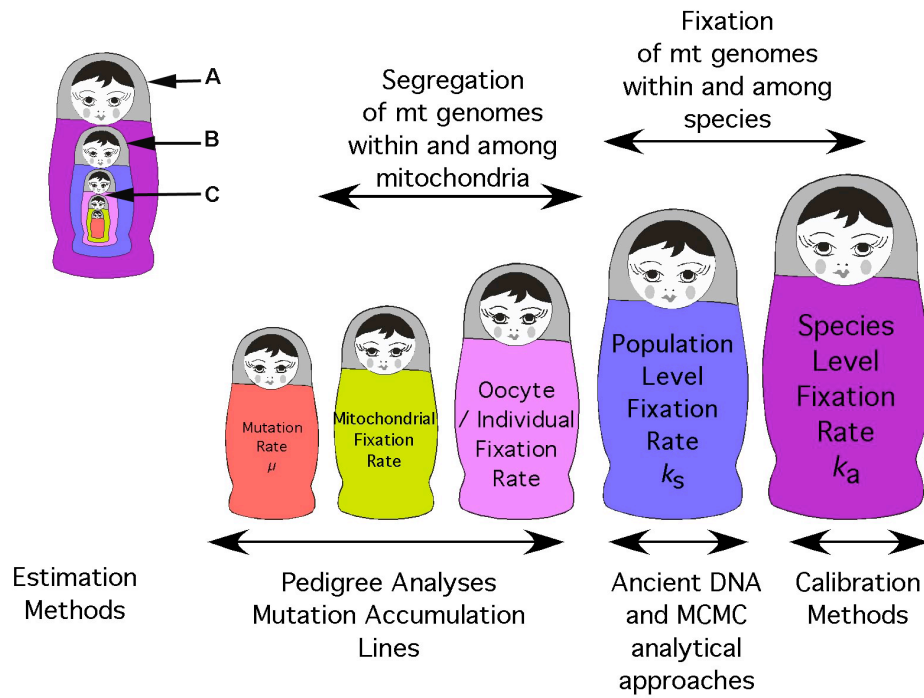
Figure 1.1: The distribution of mitochondrial genomes within mitochondria, cells (egg or oocyte), individuals and populations. Each mitochondrion contains varying numbers of mitochondrial genomes, in turn, every egg cell contains multiple mitochondria and each individual contains many mitochondrial, some of which are passed to the next generation. Various methods for inferring the molecular evolutionary rates utilize variation at each of these levels of diversity. **A.** Phylogenetic/calibration methods employ an indirect approach estimating the average rate of nucleotide substitution by comparing molecular sequence data from different species, calibrated against the time of the most common ancestor according to fossil materials or geological data. **B.** Ancient DNA methods estimate the rate of nucleotide evolution within a population or species, using analysis of serially sampled individuals from different period of times. **C.** Pedigree analysis is a direct method for estimating the mutation rates observed over a known number of reproductive events. This method accesses variation at the mitochondrial and cellular level. This figure is drawn by Vivienne Ward from University of Auckland for the purpose of this study.

1.3 The Estimation of Evolutionary Rates in mtDNA

Estimating an accurate and unbiased rate of molecular evolution for the mitochondrial genome and its variability across the molecule and among species is important to many different fields in evolutionary biology. Knowledge of variation in mtDNA mutation rates provides some insights into the pattern of mtDNA variation within species and among animals (Bazin et al. 2006; Lynch et al. 2006; Nabholz et al. 2008 b).

There are various methods for inferring the evolutionary rates using DNA nucleotide sequences obtained from different organisms. These widely employed estimates

are typically indirect, being based on calibrating against the fossil records (Shields & Wilson 1987) or some biogeographical events (Waters et al. 2000) such as phylogenetic split between human and a closely related primate (chimpanzee), or the timing of a particular human settlement, such as that of New Guinea (Kumar et al. 2005; Patterson et al. 2006; Wall 2003; Hasegawa et al. 1985; Jones 1979). In addition to this phylogenetic approach, some direct methods such as pedigree and ancient DNA method have been recently developed (Millar et al. 2008 b; Denver et al. 2000; Howell et al. 2003; Santos et al. 2005; Lambert et al. 2002; Shapiro et al. 2004).

Using a phylogenetic or calibration approach, an average rate of nucleotide substitution is typically calculated by measuring the sequence difference among living taxa and calibrating with the age of the most common ancestor according to fossil materials (Shields & Wilson 1987; Lovette 2004; Pereira & Baker 2006 ab; Päckert et al. 2007) or biogeographic data (Marko 2002). Shields and Wilson (1987) implemented this method in order to estimate the rate of mtDNA evolution in birds, using two genera of geese. They proposed a phylogeny of two species from the genus *Branta* (Branta and Canada geese) and three species from the genus *Anser* (the white-fronted goose, Ross´ goose and the snow goose). They calculate the sequence divergence between species of *Anser* and *Branta* at about 9% at the midpoint root of the phylogenetic tree. The common ancestor of these two geese genera was dated approximately 4 - 5 million years (Myr) according to the fossil records. Dividing 9% sequence divergence by 4.5 Myr divergence time gave a mean substitution rate of approximately 0.02 substitutions per site per million years (s/s/Myr) for the mitochondrial genome (Figure 1.2).

A subsequent study by Quinn et al. 1991 showed that there is a sequence difference of 0.013 between the two *Branta* subspecies based on mtDNA RFLP data. In contrast, the corrected sequence difference in part of the HVR-I between two subspecies of *snow* goose has been estimated at 0.135 (Quinn 1992). From the studies above Quinn (1992) concluded that control region evolves 10.4 times faster than the rest of the mitochondrial genome according to these calculation (0.135/0.013)=10.4 (Quinn 1992). This rate difference then multiplied by the Shields and Wilson (1987) rate of substitution (10.4 x 0.02) and obtained a rate of 0.208 (s/s/Myr) for the control region.
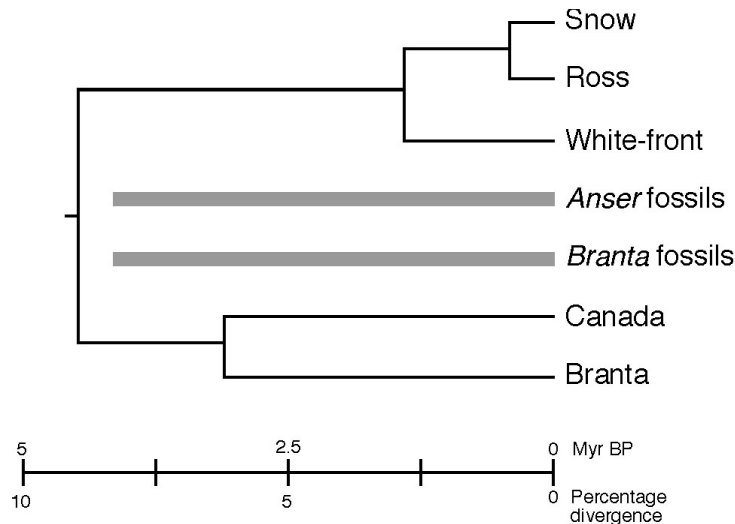
Figure 1.2: Phylogenetic tree and divergence time scale from two geese genera. The grayed lines represent the oldest fossil from each genus. The scale line indicates the time scale (Myr) and the divergence percentage (Shields & Wilson 1987). The figure is re-drawn by the Author of the thesis.

Wenink et al. (1996) extended this conclusion in order to calculate the substitution rate for the HVR-II region of mitochondrial genome in Dunlin and Turnstone (two shorebird species). The comparison between the rate of nucleotide changes in the HVR-I and HVR-II regions indicates that the rate for the HVR-II is 2.36 times slower and in absolute terms is 0.088 (0.208/2.36) (s/s/Myr). Wenink et al. (1996) calculated the substitution rate for the entire control region including both HVR-I and II at 0.148 (s/s/Myr) according to the following calculation: [0.208 + (0.208/2.36)]/2 (Wenink et al. 1996). Furthermore, the substitution rate of 0.02 (s/s/Myr) was obtained in a number of other molecular studies of birds (Randi 1996; Fleischer et al. 1998). Interestingly this rate is similar to the value commonly accepted for mammals (Brown et al. 1979; Wilson et al. 1985).

Since the evolutionary rates vary across organisms many studies have used phylogenetic methods for estimating the molecular rates and divergence times using the relaxed molecular clock model, which allows for various degrees of nucleotide changes throughout the tree (Takezaki et al. 1995; Cooper & Penny 1997; Sanderson 1997, 2002; Rambaut & Bromham 1998; Thorne et al. 1998; Cutler 2000; Huelsenbeck et al. 2000; Aris-Brosou & Yang 2002; Thorne & Kishino 2002). Although the number of these studies is increasing, the accuracy of using the relaxed molecular clock is still debatable (Ho

et al. 2005 b; Peterson 2006). Moreover, phylogenetic methods cannot be used for estimating the rates of molecular evolution within species (intra-specific) because of the absence of calibration points within a short timescale. In addition, it is likely that this rate that is derived from inter-specific comparisons produce overestimate of the age of intra-specific divergence events (Ho et al. 2005 b, Ho & Larson 2006; Peterson 2006; Lambert et al. 2002). Conversely, it is inappropriate to use intra-specific evolutionary rates to study molecular evolution between species, as performed by Zink (2002).

Several studies have shown a higher rate of molecular evolution within species in comparison to the rates obtained using inter-species comparisons (Lambert et al. 2002; Denver et al. 2000; Howell et al. 2003; Santos et al. 2005), using the ancient and pedigree DNA methods. Pedigree analysis is a direct approach to the estimate of mutation rates, which is based on observed changes in DNA sequences within known families. In most of the studies mutation rates based on pedigree analyses greatly exceed the evolutionary rate inferred by phylogenetic approach. Phylogenetic estimates of evolutionary rates in control region sequences of human mDNA range from 0.02 - 0.26 (s/s/Myr) (Vigilant et al. 1991; Horai et al. 1995), while pedigree analyses estimated this rate about 2.5 (s/s/Myr) (Howell et al. 1996; Parsons et al. 1997; Santos et al. 2005). Thus, according to the pedigree approach, sequence divergence rates in the human mtDNA control region are about 10 times higher than rates derived by indirect phylogenetic analyses (Howell et al. 1996; Parsons et al. 1997). This discrepancy can be a result of mutational hotspots in control region of mtDNA or simply a result of mitochondrial disease state of the individuals used for these studies (Jazin et al. 1998). In addition, Denver et al. (2000) performed a direct estimation of the mutation rate by sequencing 10,428 base pairs (bp) of the mtDNA of 74 *Caenorhabditis elegans* mutation accumulation lines over 214 generations. The lines followed single-progeny descent and the reported mutation rate is two times higher than the indirect estimate previously performed (Denver et al. 2000). Sigurdardo´ttir et al. (2000) estimated the mutation rate in control region of human mtDNA, using the pedigree analysis. They reported a higher rate compared with the phylogenetic estimates, although the authors used very similar approaches for pedigree and phylogenetic analyses (Sigurdardo´ttir et al. 2000).

Different rates have been reported by various studies but a consistent rate has not

yet emerged, perhaps because different studies have analysed a relatively small number of transmission events (Howell et al. 2003). The question is if pedigree rates are different from rates derived by phylogenetic analysis, what is the cause for this discrepancy? This discrepancy is unlikely to be caused by a single factor or evolutionary process. There are several proposed factors such as mutational hotspots (Macaulay et al. 1997), random genetic drift and selection (Parson et al. 1997), accumulated mutations in mtDNA of old sampled individuals (Michikawa et al. 1999) and finally the recent theory of time dependency for molecular rates (Ho et al. 2005 b), which has been specifically formulated to explain the discrepancy between pedigree and phylogenetic rates.

Ancient DNA technology provides an opportunity for evolutionary biologists to directly estimate evolutionary rates by analysis of individuals belonging to different periods of time (Lambert et al. 2002; Green et al. 2006; Shapiro et al. 2004; Edwards et al. 2007, Saarma et al. 2007). However, the main problem in using ancient DNA methods for estimating the evolutionary rates is that there are typically small numbers of preserved samples of known ages and geographical distribution available. In 2002, Lambert and his colleagues employed the ancient DNA technology for estimating the evolutionary rates of the mtDNA HVR-I in Adélie penguins (*Pygoscelis adeliae*) over the significant geological time frame (Lambert et al. 2002). By using the Markov chain Monte Carlo integration and a least-squares regression they estimated an evolutionary rate approximately two to seven times higher (0.96 s/s/Myr) than the rates previously estimated using an indirect phylogenetical approach. The high evolutionary rate reported by Lambert and his colleagues is in general agreement with the rate estimated by Parson and his coworkers for HVR-I in humans (Parsons et al. 1997). The substitution rates obtained from ancient DNA analysis have typically been high, falling between the rates estimated from pedigrees and phylogenetic analysis (Ho et al. 2007 b).

To resolve this contradiction, Millar et al. (2008 b) argued that in order to compare molecular rates accurately over different time scales it is required to specifically estimate the rates over a short and long period of time for each species or group of species. Therefore, the extant natural populations of the species from which a large number of pedigree samples can be collected along with a large numbers of well preserved ancient samples of

the same species are required. Adélie penguins represent the ideal model for estimating the molecular rates over different evolutionary time periods.
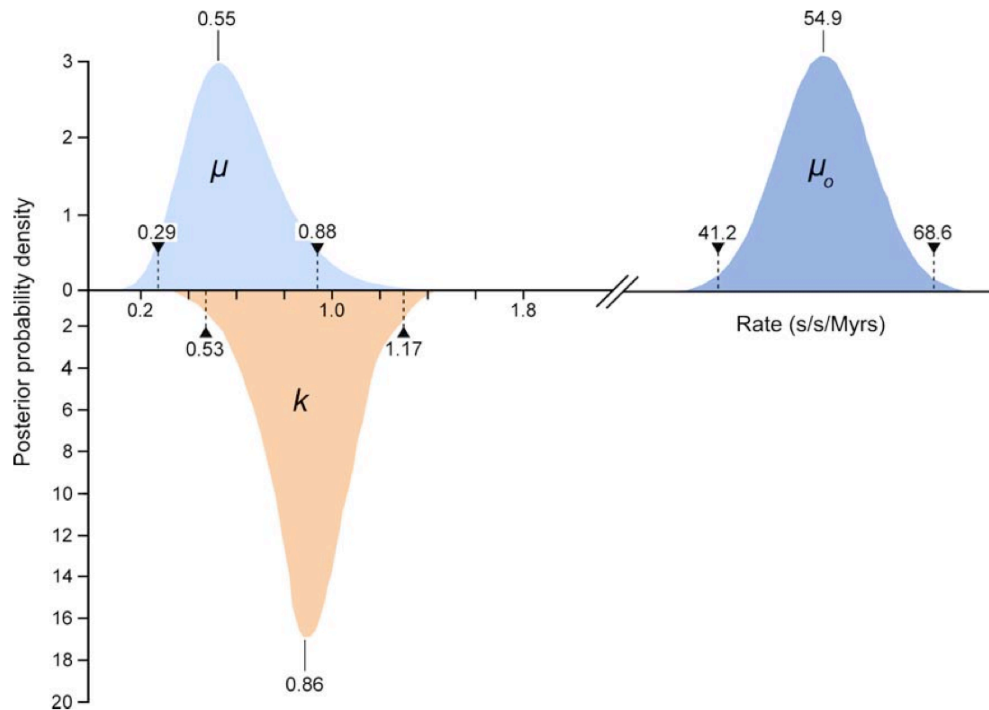
Figure 1.3: The observed rate of heteroplasmies (μ_0) of the mtDNA HVR-I region of Adélie penguin. The mutation rate (μ) estimated from pedigree data with modeling the inter-generational persistence of the heteroplasmies and the evolutionary rate (k) estimated from ancient data are shown. The mean and 95% confidence intervals are given (Millar et al. 2008 b). [For reprint this figure no permission is required under the creative commons license (CCAL) of PLOS Genetics].

They compared the molecular rates of HVR-I of mtDNA of Adélie penguins using both the pedigree and ancient DNA data (Millar et al. 2008 b). The pedigree data consisted of 344 bp of mtDNA HVR-I of 508 families of Adélie penguins (both parents) with 915 chicks. They detected 62 germ line heteroplasmies in both mothers and their offspring, which is consistence with a maternal inheritance of mtDNA transmission. Using a recently reported mathematical model (Hendy et al. 2009) and pedigree approach, they calculated the mutation rate (μ) at 0.55 s/s/Myr (HPD: 95%, CI: 0.29-0.88). In comparison, the substitution rate (k) of the HVR-I region was estimated about 0.86 (s/s/Myr) (HPD: 95%, CI: 0.53-1.17) based on ancient DNA sequences obtained from

162 known age sub-fossils of Adélie penguins (Figure 1.3). These data show that the mutation rates and substitution rates are not significantly different and suggest the HVR-I region of mtDNA is evolving neutrally in this species. Therefore, the hypothesis of time dependency of molecular rates cannot be supported with these data. In addition, Millar et al. (2008 b) explained that previous studies have treated heteroplasmies differently. Some studies have included the heteroplasmies in the estimation of mutation rates (Parsons et al. 1997) while some have not (Sigurğardóttir et al. 2000). The assumption of full contribution of each heteroplasmy to the mutation rates (Howell et al. 2003; Santos et al. 2005) can lead to an overestimation of the actual mutation rates (Millar et al. 2008 b).

1.4 Rapid Molecular Evolution in a Living Fossil

In 2008, Hay and her colleagues estimated the evolutionary rates for tuatara mtDNA HVR I-II using an ancient DNA approach (Hay et al. 2008). Tuatara (*Sphenodon*) is an enigmatic New Zealand reptile that coexisted with dinosaurs and has changed little morphologically from its Cretaceous relatives. Tuatara have a long generation time (sexual maturity at 10 - 15 years), low body temperatures, slow rates of growth (~50 cm in 35 years) and a slow reproductive rate (at 2 - 5 years intervals) (Cree et al. 1990; Cree et al. 1991; Thompson & Daugherty 1998). It has been argued that generation time, metabolic rate, body temperature and body size (Gillooly et al. 2005; Martin & Palumbi 1993) modulate the rate of neutral molecular evolution. In tuatara, the rare combination of the above features suggests a slow rate of molecular evolution in mtDNA.

In contrast to this notion, Hay et al. (2008) reported a high rate of molecular evolution for tuatara HVR (1.56 s/s/Myr, HPD: 95% CI: 0.83-2.34) compared with other vertebrates (Figure 1.4). In this study they calculated the rate by using nucleotide sequences from some known-age ancient tuatara and some modern individuals using Baysian statistics based on a Markov chain Monte Carlo (MCMC) approach implemented in the software BEAST (Drummond & Rambaut 2007). This study was the first attempt to directly estimate the molecular rate in a living fossil and quantify the neutral evolutionary changes in a species with such unique physiological and life history traits.

1.5 The Evolutionary Rate of Tuatara Revisited

Subsequently, Miller et al. (2008) argued that the serially sampled dataset used in the Hay et al. (2008) study was not large enough to estimate an accurate rate of evolution. Consequently they suggested that the rate was an overestimate as a result of not incorporating the population structure into the analysis by Hay et al. (2008). They argued that the high molecular rate reported might be a result of low genetic variability (~2%) in the Hay et al. (2008) dataset. Subramanian et al. (2009 b) responded that the nucleotide diversity of other species used for this type of study ranged from 0.2 - 6.2%. Since the diversity of the tuatara dataset was approximately 2%, this value fits within the former range and the nucleotide diversity of only three species (horse, bison and brown bear) significantly exceeding the one for tuatara (P > 0.05) (Figure 1.5). The evolutionary rates for these species (< 0.4 s/s/Myr) are lower than that estimated for tuatara (1.56 s/s/Myr). Thus the reported low nucleotide diversity did not bias the Baysian analysis of molecular rates as suggested by Miller et al. (2008).

Miller et al. (2008) also reanalysed the ancient dataset presented by Hay et al. (2008) and estimated a molecular rate for tuatara of 0.076 s/s/Myr (HPD: 0.0016-0.32). This rate is much lower than the rate previously estimated by Hay et al. (2008). They performed three runs of randomization tests by randomising the ages of ancient samples. They estimated the rate slightly lower than 1.56 s/s/Myr (HPD: 95%, CI: 0.83-2.34), but not significantly different. According to these results they concluded that the dataset for Hay et al. (2008) does not contain the sufficient information for calculating precise and reliable molecular rates. However, the approach used by Miller et al. (2008) is problematic because modern samples were not used in the randomisation test (Subramanian et al. 2009 b). Since the modern samples represent a considerable portion of the full dataset (55% or 41 out of 74 sequences), by eliminating them from the analysis the resulting rate is unlikely to be different from that estimated using the true ages of all samples. Hence, Subramanian et al. (2009 b) repeated the Miller et al. (2008) analysis using a complete randomization of both ancient and modern dataset. The resulting rate was much lower than the rate of 1.56 s/s/Myr, thereby indicating that there is an evolutionary signal in Hay et al. (2008) dataset.

Subramanian et al. (2009 b) also conducted a simulation study both with randomising the ages of ancient simulated DNA sequences and modern ones. The result of

randomising of only ancient simulated sequences gave the same pattern as Miller et al. (2008). As expected, by incorporating both ancient and modern simulated sequences, the estimate rate was much less than 1.56 s/s/Myr, again indicating strong support of the original rate estimate in Hay et al. (2008) (Figure 1.6).
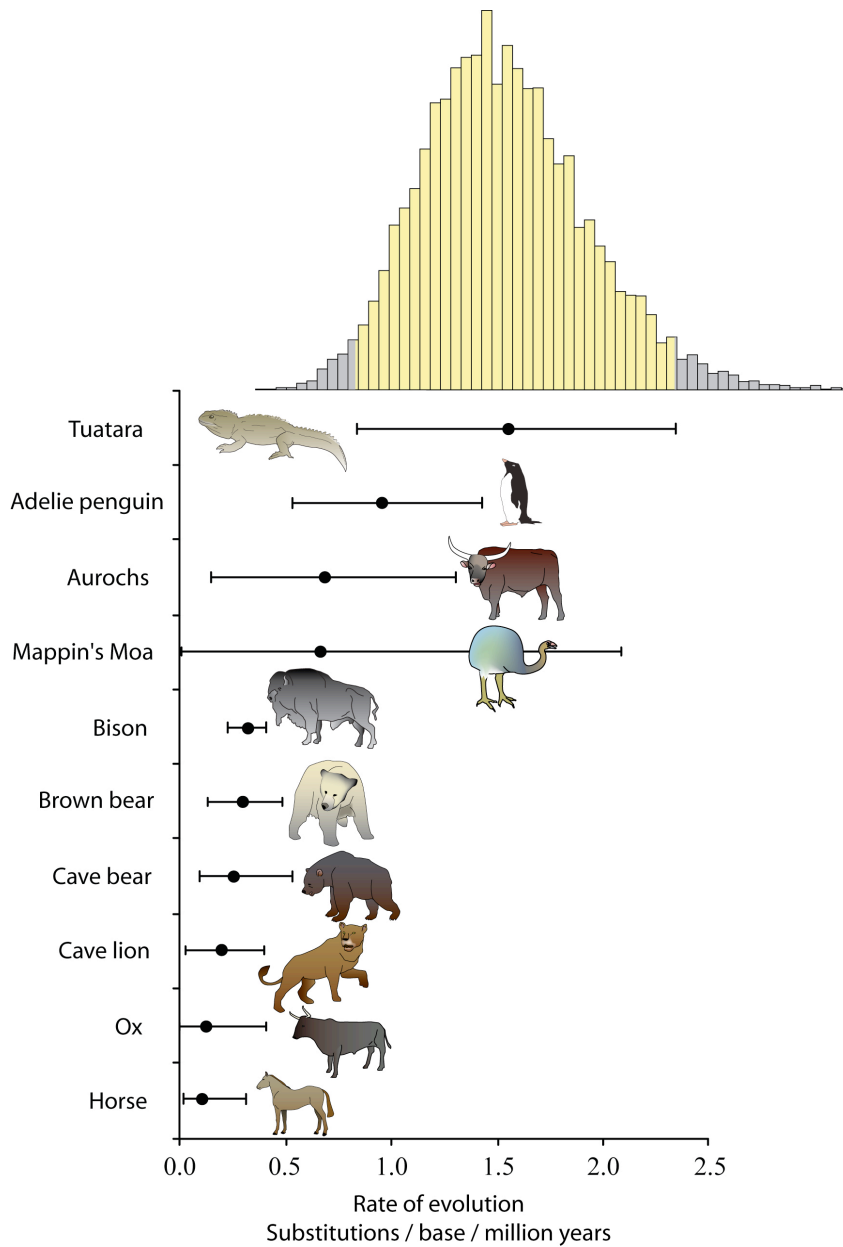
Figure 1.4: A comparison of the evolutionary rates of HVR regions of tuatara mtDNA with other vertebrates (Hay et al. 2008). The rates for nine other species were obtained from previous studies that used ancient DNA approach (Lambert et al. 2002; Edwards et al. 2007; Shapiro et al. 2004; Saarma et al. 2007; Ho et al. 2007 a). [For reprint this figure, no permission is required from Elsevier as part of author's rights].
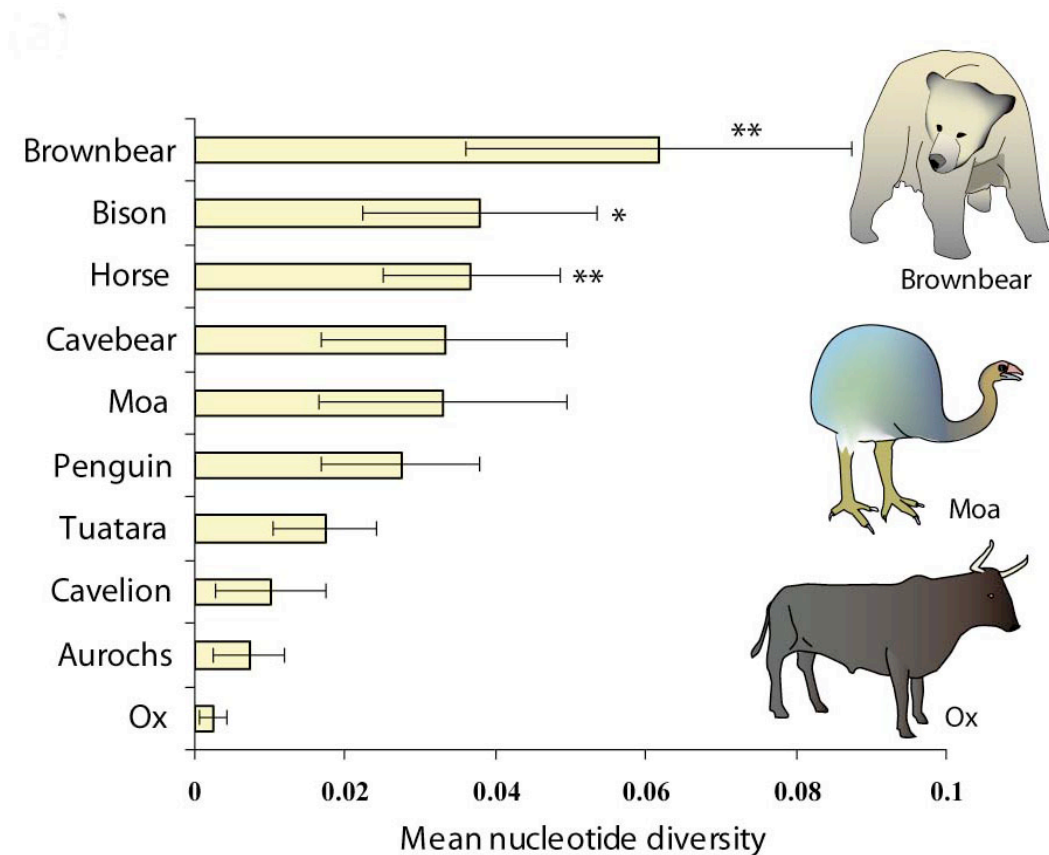
Figure 1.5: Nucleotide diversity in brown bear, bison and horse are significantly greater than tuatara (*P<0.05 and **P< 0.01) (Subramanian et al .2009 b). [For reprint this figure, no permission is required from Elsevier as part of author's rights].

Miller et al. (2008) also suggested that perhaps the high rate of evolution in tuatara HVR is a result of highly genetically structured populations in tuatara. They suggested that there is evidence of such structure in tuatara, based on microsatellites data (MacAvoy et al. 2006). They indicated that offshore island (modern) tuatara populations are genetically structured between and within islands (MacAvoy et al. 2006). In reply, Subramanian et al. (2009 b) suggested that microsatellites typically show much more variation and reflect a different time and geographic scale to that of mitochondrial DNA. Furthermore, although the ancient and modern samples from the Hay et al. (2008) study are geographically distinct they are not genetically differentiated in terms of mitochondrial DNA and the ancient and modern samples are mixed together in phylogenetic network analysis (Figure 1.7) (Subramanian et al. 2009 b).
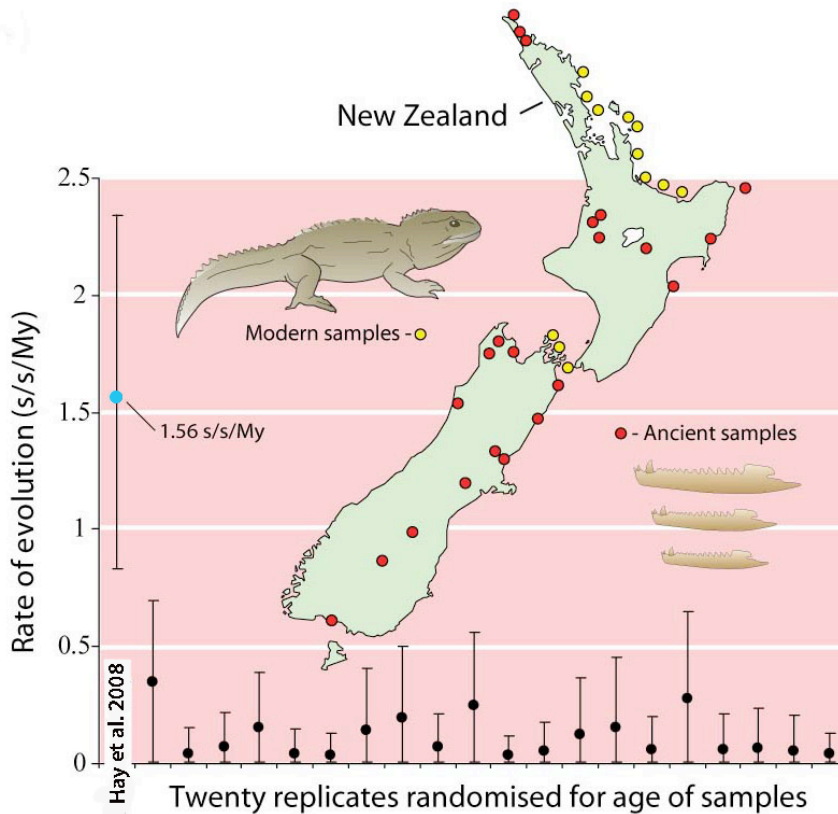
Figure 1.6: Estimated evolutionary rates for the tuatara HVR region calculated by Subramanian et al. (2009 b). The blue dot indicates the point estimate (1.56 s/s/Myr) and the 95% HPD are also given. The result of 20 replicates that the age of all samples including modern and ancient was randomized and the 95% HPDs are shown in the x-axis. The geographical distribution of the ancient (red dots) and modern samples (yellow dots) is depicted in the map of New Zealand. Location of modern and sub-fossil sites and the sample details of tuatara ancient bones used in this study are given in Appendix D, Subramanian et al. (2009 b), supplementary. [For reprint this figure, no permission is required from Elsevier as part of author's rights].

Another criticism of the high evolutionary rates estimated for tuatara mtDNA is that the Baysian analysis assumes either a constant population size or exponential growth while, tuatara populations declined dramatically after human arrival in New Zealand approximately 730-800 years ago (Duncan et al. 2002; Crook 1973; Gaze 2001; King 2003). Although this claim by Miller et al. (2008) is correct, it is not applicable to this study since most of the samples, except three of them, are older than this event.

It is worth noting that some methods like BEAST (Drummond & Rambaut 2007) that are currently used for estimating evolutionary rates are based on simple models that do not take parameters such as DNA damage, migration, bottlenecks and population

subdivision into account. Therefore, the future evolutionary rate for vertebrates using more complicated methods might be different from the rates currently reported. Nevertheless the relatively high evolutionary rates for tuatara are likely to remain high compared to other vertebrates. One explanation for the high rate of neutral evolution in tuatara mtDNA, despite its stable morphology, is that different biological processes underlie molecular and morphological evolution. Allan Wilson originally proposed this theory in 1975 (King & Wilson 1975).
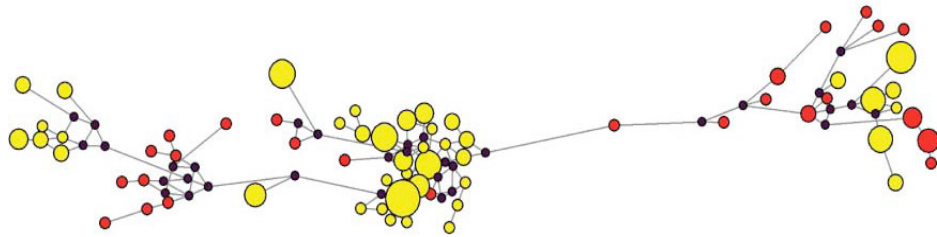
Figure 1.7: The median joining network for ancient, modern and some hypothetical haplotypes for tuatara. Ancient sequences are represented in red, modern in yellow and hypothetical intermediate haplotypes in black. The length of the branches indicates the number of the nucleotide differences and the size of the circles are representing the portion of the samples (Subramanian et al. 2009 b). [For reprint this figure, no permission is required from Elsevier as part of author's rights].

1.6 Time Dependency of the Molecular Rate Estimates

As reviewed above, recent studies show a wide range of rate estimates among genes and species based on different methods (Ho et al. 2005 b; Nei 2005). The evolutionary rates based on the phylogenetic (species-level) studies are the slowest and the rates based on the pedigree (population-level) studies typically are faster (Ho et al. 2005 b). This discrepancy in molecular evolutionary rates is a matter of controversy and there is a debate over which rates are relevant to particular studies.

Some authors have suggested a theory of time dependency of molecular rates and claimed that the rate of mutation (rate estimated over short period of time) and the rate of substitution (rate estimated over long period of time) are different among genes, among taxa and among different sites in DNA sequences (Ho et al. 2005 b; Penny 2005). This theory was first presented by Wayne et al. (1991) who suggested that, based on a study of carnivores and primates, substitution rate estimates decrease linearly with increasing time

(Wayne et al. 1991). In 2004, a similar pattern in birds was discovered by Garcia-Moreno, but the decline was exponential rather than linear (Garcia-Moreno 2004). Ho et al. (2005 b) observed some effects of time-dependency on molecular rates in three data sets including two protein-coding genes (avian and primate) and one mtDNA control region in higher primates. They calculated the higher molecular rate between generations (pedigree data) and lower rate for the local and widespread populations. This theory is based on the relationship between the calibration time and the rate of changes, which could be described by a vertically translated exponential decay curve, the so-called 'lazy jay' curve (Ho & Larson 2006; Penny 2005) (Figure 1.8). Ho et al. (2005 b) suggested that it takes 1 - 2 (Myr) for the mutation rate to reduce to a constant level of substitution rate (Ho et al. 2005 b). The authors removed the effects of any sequencing errors, calibration errors and the saturation of mutation rates for fast evolving positions. Using a rate curve estimated from the D-loop data, Ho et al. (2005 b) showed that the timing of many recent evolutionary events such as divergence of human and Neanderthal has been overestimated by previous studies and they should be reconsidered.
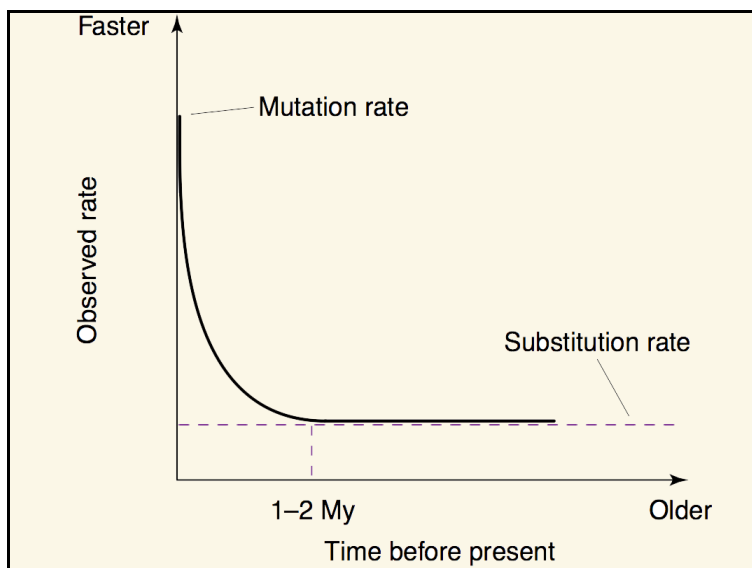
Figure 1.8: The theoretically expected decline in molecular rates with increasing time over which they have been estimated- namely the time dependency of the molecular evolutionary rates. The Mutation rates continuously decrease and in 1-2 (Myr) reach to a plateau similar to molecular evolutionary rates estimated over long timescales (Ho & Larson 2006). The figure is re-drawn by the Author of the thesis.

Penny (2005) has suggested that the short-term acceleration in molecular evolution (when time is thought to go backwards) is the result of an increasing persistence time of mutations in populations from deleterious mutations to neutral ones (Penny 2005).

1.7 Resolving the Conundrum

There are a number of possible reasons for the apparent difference in estimated times of various evolutionary events. Firstly, the estimation of inter-species divergence based on the neutral sites (synonymous positions) and in the HVR regions is most likely to be affected by the occurrence of multiple nucleotide substitution at the same position (saturation) and consequently the underestimation of molecular rates over a long timescale (Heyer et al. 2001; Nabholz et al. 2008 a; Soares et al. 2009; Subramanian et al. 2009 a). Secondly, the presence of short-lived slightly deleterious mutations among intra-species populations that could lead to the overestimation of the molecular rates over short timescales (Soares et al. 2009; Kivisild et al. 2006; Endicott & Ho 2008; Subramanian et al. 2009 a; Penny 2005). Although estimating evolutionary rates based on intra-species comparison is not affected by the nucleotide substitution saturation, it is influenced by the presence of short-lived deleterious mutations. On the other hand, phylogenetic approaches are not suitable for estimating the molecular rates within species as a result of lack of calibration points within a short timescale.

Although previous studies (Ho et al. 2005 ab; Ho & Larson 2006; Penny 2005, Millar et al. 2008 b) proposed some interesting theories and suggestions in order to increase the accuracy and decrease the level of discrepancy among different rates reported, the concept of molecular rate variation with time is still in debate. In order to resolve this conundrum we need more comprehensive molecular data from the complete mtDNA of both ancient and modern populations of each species. Perhaps, analysing the microsatellites with their great variability in a short time scale provides more potential to reveal the evolutionary events over narrower time scales (Shepherd et al. 2005; Harper et al. 2006; Martinez-Cruz et al. 2007). The recent suggestions that mtDNA is not a powerful marker for demography studies (Bazin et al. 2006), might suggest that nuclear DNA would be an alternative marker to reveal the evolutionary histories over time. The next generation sequencing technology provides great promise in this direction with the opportunity to

generate more genetic data for various species with a great sequencing coverage of the genome. To exploit the full potential of these new massive data sets, it is important to conduct evolutionary analyses using new and rigorous bioinformatic approaches. There are various sophisticated and novel statistical approaches and software available for the joint analysis of ancient and modern DNA (Navascues et al. 2010). This is the responsibility of the researchers to employ the appropriate approach for analyzing the data in order to prevent the misleading results.

1.8 The Rationale and Importance of this Project

Previous studies have contributed a great deal to our understanding of the important issue of molecular evolutionary rates. However, this research project is the first study to directly estimate the overall rate of molecular evolution for the complete mtDNA in tuatara; a species with extreme physiological and life history traits. The estimation of the evolutionary rates for different mitochondrial genes and determination of how these rates differ among genes can also make a significant contribution to our knowledge of the life history and biology of tuatara. Perhaps understanding the reasons for high molecular rates in tuatara will provide us with better tools for solving the controversy issue of neutral evolution.

Chapter Two

The Mitochondrial Genome: Structure, Maternal Inheritance and Mutations

2.1 Introduction

In comparison with the nuclear genome, the mitochondrial genome (mtDNA) comprises only a tiny fraction of a species' complete genome. This small molecule has been widely recognised as a powerful molecular tool in evolutionary biology for decades. More specifically, migration and population genetics studies (Harrison 1989; Avise et al. 1987; Brown 1980, 1985; Cann et al. 1987; Ingman et al. 2000; Johnson et al. 1983; Krings et al.1997; Melton et al. 1998; Redd & Stoneking 1999; Tambets et al. 2004; Torroni et al. 1993; Vigilant et al. 1991) have benefited from analyses of mtDNA sequences in many diverse ways. This is due to several unique characteristics of mtDNA including its high copy number, predominantly maternal inheritance, typically a lack of recombination, generally high mutation rates and often, high variability within and among species and even populations (Avise et al. 1979; Brown & Vinograd 1974).

There are also some difficulties associated with this molecule such as evidence of mtDNA recombination in some species (Eyre-Walker et al. 1999; Ballard & Whitlock 2004), adaptive evolution (Hurst & Jiggins 2005; Bazin et al. 2006) and mutational hot spot in some regions of the genome (Galtier et al. 2006). Understanding the processes that govern the evolution of mtDNA genome is crucial to the proper use of this marker in different types of evolutionary studies. Although many studies have been aimed at the estimation of molecular rates (both mutational and evolutionary) of mtDNA (Lambert et al. 2002; Green et al. 2006; Shapiro et al. 2004; Edwards et al. 2007, Saarma et al. 2007), the extent of the variation across species and lineages and the reason for mitochondrial genome hyper-mutability remain only partially understood.

2.2 Mitochondrial Genetics: the Basics

Mitochondria are found in all nucleated cells and are the principal generators of cellular energy via oxidative phosphorylation (OXPHOS), incorporating the electron transferring respiratory chain (complexes I-IV) and the ATP synthase (complex V). Mitochondria are the only known location of extra chromosomal DNA within the cell (except chloroplast in plants) and are under the genetic control by both nuclear and mitochondrial DNA. Mitochondria have a double layer membrane; the outer membrane separates the organelle from the cytosol whereas the enzyme complexes for oxidative phosphorylation are embedded in the inner layer (Hatefi 1985). The mtDNA comprises a small, circular double-stranded molecule of approximately 16,000 bp in length. The structure and gene organization of mtDNA is relatively conserved among animals (Wolstenholme 1992) as diverse as human (Attardi et al. 1976), mouse (Battey & Clayton 1978), rat (Kroon et al. 1977), frog (Dawid et al. 1976) and fruit flies (Dawid et al. 1976; Goddard & Wolstenholme 1978). The two strands that compose the mtDNA are different in base composition and therefore are referred to Heavy (H) and Light (L) strands. The difference in base composition (G+T) results in different buoyant densities in denaturing caesium chloride gradient (Kasamatsu et al. 1974).
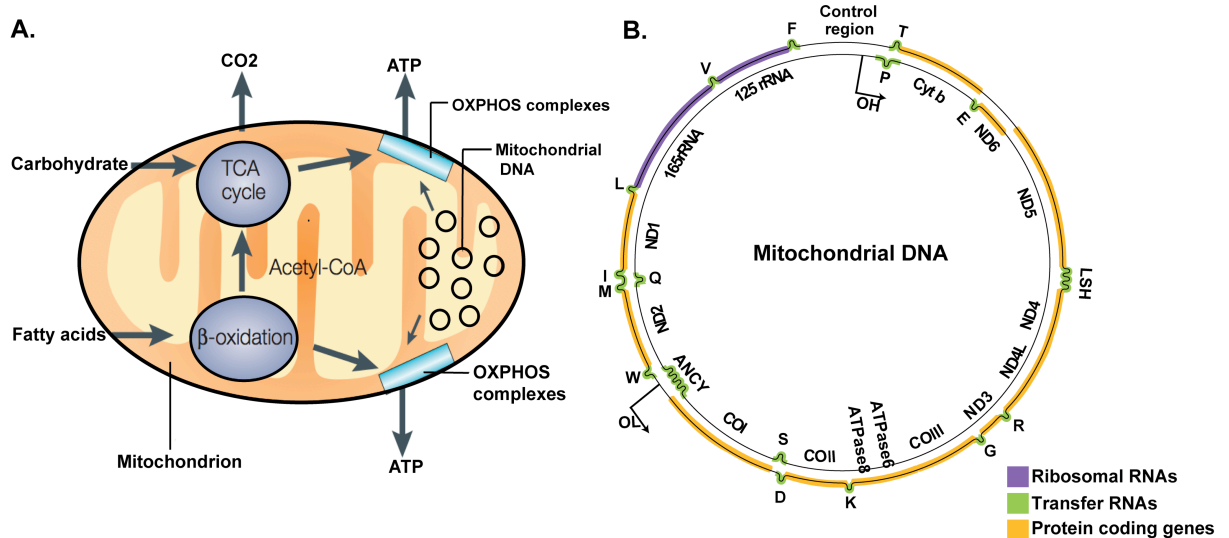
Figure 2.1: The role of mitochondrial genome in cellular energy production. **A.** mtDNA contributes to the enzymatic complexes of the OXPHOS system within the inner membrane of the mitochondrion which synthesize ATP as a source of cellular energy. **B.** mtDNA in vertebrates comprises 13 protein-coding genes, 2 rRNA (12sRNA, 16sRNA), 22 tRNAs and control region. mtDNA has two origins of replication (OH and OL). Reprinted by permission from Macmillan Publishers Ltd: [Nature Reviews Genetics](Taylor & Turnbull), copyright (2005).

In vertebrates, the mitochondrial genome encodes 13 essential polypeptides involved in OXPHOS system (Anderson et al. 1981; Macreadie et al.1983; Chomyn et al.1985; Chomyn et al. 1986), 2 ribosomal RNA (rRNA) and 22 transfer RNAs (tRNAs) for their translation within the organelle (Figure 2.1). The genes for 2 rRNAs, 14 tRNAs and 12 polypeptides are encoded by the heavy strand while the light strand is responsible for encoding 8 tRNAs and one polypeptide. In addition to the functional regions mtDNA contain a highly variable control region (CR), which has mainly regulatory function and has been particularly important in evolutionary studies. The CR consists of two hyper variable regions called Hyper Variable Regions HVR-I and HVR-II. The mitochondrial genome has an exceptionally economical organization, being very compact in size. Apart from one regulatory region, the majority of the genes lack introns and also the intergenic sequences are either absent or reduced down to couple of bases (Ojala et al. 1981). In addition to that, tRNAs and rRNAs molecules are unusually small and even some of the protein coding genes are overlapping (Wolstenholme 1992). The genetic codes on mtDNA deviate from the standard codes and even differences in codon usage have been documented in a number of species (Osawa et al. 1992). For example the codon TGA in

vertebrate mtDNA codes for tryptophan, rather than being a stop codon. Also, AGR (R: A or G) is a stop codon in the mtDNA genome of vertebrates, while it codes serine amino acid in mtDNA of echinoderms and argenine in yeast mtDNA.

Although the structure and gene organisation of mtDNA is highly conserved among various taxa, the mitochondrial genome of *Sphenodon* has a distinctive gene order compared to other vertebrates. Mitochondrial genome of *Sphenodon* has two copies for tRNA-Lys gene while the other genes such as ND5, tRNA-His and tRNA-Thr are absent (Rest et al. 2003). Moreover, this genome lacks the origin of replication in light strand and this is a common trait among all birds, crocodilians and *Sphenodons*. Another unique feature of this genome is the existence of two copies of control region with the size of 823 bp and 926 bp respectively. The sequences of these two D-loops are identical in each individual for the initial 750 bp in 5´ direction but can be different thereafter (Rest et al. 2003) (Figure 2.2).
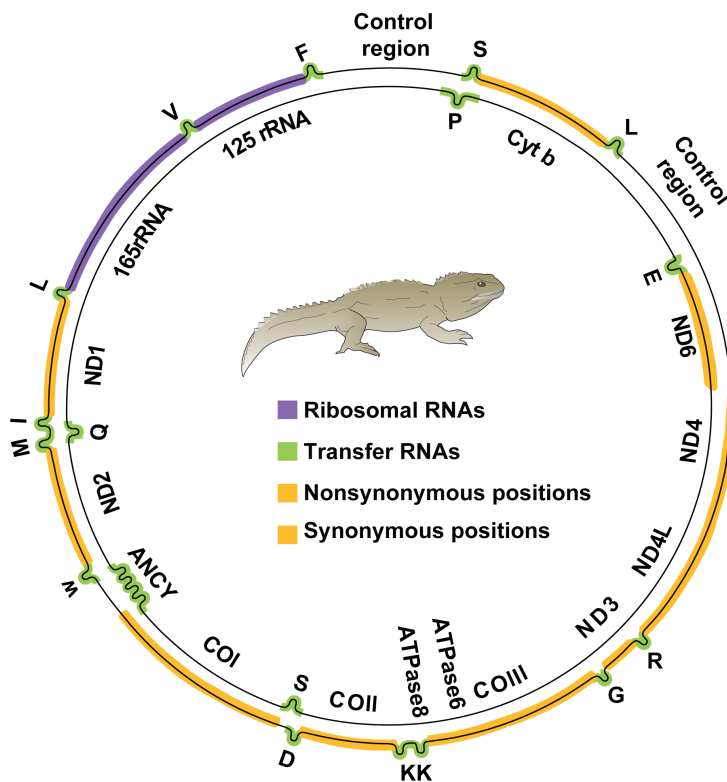
Figure 2.2: The gene structure and organization for mitochondrial DNA of tuatara (Rest et al. 2003). The position of two control regions, 21 tRNAs, showing by their single-letter amino acid code, two rRNAs and 12 protein coding genes are given. The figure is drawn by the Author of the thesis.

2.3 High Copy Number of Mitochondrial DNA

Mitochondrial DNA is present in high copy numbers in cells. The average somatic cell contains two copies of nuclear DNA (nDNA) in nucleus but it has hundreds to thousands copies of mtDNA (Robin & Wong 1988; Michaels et al. 1982; Shuster et al. 1988; Wiesner et al. 1992) (Figure 2.3). The high copy number and the location of the mtDNA make the recovery of this molecule easier. This feature of mtDNA is particularly important in ancient DNA and molecular evolutionary studies.
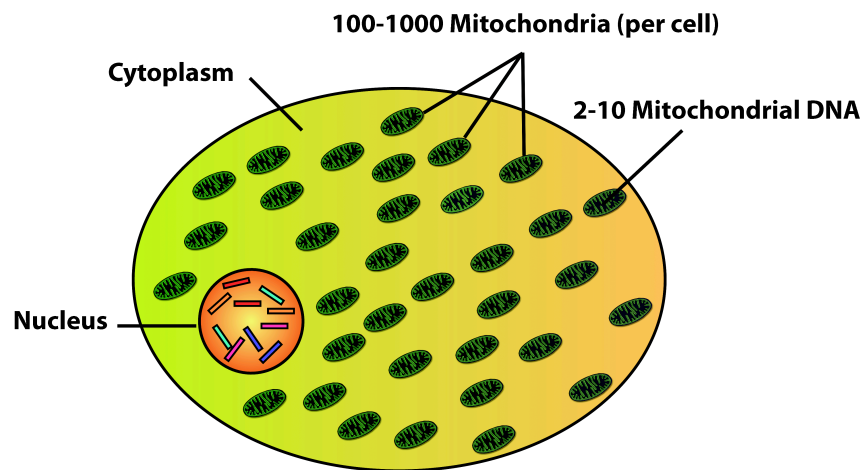
Figure 2.3: This schematic represents the relative number of nuclear DNA to mitochondrial DNA in a somatic cell. Each cell comprises two copies of nuclear DNA and 100 - 1000 copies of mitochondria. Each mitochondrion is estimated to contain 2 - 10 mtDNA copies. The figure is drawn by the Author of the thesis.

2.4 Inheritance of Mitochondrial DNA

The predominantly maternal inheritance of mtDNA in vertebrates is widely accepted as a standard model of mtDNA transmission (Stoneking 1993; Stoneking & Soodyall 1996; Wallace et al. 1999). However, a number of studies have challenged this model by showing the paternal inheritance of mtDNA in species such as *Drosophila* (Kondo et al. 1990), honeybees (Meusel & Moritz 1993), sheep (Zhao et al. 2004), mouse (Gyllensten et al. 1991; Kaneda et al. 1995), hybrid bird (Kvist et al. 2003), mussels (Zouros et al. 1992) and human (Schwartz & Vissing 2002).

The enzymatic destruction or elimination of sperm mtDNA in the oocytes could be explained by different mechanism (Lima-de-Faria 1983) (Figure 2.4). The lack of paternal mitochondrial genes in sperm has been documented in cray-fish (Moses 1961). In some cases such as ascidians the nucleus of the sperm enters the egg while the rest remains outside (Ursprung & Schabtach 1965). Another mechanism, which has described in sea urchin, is the destruction of the sperm mitochondria by the egg cell after fertilisation, via oxidative damage and proteolytic processes (Anderson 1968). Maternal inheritance of mtDNA could be simply a result of dilution of paternal mtDNA in the oocyte. This occurs as a result of high copy number of mitochondria in oocytes ($>10^5$) (Michaels et al. 1982; Piko & Matsumoto 1976), compared to sperm (50 - 75) (Hecht et al. 1984). Therefore, even if fertilisation involves a complete mixing of paternal and maternal mtDNA, the contribution of paternal mtDNA to the zygote's pool would be relatively small. However, most evidence suggests that the paternal contribution is almost zero and mtDNA is exclusively maternally inherited. Therefore, mtDNA has one-quarter of the effective population size of a nuclear gene.

This uni-parental mode of inheritance of mtDNA enables researchers to trace a maternal lineage far back in time without the confounding effect of paternal inheritance of mtDNA and the nuDNA recombination (Pakendorf & Stoneking 2005). However, the paternal inheritance of mtDNA has been reported from a patient suffering from mitochondrial myopathy (Schwartz & Vissing 2002). After this report the strict maternal inheritance of mtDNA was generally called into the question and this research opened the way to more studies about this feature of mtDNA (Bromham et al. 2003). Further studies have not shown any trace of paternal mtDNA in other patients suffering from the same disease (Filosto et al. 2003; Schwartz & Vissing 2004; Taylor et al. 2003) and the paternal inheritance of mtDNA remains a rare phenomenon in human populations (Pakendorf & Stoneking 2005). Therefore, the assumption of maternal inheritance of mtDNA can still be regarded as a general rule in most animal species.
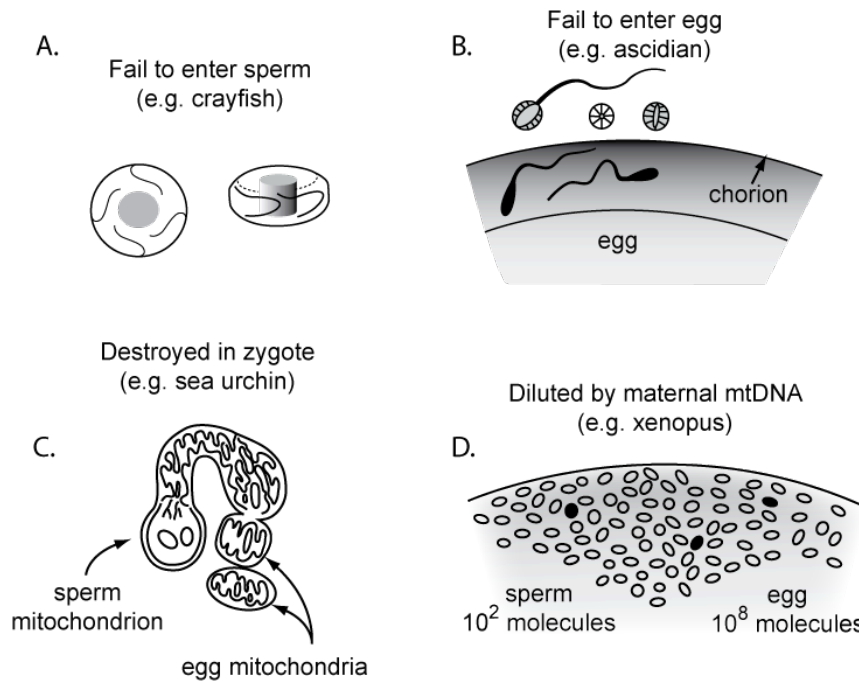
Figure 2.4: Some mechanisms for uniparental inheritance, with examples from animal kingdom, of the maternal inheritance of mitochondrial genes **A.** Peternal mitochondrial genes may be missing from sperm (Moses 1961), **B.** Fail to enter the egg (Ursprung & Schabtach 1965), **C.** be enzymatically destroyed in the egg (Anderson 1968) or, **D.** simply be outnumbered by maternal genes (Michaels et al. 1982; Piko & Matsumoto 1976). The figure is drawn by Vivienne Ward from University of Auckland.

2.5 Heteroplasmy

The presence of several thousand copies of mtDNA per cell, together with generally high mutation rates in this molecule increases the possibility of observing multiple mtDNA genotypes within a cell, a tissue, or an individual. This condition is known as heteroplasmy. Although there is a substantial bottleneck in the early oogenesis that restricts the number of mtDNA variants that can pass from mother to offspring, the actual size of this bottleneck depends on distribution of mitochondrial genotypes within and among mitochondria. Since the concentration of mutant variants is generally low, the majority of the mtDNA copies in each individual are likely to be identical (homoplasmy) (Bendall et al. 1996; Brown et al. 2001; Poulton & Marchington 2002; Hauswirth & Laipis 1982; Olivo et al. 1983; Koehler et al. 1991; Jenuth et al. 1996). Heteroplasmy is of importance in evolutionary studies and forensic applications as it is direct evidence for the generation of new mtDNA variants within a species (McLeod & White 2009).

Heteroplasmy is often difficult to detect accurately since several mutant variants are

likely to coexist at any time in a cell or even within a single mitochondria because there are typically multiple genomes per mitochondria. Some of these variants will be at very low concentrations. In practice, with Sanger sequencing methods, heteroplasmies can be recognised only when the frequency of a particular mutant variant accumulates to a level that easily distinguishes it from background noise in electropherograms. Insensitive methods such as direct Sanger sequencing of the PCR products in most cases have not provided a reliable estimates of the occurrence of heteroplasmy since distinguishing real heteroplasmies from the background signals in electropherograms is difficult. Millar et al. (2008 b) suggested the use of the detection threshold of 23% for Sanger sequencing method in order to maximise accuracy and minimise the false positive heteroplasmy detections. They estimated the ratio of heteroplasmic variants in mothers and chicks from the relative peak heights in DNA electropherograms for each individual.

2.6 Mutation in the Mitochondrial Genome

Mitochondrial DNA is the most popular marker in molecular evolution and diversity in animals is largely because of its elevated mutation rate. Generally the mutation rate in mtDNA is an order of magnitude higher than the nuclear genome (Wilson et al. 1985; Brown et al. 1979) with an estimated rate of 0.017 (s/s/Myr) for the entire mitochondrial genome excluding the control region (Ingman et al. 2000). The observed mutation rate in the entire mtDNA is non-uniform, with a low rate in protein coding regions, and a relatively high rate in the control region. The mutation rate for RNA regions generally falls between these two rates. It is worth noting that within the control region the mutation rate is heterogeneous. Some mutational hotspots mutate faster than the average sites (Excoffier 2002; Heyer et al. 2001; Meyer et al. 1999; Stoneking & Soodyall 1996). The majority of mutations in mtDNA are nucleotide substitutions and deletions. Deletions are the cause of some maternally inherited diseases while many substitutions appear to be neutral. There are several clinical syndromes such as myopathy, epilepsy, heart failure, diabetes and thyroid disease that have a high probability of mtDNA involvement, although there is a long list of other potential genetic causes for this disease (Taylor & Turnbull 2005).

In principle there are two likely reasons for the high nucleotide substitution rate of mtDNA in comparison with nuDNA. It could be due to a high mutation rate or/and a high fixation of mutations. The high mutation rate is likely to be the result of an inefficient repair system or the absence of proofreading replication enzymes in mtDNA, in comparison to that in the nucleus (Shadel & Clayton 1997; Grossman et al. 1973). In addition, the absence of a functional enzyme responsible for the removal of the thymine dimmers from the mtDNA strands has been well documented (Clayton et al. 1974). Other DNA repair mechanisms in mtDNA are also inefficient, although are not totally lacking (Lansman & Clayton 1975). There are also some additional characteristics of mtDNA that likely influence the high rate of mutation. For instance, mtDNA contains a high concentration of mutagenic oxygen free radicals that is a consequence of its function in energy generation through OXPHOS (Shadel & Clayton 1997). Mitochondrial DNA may also have a higher turnover rate than nuclear DNA, requiring more replications per unit time. The high susceptibility of mtDNA to free oxygen radicals and the other damaging agents could be due to absence of DNA-binding proteins such as histones which protects nuclear DNA (Shadel & Clayton 1997). Another explanation for the high nucleotide substitution rate of mtDNA is the higher probability of fixation of nucleotide substitutions. This could arise from the low functional constraints on mitochondrial gene products. In mtDNA, a mutation that inactivates a copy of a gene has a minor effect on the fitness of the organism due to the existence of multiple copies of mitochondrial genes (Brown et al. 1979; Wilson et al. 1977). There is also variation in mtDNA mutability across species and lineages, which has been argued to be explained by a number of factors such as generation time, metabolic rate and longevity (Nabholz et al. 2008 a).

According to generation time hypothesis, the generation time and the rates of neutral substitutions are inversely related (Li et al. 1987; Ohta 1993; Mooers & Harvey 1994; Bromham et al. 1996). Species with longer generation time are likely to experience less DNA replication per year and therefore the smaller number of replication error per unit of time. However, this model only explains the mutations that result from replication error. This error rate is constant among different species and could also affect the nuclear genome.

The metabolic rate hypothesis (Martin et al. 1992; Martin & Palumbi 1993;

Martin 1995; Nunn & Stanley 1998) is a replication-independent theory proposing a direct relationship between metabolic rate and neutral mutation rates. The existence of mutagenic free radicals during mitochondrial respiration increases the probability of mutations occurring in mtDNA. It is proposed that the metabolic rate varies with body mass and temperature among different organisms and affects the nucleotide mutation rate (Gillooly et al. 2005).

The longevity hypothesis is proposing an inverse relation between longevity and mutation rate. This hypothesis makes a link between the production of mutagenic free radicals and aging of the organism. Long-lived organisms are hypothesized to hold a reduced mutation rate in comparison with short-lived organisms (Barja & Herrero 2000; Kujoth et al. 2007). However, it needs to be acknowledged that these theories do not make specific predictions about the rates of molecular evolution we would expect to find in the case of any particular animal species with a given longevity and metabolic rate.

2.7 Summary

Mutations in mtDNA are unusual in two respects. First, the rate of mutation in the mitochondrial genome is high with respect to that recorded for the nuclear DNA. Second, in the mitochondria, the biological processes that govern how mutations pass through one generation to the next is very different to those that govern inheritance of nuclear mutations. Mitochondrial DNA is exclusively maternally inherited in most species and therefore it has one-quarter of the effective population size of a nuclear gene. Mutations arise more frequently in mtDNA than in nDNA due to its replication system. During the replication, some regions of the mitochondrial genome spend more time in the vulnerable single-stranded form than nDNA. This makes these regions more susceptible to the free mutagenic radicals and results in a high frequency of mutations. The underlying biological processes that determine the high mutation rates observed in mtDNA are likely to apply across a broad range of species. Hence, we would predict similar mtDNA mutation rates among a wide spectrum of vertebrate species for example. A further consequence of this, and importantly, any study that attempts to estimate mtDNA mutation rates is potentially relevant to more than the single species being investigated.

Chapter Three

Tuatara (*Sphenodon*): Species Classification and Geographical Distribution

3.1 Introduction

Tuatara were originally classified as lizards, when John Gray, in 1831, on receiving a skull from New Zealand named it *Sphaenodon* (now *Sphenodon*) (Gray 1831). In 1867, Albert Günther announced that tuatara are not lizards and proposed a separate order, Rhynchocephalia, for tuatara and its fossil relatives (Günther 1867). This assessment was based on a number of skeleton features of tuatara. They have two temporal openings on each side of the skull behind the orbit, with complete arches (Günther 1867). Günther's conclusion about tuatara was supported by a comparison with fossil remains from the Upper Jurassic period found in Germany and England, which is known as *Homoeosaurus*. The skeleton and impressions of skin of species belonging to this genus were almost identical to tuatara (Crook 1975; Wettstein 1931). The Rhynchocephalia have preserved many primitive characteristics of an ancient reptile group in Triassic (Thecodontia), which is assumed to be the ancestor of all reptiles. In conclusion, tuatara appear to be the last remaining member of a distinct order Sphenodontia (Fraser 1988; Benton 1993), which

was represented by many species during Triassic period about 220 million years ago. All species apart from tuatara declined and eventually became extinct at the end of the Cretaceous about 65 million years ago (Cree & Butler 1993; Gaze 2001). Therefore, tuatara is of extraordinary zoological interest and is regarded as the most distinctive surviving reptilian genus in the world (Carroll 1988; Benton 2000; Fraser 1988).

3.2 Geographical Distribution, Taxonomy and Biology of Tuatara

Fossil tuatara have been recorded from Europe, Africa, England and north America suggesting that they were once widespread across the Pangaea super continent and were probably isolated in New Zealand when it broke away from Gondwanaland about 80 million years ago (Cree & Butler 1993). They were widely distributed over the North and South Islands before Polynesian and European times. They declined during the last 800 years after human arrival (Crook 1973; Gaze 2001; Duncan et al. 2002; King 2003) and probably became extinct on the mainland by the late 1700s (Newman 1878; Buller 1894). Tuatara are currently present in the wild on 12 island groups (Figure 3.1) and has been recognised as a species in need of active conservation management.

There are a variety of possible causes of the extinction of tuatara on the New Zealand mainland. There is good evidence of tuatara population decline after the introduction of rats such as Norway rat (*Rattus norvegicus*), the Ship rat (*R. rattus*) and particularly the Pacific or Polynesian rat (*R. exulans*) (Crook IG 1973). Apart from rats and other introduced mammals, humans indirectly threaten this species by reducing the amount of suitable environment through activities such as farming, burning and poaching (Cree & Butler 1993).

The current taxonomy based on genetic data has suggested two species of tuatara, one comprising two subspecies (Daugherty et al. 1990 b; Whitaker & Daugherty 1991). The latter species (*S.punctatus*) includes the northern tuatara subspecies (*Sphenodon punctatus punctatus*), which is distributed on nine island groups around the North Island. The second subspecies (unnamed) is found on islands in Cook Strait: namely Stephens and Trio islands. The second species (*S. guntheri*) is found on Brothers Island and also in Cook

Strait. The latter species currently consists of only a few hundred individuals (Cree & Butler 1993; Gaze 2001; Nelson et al. 2002).

Although tuatara were originally recognised as lizards, they have a number of distinct features that separate them from that group. Tuatara have a unique dentition, a small bony extension of the ribs, a complete lower temporal bar in the skull, and a gland beneath the skin on the head that contains the simple third eye. Finally the lack of a visible ear opening and copulatory organs in males are among the special characters that suggest the differences between tuatara and lizards (Newman 1987; Wright & Kevin 1994).

Tuatara have a long generation time and a slow reproductive rate. Males are typically sexually mature when the snout-vent-length (SVL) reaches about 180 mm, although they may not be sexually active until they grow longer. Females become sexually mature by the time that the SVL is about 170 mm. On average female tuatara lay 9-10 eggs in the nest that are then back-filled and guarded for several days before being abandoned. The incubation time for eggs is approximately 11-16 months with the success rate of 42%, and the average time between two clutches is around four years (Cree 1994). The temperature greatly affects the sex of the hatchlings (TSD: Temperature-dependent Sex Determination) (Newman *et al.* 1994). For example, the sex ratios (F:M) were 4:13 at 22°C, 31:3 at 20°C and all tuatara hatched were female at 18°C (Cree et al. 1995).

3.3 Genetic Variation among Different Tuatara Populations

Low genetic variation has been reported in tuatara for allozymes, mtDNA, and nDNA sequences. This low variation in tuatara was suggested to be a result of a bottleneck during the Pliocene/Pleistocene glaciations cycle (Hay et al. 2003). These data indicate genetic divergence between northern populations and Cook Strait populations. However, there are varying opinions about the history and taxonomic relationship between the Cook Strait and North Brothers Island populations. According to the allozymes data, the North Brothers population (*S. guntheri*) has widely diverged from other *Sphenodon* populations.

New Zealand

North Island

South Island

Cook Strait

△ Northern, *S. punctatus*

▽ Western Cook Strait, *S.punctatus*

□ Brothers, *S.guntheri*

○ Sub-fossil records

0   100   200
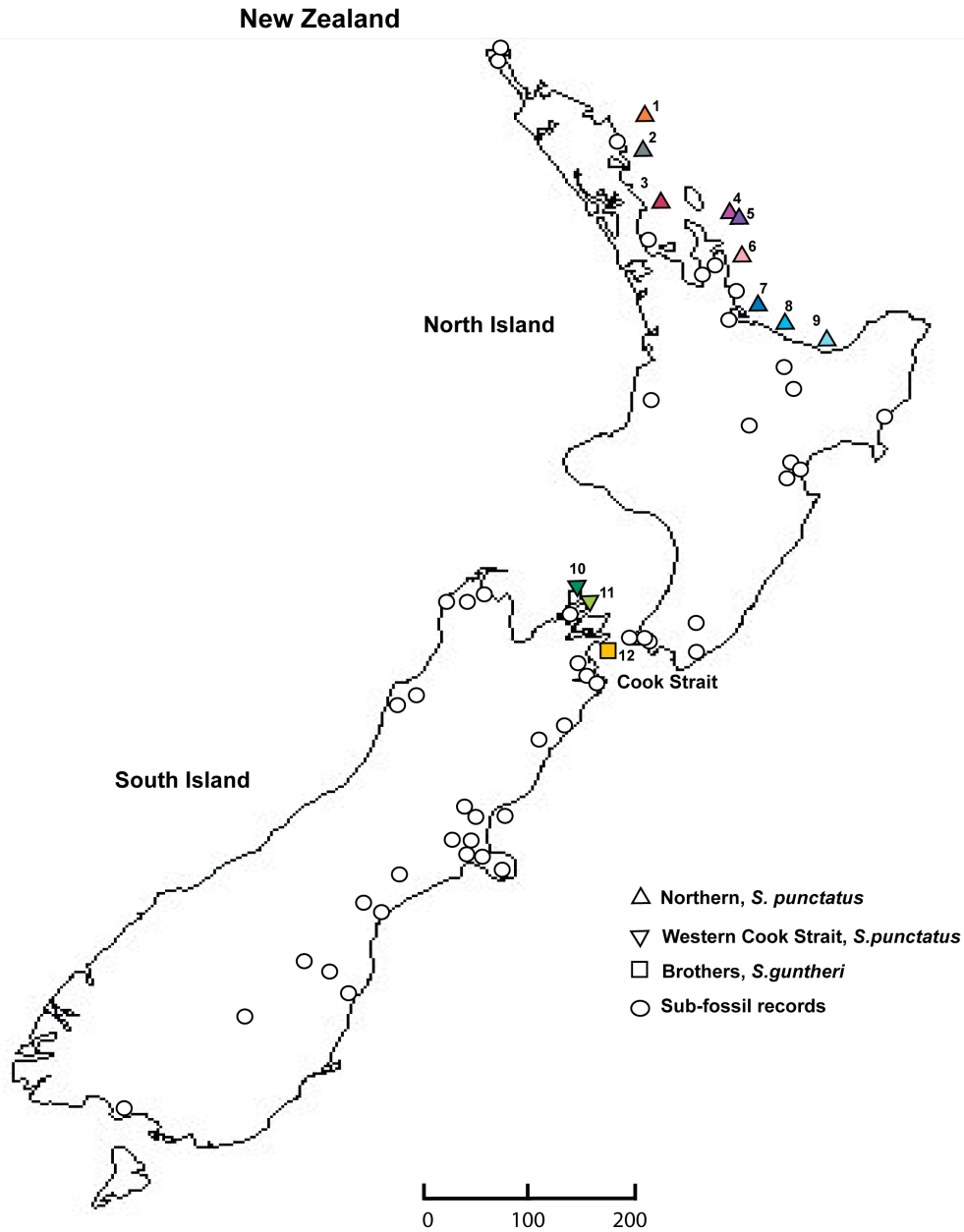
Figure 3.1: The geographic distribution of sub-fossil and extant populations of *Sphenodon* in New Zealand. Tuatara wild populations are currently present on 12 island groups which are marked as numbers; (1) the Poor Knights Islands (Tawhiti Rahi, Aorangi, Aorangaia, Stack B), (2) the Hen and Chickens Islands (Hen, Lady Alice, Whatupuke, Coppermine), (3) Little Barrier Island, (4) Cuvier Island, (5) the Mercury Islands (Stanley, Red, Middle and Green), (6) the Aldermen Islands (Ruamahua-iti, Ruamahua-nui, Hongiora, Hernia), (7-9) the Bay of Plenty Islands, (7) Karewa Island, (8) Motunau Island, (9) Moutoki Island, (10) Stephens Island, (11) the Trios Islands (Middle, North, South), (12) North Brother Island (Hay et al. 2004). The sub-fossil records are obtained from Crook (1975) and shown in circles. The figure is drawn by the Author of the thesis.

However, the data based on mtDNA suggest little genetic divergence between Cook Straits and North Brothers Island populations. This is in conflict with un-rooted gene tree derived from allozyme data (Hay et al. 2003) (Figure 3.2). Hay et al. (2003) suggested that as there is no suitable outgroup for these two un-rooted gene trees, it is likely that the data sets for allozymes and mtDNA reflect different aspects of tuatara's evolutionary history (Hay et al. 2003).
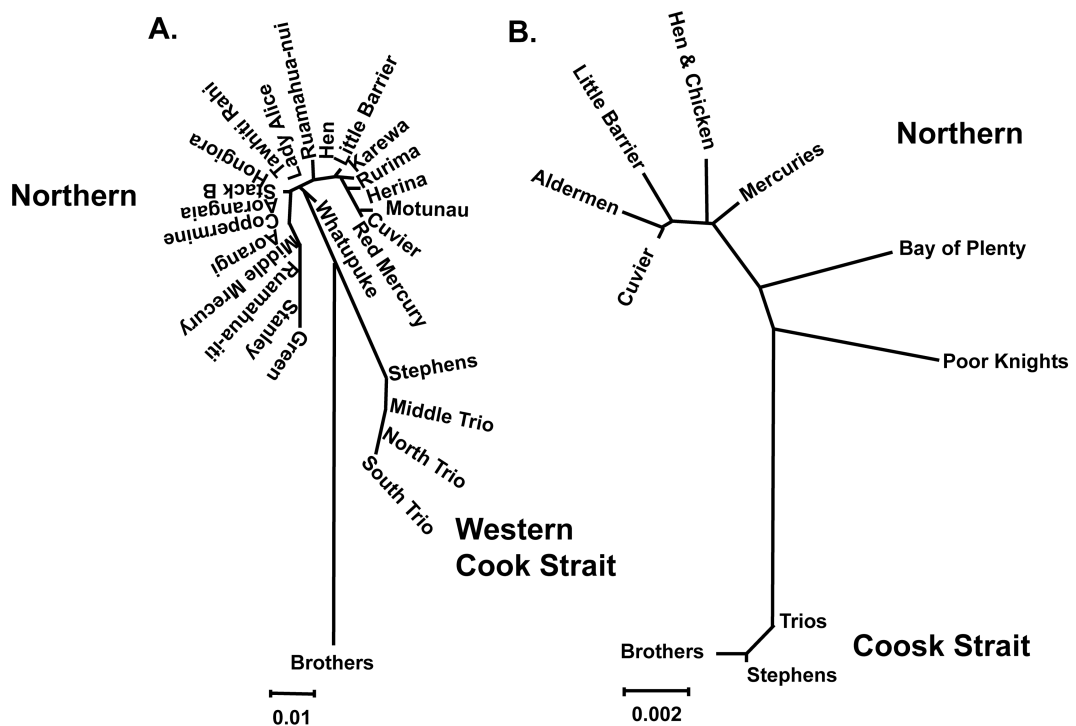
Figure 3.2: Un-rooted gene tree obtained from different tuatara populations based on **A.** allozymes data and **B.** mtDNA sequence data (Hay et al. 2003, 2004). The figure is re-drawn by the Author of the thesis.

3.4 Nuclear Mitochondrial Pseudogenes as Molecular Outgroup

The presence of an appropriate out-group is a requisite for phylogenetic studies as this provides the root for any phylogenetic tree and thus decides the direction of the evolution among the taxa of interest. As tuatara is phylogenetically an isolated taxon there is no close relative suitable as an out-group for providing the deep phylogenetic trees. To circumvent this problem Hay et al. (2004) suggested a molecular out-group for this living fossils rather than an organism out-group (Hay et al. 2004). The latter authors used the

nuclear copy of the mitochondrial (Numt) (Lopez et al. 1994) cytochrome b (Cyt b) gene as molecular outgroup. They constructed a minimum evolutionary tree based on 445 bp nucleotide sequences obtained from both nuclear and mitochondrial copies of cytb gene. The nuclear copy of cytb clearly places the root of the mitochondrial cytb tree between the northern and the southern (Trios, Stephens and Brothers) tuatara populations. This result supports the previous un-rooted mtDNA gene tree constructed by Hay et al. (2003) (Figure 3.2.B), suggesting differentiation among the northern and the Cook Strait populations (Hay et al. 2003; Hay et al. 2004).

3.5 Evolution of MHC in an Ancient Reptilian Order (*Sphenodontia*)

As reptiles are the sister taxon of mammals and birds, they are important for understanding the evolutionary history of many classes of nuclear loci such as the major histocompatibility complex (MHC) genes. MHC genes provide a link between the ancient ectothermic (fish and amphibians) and modern endothermic clade (mammals and birds). Therefore, an analysis of MHC gene sequences of tuatara could provide some information about the initial arrangement of MHC genes, from which avian and mammalian species diverged. Such analyses could enable us to estimate the time when the ancestral MHC gene from an ancient lineage diverged to the more sophisticated arrangement characteristic of eutherian mammals (Miller et al. 2006). To address this question Miller et al. (2006) generated different MCH class I-II sequences from tuatara and compared them to other vertebrates. DNA sequence analysis of MHC I-II in tuatara does not strongly group them with the other reptiles in phylogenetic tree. This probably reflects the antiquity of the *Sphenodon* lineage and the scarcity of closely related species for these living taxa (Miller et al. 2006; Miller et al. 2005). Maximum likelihood analysis of MHC-I sequences of reptiles revealed gene duplication within reptilian orders, however the evolutionary relationships among these sequences are still unclear (Miller et al. 2006).

Chapter Four

Materials and Methods

4.1 Materials

4.1.1 Tuatara Samples

In this research two different sources of samples were used. The ancient samples were exclusively sub-fossil jaw bones (Figure 4.2), which were kindly provided by three different museums in New Zealand (Table 4.1). The ancient bones were collected during the last 50 years and stored in cool dried museum conditions. The ancient samples were radiocarbon dated (^{14}C) at the Rafter Radio Carbon Laboratory, Institute of Geological and Nuclear Sciences, Lower Hutt, New Zealand (Appendix C). The estimated age of the samples used for this study range from 696±30 to 5,191±35 yrBP (years before present, i.e., 1950) (Table 4.1). The modern blood samples comprised both blood and toe pad tissues from different tuatara individuals derived from eight island groups around New Zealand (Table 4.2). Modern samples were collected under the ethics permit number 2009R12 from VUW animal ethics committee and stored in ethanol (EtOH) at -20°C. The geographical locations of the ancient and modern samples are shown in Figure 4.3.
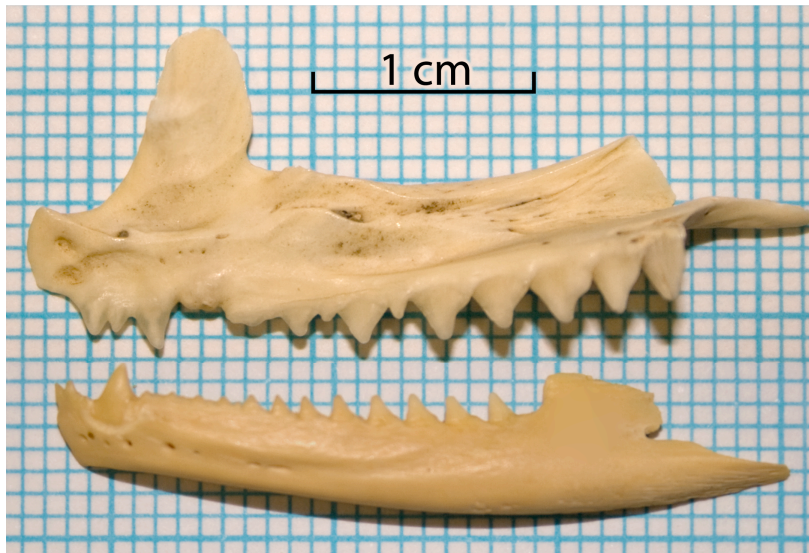
Figure 4.1: Typical examples of ancient tuatara material used for sampling. The two sub-fossil jaw bones shown above (upper: maxilla, lower: dentary) were provided by Te papa museum in Wellington. The photograph is taken by the Author of the thesis.
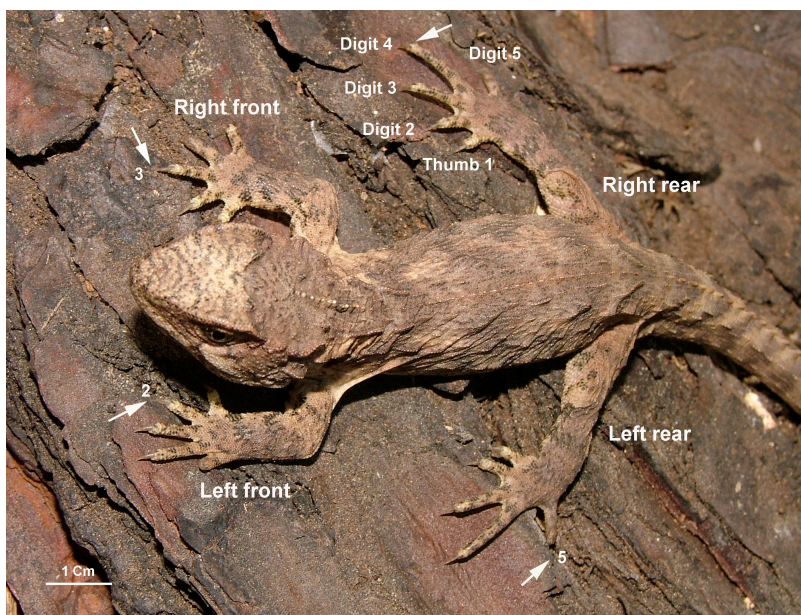
Figure 4.2: The method used for individual identification of tuatara juveniles, using the removal of toe pad tissues. For each individual juvenile tuatara, a toe is removed from each of the four limbs. Each toe is assigned a number; thumbs are '1' and digit numbers increase from proximal to distal. The code for any individual reads in the following order: left front, right front, left rear, right rear. For example if each of the digits indicated in this figure were removed, this juvenile tuatara would be designated '2354'. The snout-vent length for this juvenile is measured about 62 mm, and the total length is 130mm. The photograph is taken by Susan Keall from Victoria University of Wellington.

Table 4.1: Tuatara sub-fossil sample details used in this study. Age yrBP (years Before Present, where present=1950)

Region	Location	Site	Accession No.	Museum	Ancient Sample Sequenced, Age (yrBP)	Radiocarbon Reference No.
Northland	Tom Bowling		AU 4030	Department of Geology at Auckland University	T31 (1684 ± 35 yrBP)	R29386/2
Eastern North Island	Gisborne	Wainui	MNZ S44276	Museum of New Zealand, Te papa Wellington	T91 (696 ± 35)	R29328/3
	Wairarapa	Ruakokopatuna	MNZ S44277.S4		T87 (3,145 ± 35)	R 29386/12
Western South Island	Nelson	Tarakohe, NorthTakaka	MNZ S43850		T106 (5,191 ± 35)	R29386/16
	Nelson	Sims Cave, Takaka Valley	MNZ 33339 (26/1)		T82 (5,191 ± 35)	
Eastern South Island	Marlborough	Marfell Beach	CM REP 335	Southland Museum and Art Gallery, Invercargill	T142 (1,189 ± 35)	R 29386/22
	North Canterbury	Holocene Cave, Mt Cookson	MNZ 33366 (8/2)	Museum of New Zealand, Te papa Wellington	T110 (1,794 ± 35)	R29386/17
	Otago	Knobby Range, North Alexandra	MNZ S34098		T102 (1,546 ± 35)	R29386/15
Southland	Wakapatu		SM 6	Southland Museum and Art Gallery, Invercargill	T151 (2,321 ± 35)	R29386/24
			SM 1		T60=T71 Not dated	
			SM 28		T62 Not dated	

Table 4.2: The geographical locations of sampled modern tuatara populations

Island group	Island	Sample ID	Island group	Island	Sample ID
Stephens	Stephens	FT4244	Poor Knights	Tawhiti Rahi	FT2522
		FT4251	Hen & Chickens	Hen	FT2409
		FT4253	Little Barrier	Little Barrier	FT7200
		FT4254			FT7205
		FT4850			FT7613
		FT4859			FT7625
		FT4889			FT7896
		FT4906			FT7899
		FT4911			FT7930
		FT4913	Cuvier	Cuvier	FT2917
		FT4915	Mercury	Green Mercury	FT2391
		FT4926			FT2392
		FT4961		Middle Mercury	FT2365
		FT5002			FT2369
		FT5044		Red Mercury	FT2654
		FT5093		Stanley	FT2704
		FT5098	Bay of Plenty (Motunau)	Motunau	FT2139
		FT5099			
		FT5102	Brothers	North Brothers	FT200
		FT4994			FT210
		FT5113			FT205
		CD1333			
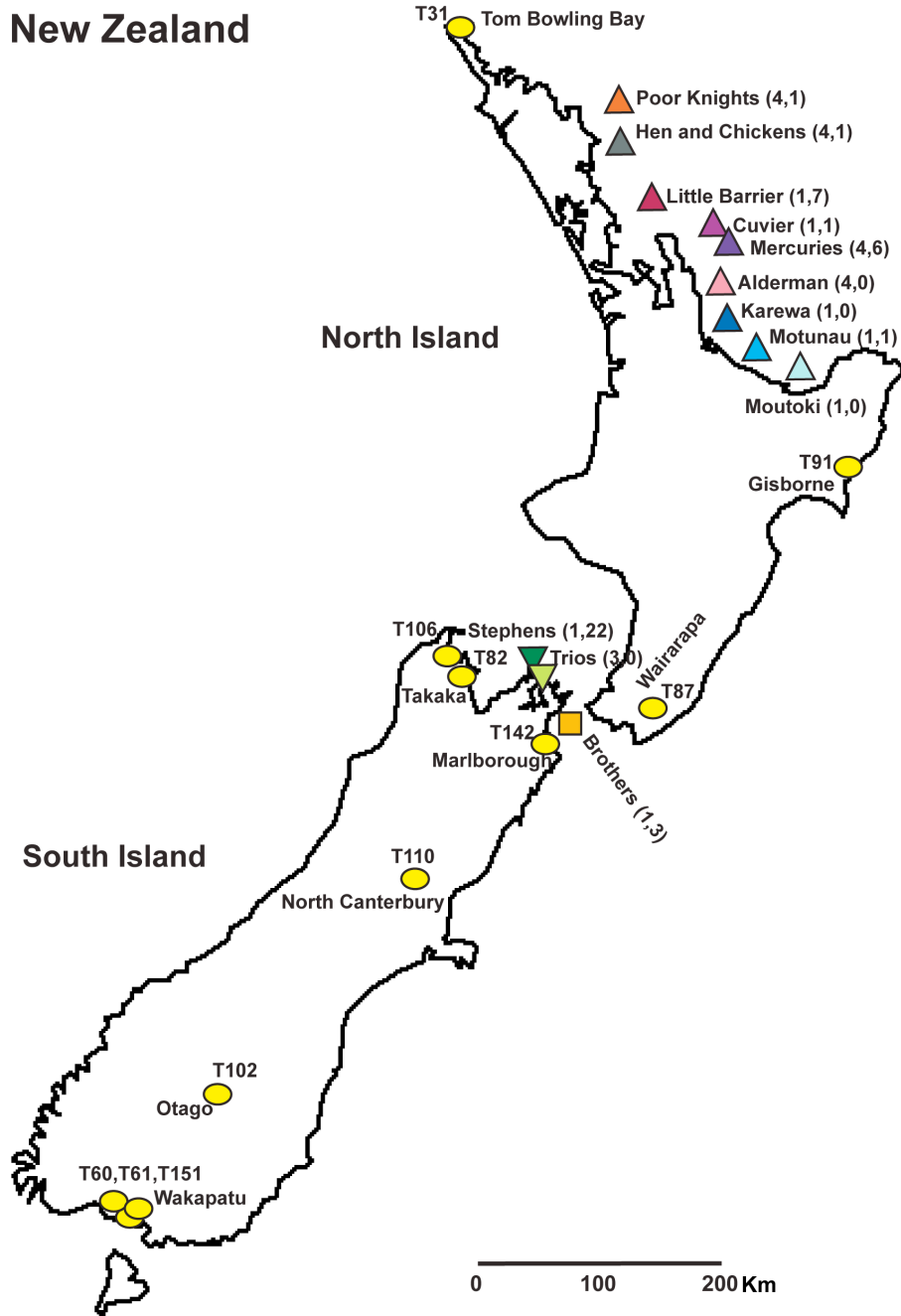
Figure 4.3: Location of modern and sub-fossil sites of tuatara samples used in this study. The island groups are marked as a colored rectangle or colored triangles. The number of islands in each island groups and the number of samples per island group are indicated in brackets. The sub-fossil sites are marked as yellow circles with the name of the location of each sub-fossil sample. The figure is drawn by the Author of the thesis.

4.1.2 Oligonucleotide Primers

Table 4.3: The external primers (5′ to 3′ direction) designed for amplifying the complete tuatara mitochondrial genome (set A)

Primer Name		Primer Sequence	
F01	GGCGTAGCAAAACCTGGATA	F09	AACTAGGCTTGGCCCCATT
R01	GCATATGGGCGTACTGGTTC	R09	CCTATATGGGCGATTGAGGA
F17	CCTTATTCATTTGATCCGTCCT	F25	TTGGCCACCAATGATACTGA
R17	GGTTCTTTCTTGTTTGAGTAGTCG	R25	TGGCAGGTGTGTTGAGATTG
F33	CATGGCCTATTACAGGAGCAA	F41	AACAACCTAAATCTCCTTCAATGC
R33	TGGGGGATAGGCTAGAGTGA	R41	GGGAGTTTTGGCTTGCTAGT
F49	TGATCTGGCACGTATTAACCA	F57	CGAAACCATTCGAAAACACC
R49	TTTGTTGGGTCTGGGTTTTC	R57	CTCGGGTTGCATGAAGGTAT
F65	CGCGCATAATATCCTATTCCA	F73	TCAAGTTGAAAAAGGCTCTGG
R65	GTCCGTTACACACCCATGC	R73	CCTGGGGTTCTTTTCTATCTCC
F81	CCGAAAACCTCATTTAGACCA	F11	TAGCCCTAACCCTGCTTTCA
R81	GATTGCTCCGGTCTGAACTC	R11	TTTGTGGGAGTAAAATGTGGTG
F03	CAATCTTATGATCGGGCTGA	F27	GACTGATCATGGGCCACAAC
R03	AATGGGGCTCGGTTTGTT	R27	GCACATGGTTGTGGGATAAA
F19	ATTATTGCCATCCCAACAGG	F43	ATTCGAAGCAACGCTTATCC
R19	AGGGGGAATCAATGGGTTAG	R43	ACATGCACCTCACAGAAAGAAA
F35	GGGGTAACAGTCACCTGAGC	F59	AAACCTCTTGTCCGCCATAC
R35	CCCAAAATGGTGGTCTATGG	R59	GTGCTCCTAGCAGGTCCTTG
F51	CCAACACCAACCAAAAATCC	F75	TACCCCGAAATCAAGTGAGC
R51	GAGGTCGAGAAGGAATGCAC	R75	TTAAGGCCTACGGTTTGGTG
F67	CCCCCACGAACATAGATTAGC	F13	TTTCTGAGGAAACCTGAAAGATT
R67	TRCATGGTGTGGTTGTRGGT	R13	TCAGCGTTTGTATGGCAAAG
F05	AAATCCAGGGCAAGGAATAA	F29	CGGTGATAATACCAGCAATGG
R05	GTGATCCTTTATTCAGGCACAT	R29	AGGCTAGGGCCATGGTTAAT
F21	CGCACTAACATTCATCGGAGT	F45	CAATGGTCTTAGCCGCAATC
R21	GGGGGTGTCCATGTAGTCAT	R45	TCCCGTTAGGCTTCATTCTG
F37	TTCCAATCATTAAGTTCTAGACAACC	F61	GAAAATTTCACCCCAGCAAA
R37	GCGGTAGGGGTGTTAGATTG	R61	TGGGGGTTCTACTGGTTGTC
F53	CGCGCATAATATCCTATTCCA	F77	AAAGATTTCTCCGTGCATCC
R53	GTCCGTTACACACCCATGC	R77	TGTTCCGAGTTCCTTCTGGT
F69	TGCCTCAATGACACAAAGATT	F15	CCGAGGAGAATTAAGCTATCCA
R69	GCTGGCACGAAATTGACC	R15	CGGGTATACGGTTCATCCTG
F07	CTAATGCACCCCTCCATCAC	F31	ATTTCCTCCCAGAAGGTTCC
R07	TCTCATTGGGCAGTTTGTCA	R31	AGGCTTGAATAAGTGCTACTGC
F23	CAAGCTAGTTGAGCAACCACA	F47	TGAACGTATTAACAATCGAACACT
R23	GGCGTCTGTTGCGTTAATGT	R47	GGATGCCAGGTGGTTAGAGA
F39	GACATGAACTTCCCTAATTTTAACC	F63	TCTCGCTCAAGAAAAGAGGA
R39	CCAAGCATAATAAAGCGGAGA	R63	AYTCGTGAGGTCGAGAAGGA
F55	CCCCCACGAACATAGATTAGC	F79	CCGTGCAAAGGTAGCGTAAT
R55	CCCCCACGAACATAGATTAGC	R79	CCAAAACAGAGGGGGTCATT
F71	TTGGGGCCAAGATAAAATTG	F81	CCGAAAACCTCATTTAGACCA
R71	GCGGTAATGCTGTTTGACCT	R81	GATTGCTCCGGTCTGAACTC

Table 4.4: The internal primers (5′ to 3′ direction) designed for amplifying the complete tuatara mitochondrial genome (set B)

Primer Name		Primer Sequence	
F02	CAACCTATTGCAGATGGCCTA	F10	GGTGAGGAGGAATGAACCAA
R02	GTTTGTGCTACCGCTCGAAT	R10	TCCTGTTATGGGTGGTAGGC
F18	TCTTACCTGGATTTGGGATAATCT	F26	TTGGTATCAGCAGAAGATGTCC
R18	TGCTTCTCATTTAATTATTCCTCCA	R26	CTAAAAGGCTGGTGCTGCAT
F34	TTTGTTTATTGTCTCCGAGGTG	F42	AGATCACATTTCATCCCCACT
R34	CCGATTTTCGTTTCCCTTCT	R42	TTAAGGCGTTCTGGTTGGTT
F50	CAACTGTCTCTTCCCCAACC	F58	CCGCAATATACATGCAAACG
R50	GGGGTGTTAGGGCTTTTGAT	R58	GGTTAGTGTGGCGTTGTCAAT
F66	TCCTTTCAACCACATATCCAG	F74	AACTAGCCCGCTACCCAGAT
R66	TTRACTCTTGGTTTAGGGGTTT	R74	TTGCCACAGAGCCGAATTAT
F82	TGTTCAACGATTAACAGTCCTACG	F12	ACCACTCCACCAAACACCTC
R82	GGCGAGAGGAGTTGAACCT	R12	GAGGCCTGCTCATCTAGGAA
F04	AACCCTAAACACCACGCAAG	F28	CAATAGTCCTAGTGTGGCTCTCA
R04	CTCGGACCCATAGGAATCAA	R28	GCCCTGTTACAAGGTGTTGG
F20	CCTTAGACATTATTCTCCACGACA	F44	CTGAACAACACCCTATTAATCACA
R20	TCATGTGGAGTATGCATCTGG	R44	TTATTGGTGTGAGCATGGTG
F36	TGCCTCCTACGACAAATCCT	F60	AGGATCGAACAACCCAACAG
R36	TGGTTGTGTGTTTGGGAGTC	R60	CAACACCCCTCCCATTTTATT
F52	GGYACGCCGCCTTATATT	F76	AAATAGGAAACAGCCTATACAAGAGG
R52	GGAAGATCCGTAAGCACTGG	R76	AATTATTGTGTAATGACCAGTAGGTTG
F68	AAGGAAGCCTTATGCATTTTT	F14	CAGGCCTGATAAGAAAAGGAA
F68	CACTGGTGTGCTGATACTTGC	R14	AAGCATGGGCTGTTACGATT
F06	CCTTTGACACACATCCATGC	F30	ACGCAAGGCCATAAATGAAC
R06	GCTTGTGGATAGGGCGATAA	R30	CTACCACAAGGGCTGGGATA
F22	CCTTTGAGAGGCTTTCTCTTCA	F46	TCATTTGCCTTCGACAAACA
R22	GCAATTTGTATGGGGTAGGC	R46	AAAGTCATCATGTGGAGAGAAGG
F38	CTCGCCTCCCATTTTCTATG	F62	CACCGCCCAATCAGTAAAAT
R38	GAACTAACTCTTCATTCGGCTCA	R62	TTTATTTAAAACTCCGGTTCTGG
F54	TCCTTTCAACCACATATCCAG	F78	AGGCATGCACCTAGGAAAGA
R54	TTRACTCTTGGTTTAGGGGTTT	R78	TTTCGCCAGGAGACAGTTG
F70	ACACCCCCAAGGGTACACAG	F16	TATTGACATCCGCCTGAACA
R70	GGGGATCTAATCCCAGTTTGA	R16	TCCTGCTGCTAGTACGGGTAA
F08	GCCGTAGAAGCCACCATAAA	F32	CCCATATGACCGGCTACAGT
R08	TAGTGGGGCCAGTTTTTGTC	R32	AAATCATACGGCGAGACCTG
F24	TCCGTGTTCTACATAATCACCA	F48	TTACAGGAGCTGGCGTAGTG
R24	GGAGTCAAATGAGAGGTCTTTGT	R48	AGAGTTGACTCGGGGTCAGT
F40	GCCTGTCATTACAACGCTCA	F64	GGYACGCCGCCTTATATT
R40	GGGAGAAGGAAAATTGTTGGA	R64	GGAAGATCCGTAAGCACTGG
F56	CGCTTTTAGAGGGGAGGAAC	F80	CCTTGTGCAAAAGCAAGGAT
R56	TACGTCGGCGGTGTAATGTA	R80	CCTGGGGTAACTTGGTTCAA
F72	TCGCCAGTCTACCTTGTGAA	F82	TGTTCAACGATTAACAGTCCTACG
F02	CAACCTATTGCAGATGGCCTA	R82	GGCGAGAGGAGTTGAACCT

Table 4.5: The most used primers (5′ to 3′ direction) for amplification of tuatara mitochondrial genome, using long-range PCR.

Primer Pairs (F/R)	Product Size (bp)	Primer Pairs (F/R)	Product Size (bp)
F01/R05	1,064	F41/R47	1,298
F05/R10	1,208	F47/R52	1,243
F10/R15	1,274	F52/R57	1,191
F15/R20	1,240	F57/R62	1,202
F20/R25	1,165	F62/R67	1,210
F25/R31	1,304	F67/R72	1,257
F31/R36	1,242	F72/R77	1,156
F36/R41	1,160	F77/R02	1,287
F01/R21	4,121	F41/R63	4,296
F20/R43	4,403	F61/R02	4,535

4.2 Methods

4.2.1 Contamination Precaution

Standard authentication criteria for ancient DNA studies, including multiple extractions, and independent PCR amplification, have been followed. In order to prevent possible contamination all stages of the work were carried out under sterile conditions. All steps (bone cutting, surface removing, powdering, DNA extraction and PCR set up) were carried out in separate rooms in a dedicated ancient laboratory at Massey University in Auckland that undergoes regular decontamination by the overnight UV irradiation. The protective cloths, sterile latex gloves, disposable mask and the face shield were used while performing the experiments in ancient laboratory.

4.2.2 DNA Extraction

Briefly, three methods were used to extract DNA from tuatara. For the routine extraction of DNA from blood, the proteinase K, phenol/chloroform and for DNA extraction from ancient materials silica-based method were used.

4.2.2.1 DNA Extraction using Proteinase K Digestion and Phenol/Chloroform Purification

DNA was extracted from approximately 10 µl of tuatara blood in 20 - 40 µl of Queen's lysis buffer (10 mM Tris-Cl pH 8.0, 10 mM NaCl, 10 mM EDTA, 1% n-lauroylsarcosine) (Seutin et al. 1991) following the method of Sambrook et al. (1989). Proteins were digested by the addition of 400 µl of SET buffer (20 mM Tris-Cl pH 7.6, 50 mM NaCl, 1 mM EDTA), 20 µl of 20 mg/ml proteinase K (Blin & Stafford 1976) and 20 µl of 10% SDS. The mix was incubated overnight at 55-65°C to allow digestion.

After digestion, the DNA was purified by phenol/chloroform extraction. Proteins were denatured and removed from the aqueous phase by the addition of 400 µl of phenol (pH 7.8-8), mixed for 30 min by rocking and the phases were separated by centrifugation at 16,060 g for 5 min. The supernatant was re-extracted with 400 µl of phenol-chloroform-isoamyl alcohol (25:24:1) and mixed for 30 min by rocking and the phases were separated by centrifugation at 16,060 g for 5 min. The supernatant was re-extracted with 400 µl of 24:1 chloroform-isoamyl alcohol followed by mixing and centrifugation at 16,060 g for 5 min. The resulting aqueous phase was carefully removed to avoid the PCR-inhibiting organic phase.

The DNA was precipitated by adding 40 µl of 3M sodium acetate (NaOAc, pH 5.2) and 800 µl of 100% EtOH, chilled to -20°C. The mix was gently rocked for 15 min and then centrifuged for 30 min at 16,060 g at 4°C. The supernatant was removed and the pellet was washed and re-suspended with 500-1000 µl chilled 70% EtOH. The mix was then centrifuged for 5 - 12 min at 16,060 g at 4°C, the supernatant was removed and the pellet was left to dry at room temperature for approximately 3 - 8 min and then re-suspended in 100 µl of MilliQ water or TE buffer (10 mM Tris-HCl, 0.5-1 mM EDTA, pH 7.6) at 4°C or room temperature (RT) overnight (Sambrook et al. 1989). A negative control extraction was performed for every set of extractions. After extraction 1 - 2 µl of DNA was subjected to electrophoresis in 1.5% agarose, stained with ethidium bromide (50 ng/ml) and visualized over UV light in Gel Doc (Bio-RAD) apparatus.

4.2.2.2 DNA Extraction using Silica

In the first step, the bones were ground using mortar and pestle and 0.10 - 0.25 gr of the bone powder incubated overnight in 5 ml of extraction buffer (0.45 M EDTA pH 8.0, 0.25 mg/ml proteinase K) at room temperature in the dark. After centrifugation for 2 min at 1,520 g the supernatant was transferred to a clean Eppendorf tube and 18 ml L2 buffer (5 M GuSCN, 0.05 M Tris-HCl pH 8.0, 0.025 M NaCl) and 100 µl of silica suspension were added. Tubes were sealed with parafilm and rotated for 1 - 3 hrs at room temperature in the dark followed by centrifugation at 1,520 g for 2 min. The supernatant was then discarded and the pellet was re-suspended in 1 ml L2 buffer and centrifuged at 16,060 g for 5 sec in a mini-centrifuge. The supernatant was removed and the pellet washed in 1 ml of new wash buffer (EtOH 51.3%, 125 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) and centrifuged at 16,060 g for 1 min. The supernatant was then discarded and the washing step repeated until the silica pellet was almost white. The pellet was then dried at room temperature for about 15 min and re-dissolved in 100 µl of Trizma-EDTA (TE) buffer (1 mM EDTA pH 8.0, 10 mM Tris-HCl pH 8.0) to release the DNA from the silica. After suspending the DNA in TE buffer it was briefly centrifuged at 16,060 g for 2 min and the supernatant was transferred to a new tube to avoid traces of silica in the DNA extraction (Rohland & Hofreiter 2007).

4.2.3 Amplification of Complete Mitochondrial Genomes of Tuatara using Nested Multiplex Polymerase Chain Reaction (PCR)

Briefly, the nested multiplex PCR method of Krause et al. (2006) was used to amplify complete mitochondrial genomes of tuatara. 82 primer pairs were designed using Primer3 version 0.4.0 and PrimerQuest in two overlapping sets (set A and set B, see Materials) to cover the complete tuatara mitochondrial genome (Figure 4.4). Initially, two amplifications were performed using primer set A and primer set B separately. The amplified mix from each set was then separately diluted before second sets of amplifications were carried out using individual primers in each set.
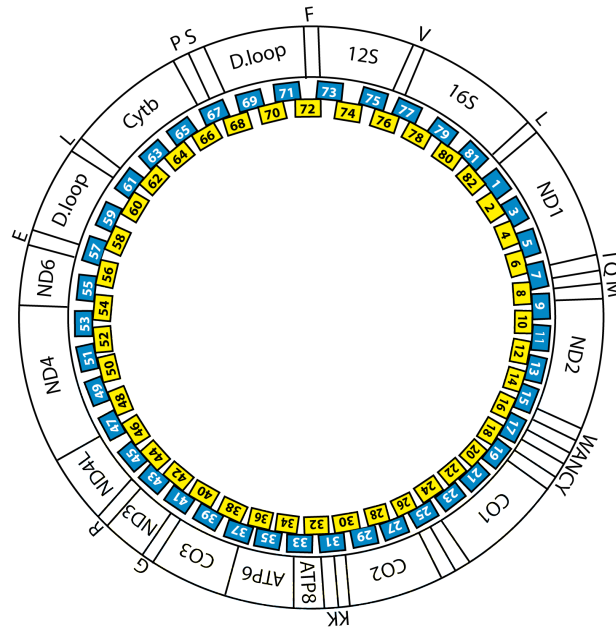
Figure 4.4: Mitochondrial genome map for tuatara, including two D-loops, 21 tRNAs, 2 rRNA and 12 polypeptide genes. The relative size and positions of the 82 PCR amplicons are shown as external (blue, set A) and internal (yellow, set B) overlapping sets. The figure is drawn by the Author of thesis.

4.2.3.1 Multiplex-PCR

For this work the multiplex-PCR approach designed by Rompler et al. (2006) was used to simultaneously amplify the entire mitochondrial genome of tuatara using a small amount of aDNA in two initial amplifications (Rompler et al. 2006; Krause et al. 2006). Briefly, multiple fragments were simultaneously amplified in one multiplex-PCR using each set of primer sets separately (41 primer pairs in each set). The initial multiplex PCR mix is then diluted and a series of second amplifications is carried out using each of 41 primer pairs of each set individually (Rompler et al. 2006; Krause et al. 2006) (Figure 4.5). Using this method, the entire mitochondrial genome of tuatara can be constructed from a small amount of aDNA.

The first step of multiplex PCR was carried out in a 20 µl volume containing 1X PCR buffer (Invitrogen), 4 mM MgCl2 (Invitrogen), 1 mg/ml BSA (Invitrogen), 250 µM mix dNTPs (Bioline), 0.15 µM each primer (41 primer pairs in each set), 1 U of Platinum® Taq DNA polymerase (Invitrogen) and 5 µl aDNA extract (regardless of the amount of DNA). One extraction control (blank) and one PCR control were added for every five PCR reactions.

The PCR mixes were amplified using an iCycler™ Thermal cycler (Bio-RAD) outside the ancient DNA lab. The amplification programme consisted of initial denaturation at 94°C for 9 min followed by 30 cycles consists of 94°C for 20 sec 52°C for 30 sec 72°C for 30 sec and a final extension of 72°C for 4 min. PCR products were stored at 4°C until further analysis. The PCR products were checked by electrophoresis in 1.5% agarose in TBE and then diluted in water in a ratio of 1:20 to 1:100 (I routinely used a dilution of 1:25) for amplification by simplex PCR.

4.2.3.2 Simplex PCR

Simplex PCR was used to specifically amplify the entire tuatara mitochondrial genome in 82 fragments. Simplex PCR was carried out in 20 µl volume containing 1X PCR buffer (Invitrogen), 4 mM MgCl2 (Invitrogen), 1 mg/ml BSA (Invitrogen), 250 µM mix dNTPs (BioLine), 0.75 µM forward primer (Invitrogen), 0.75 µM reverse primer (Invitrogen), 0.25 U of Platinum® Taq DNA polymerase (Invitrogen) and 5 µl DNA template (1:25 diluted multiplex-PCR products).

The PCR mixes were amplified using an iCycler™ Thermal cycler (Bio-RAD) outside the ancient DNA lab. The amplification programme consisted of initial denaturation at 94°C for 9 min followed by 30 cycles consists of 94°C for 20 sec 55°C (primer dependent) for 30 sec 72°C for 30 sec and a final extension of 72°C for 4 min. The PCR products were subjected to electrophoresis in 1.5% agarose, stained with ethidium bromide (50 ng/ml) and visualized over UV light. One negative control was included for every eight PCR reactions.

4.2.3.3 Repetition of Selected Samples at another Ancient DNA Facility

In accordance with accepted procedures for ancient DNA analysis, a subset of samples were re-analysed at a separate ancient DNA facility (Cooper & Poinar 2000; Krings et al. 1997). DNA was extracted from a subset of 9 ancient tuatara samples and amplified for control region sequences (using primers F52/R52, F53/R53, F54/R54) at Griffith University Ancient DNA laboratories, Brisbane, Australia. In all cases, the sequences matched those obtained from the Massey University Ancient DNA facility.
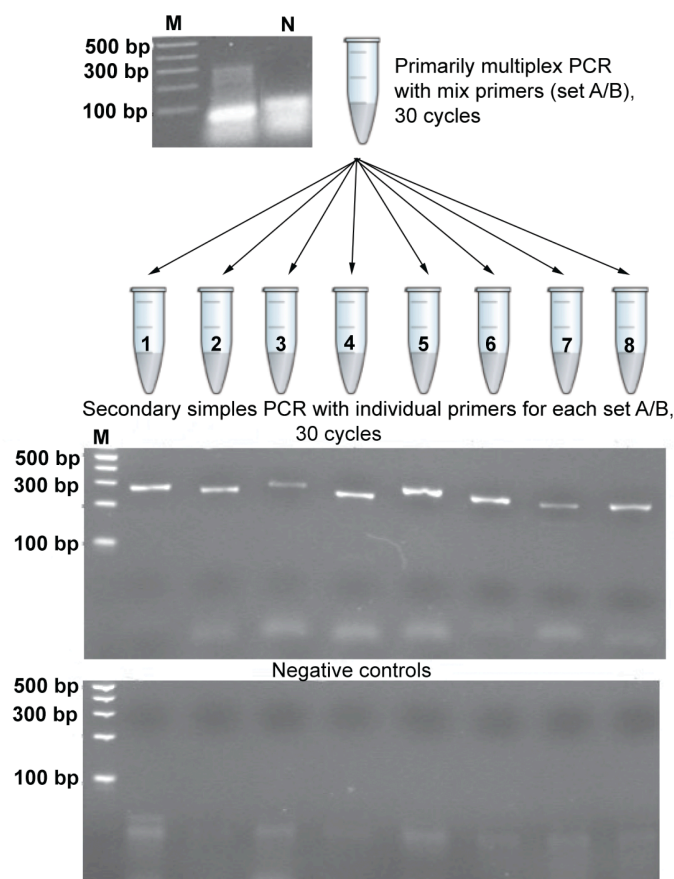
Figure 4.5: Principle of the multiplex-PCR approach and typical results. The PCR products from the primarily multiplex-PCR comprise different fragments of various sizes amplified by primer pairs in set A/B. In the second step, each fragment is specifically amplified using only one primer pair. Negative controls for the same primer pairs are shown in the lower gel. Gel lanes marked with M and N contains molecular size markers and negative control respectively. The size of PCR products ranges from 200 to 300 bp. The figure is drawn by the Author of the thesis.

4.2.3.4 Amplification of the Mitochondrial Genome using Long-Range PCR

For high quality modern tuatara DNA complete mitochondrial genomes were obtained using long-range PCR of 1 - 4 kb DNA fragments. The Long-Range PCR carried out in a 20 µl volume containing 1X PCR buffer (Invitrogen), 50 mM MgCl2 (Invitrogen), 1 mg/ml BSA (Invitrogen), 250 µM mix dNTPs (BioLine), 0.50 µM forward primer (Invitrogen), 0.50 µM reverse primer (Invitrogen), 5 U Elongase® Taq and 1 µl DNA template. The PCR mixes were amplified using an iCycler™ Thermal cycler (Bio-RAD, USA) in modern lab. The amplification programme consisted of initial denaturing at

94°C for 4 min followed by 10 cycles consists of 94°C for 30 sec, 57.5°C (primer specific) for 30 sec, 8 min extension at 68°C and another 30 cycles consists of 94°C for 30 sec, 57.5°C (primer specific) for 30 sec and 68°C for 8 min. For the following 30 cycles the extension time was increased by 20 sec after cycle one. The final extension at 68°C for 4 min was followed. The PCR products stored at 4°C for further analysis. The PCR products were subjected to electrophoresis in 1.5% agarose, stained with ethidium bromide (50 ng/ml) and visualized over ultra violet (UV) light.

4.2.4 PCR Product Purification and Sequencing

The modern and ancient PCR products were purified for sequencing by using either SPRI (Solid-Phase Reversible Immobilization) Agencourt® AMPure® XP, or Zymo DNA clean and concentrator™-25 kits according to the manufacturer's instructions. The purified PCR products were then sequenced directly in both directions on an Applied Biosystems 3730 DNA analyser (the Allan Wilson Centre Genome Service), using BigDye Terminator Version 3.1 Ready Reaction Cycle Sequencing kit (Applied Biosystems).

4.2.5 Sample Preparation for FLX Sequencing

Eight complete mitochondrial genomes from modern tuatara (Table 4.6) were sequenced using high throughput DNA sequencing technology. The entire mitochondrial genome (16 kb) of eight tuatara samples were amplified in 4 fragments approximately 4 kb in length as outlined in 4.2.3.4 and purified using Agencourt AMPure beads. After purification the DNA concentration was measured by NanoDrop spectrophotometer (Asp-3700, ACTGene). The PCR products for each sample pooled in equimolar ratios in order to avoid unequal sequence coverage rates between different fragments. For the long-range PCR products the final concentration for the pooled mitochondrial genomes were adjusted to give a concentration greater than 50 ng/μl in a final volume of 60 μl. The total amount of DNA for each sample should be greater than 3 μg (i.e. >60μl of 50 ng/μl sample), but not exceed much more than that as it results in chimera formation during the following ligation step. The optical density (OD) A260/A280 was measured with NanoDrop spectrophotometer (Asp-3700, ACTGene) to ensure this value was greater than 1.65.

Next generation FLX sequencing was carried out at the University of Otago high

throughput DNA sequencing unit. Sample-specific sequence tags (Table 4.6) were ligated to each of the eight mitochondrial genome libraries and then were combined to produce one FLX sequencing library, using the GS DNA library preparation kit according to the manufacturers instructions. The sequencing library was sequenced on 1/16 equivalent of a standard GS FLX platform according to the standard FLX routine (Margulies et al. 2005; GS FLX system manufacturer's instruction). After sequencing, the FLX sequence reads were sorted according to sample-specific tags and the rest discarded. The output was saved as FASTA or sff (standard flow format) files for analysis by FLX assembly software.

Table 4.6: The list of samples and the sequences of the specific tags used for preparing the FLX sequence library. The numbers of enriched beads for constructed libraries range from 2,200-152,400.

Sample ID	Sample-specific Tag Sequence		Emulsion PCR (emPCR)	Number of Enriched Beads for library
FT4253	MID 7	CGTGTCTCTA	emPCR 1	2,200
FT4244	MID 8	CTCGCGTGTC	emPCR 2	69,400
FT4252	MID 9	TAGTATCAGC	emPCR 3	152,400
FT4251	MID 10	TCTCTATGCG	emPCR 4	125,000
FT4246	MID 11	TGATACGTCT	emPCR 5	113,800
FT4249	MID 12	TACTGAGCTA	emPCR 6	52,200
FT4241	MID 5	ATCAGACACG	emPCR 7	75,400
FT4250	MID 6	ATATCGCGAG	emPCR 8	46,800

4.2.5 Analytical Methods

Analytical methods corresponding to different data analysis are detailed in the related data chapters that follow.

Chapter Five

A Single Species of Tuatara? Reassessment of Genetic Diversity and the Taxonomy of Tuatara (*Sphenodon*: Reptilia)

5.1 Introduction

Tuatara are significant in terms of global biodiversity and the evolutionary history of reptiles (Cree & Butler 1993). They are the sister taxon to the Squamata (including lizards, and snakes) (Rest et al. 2003) and are regarded as the most distinctive surviving reptilian genus in the world. Although they resemble most lizards, the tuatara is actually the last remaining member of a distinct order Sphenodontia. This unique reptile is exclusively found on some offshore islands around New Zealand and was once widespread on the mainland before the arrival of humans and introduced mammals around 800 years ago (Gaze 2001; Duncan et al. 2002; King 2003). The fossil record shows minor morphological changes in modern tuatara compared to its Cretaceous relatives (Carroll 1988; Benton 1993; Benton 2000; Apesteguia & Novas 2003).

Prior to 1990, New Zealand conservation legislation recognised only one species of

tuatara, namely *Sphenodon punctatus*. In 1990, Daugherty and colleagues suggested the existence of two species of tuatara based on genetic data and morphological variation (Daugherty et al. 1990 b). To support their suggestion that in addition to the widely accepted taxon *Sphenodon punctatus* (including two subspecies), a second congeneric species, *S. guntheri*, should be recognised, they reported a phylogenetic tree based on isozyme data (Daugherty et al. 1990 b). These two species are currently present on 35 islands off the coast of New Zealand. The northern and Cook Strait tuatara populations consist of 10,000 and 45,000 individuals respectively, while only 400 tuatara individuals exist in Brothers Island (Groombridge 1982; Hallidy & Adler 2002). Mainly based on allozyme differences observed by Daugherty et al. (1990 b), currently the New Zealand Department of Conservation and International Union for Conservation of Nature's Red Data Book (IUCN 2001) recognises both *S. punctatus* and *S. guntheri* and classifies *S. guntheri* as a category A (vulnerable, requiring urgent recovery work), while *S. punctatus* is in category B (low risk, requiring work in the short term). This designation has in turn influenced the planning of conservation priorities and consequently affected some of the conservation management decisions (Cree & Butler 1993; Gaze 2001; Nelson 2002). For instance, *S. guntheri* is highly protected under the Wildlife Act in New Zealand and the Brothers Island is designated as a Wildlife Sanctuary which permits are required for visit (Groombridge 1982). In addition, New Zealand's Wildlife Service (now DOC) ran a programme to investigate more about the ecology and factors affecting the survival of these two species (Crook, 1973), which in fact affects the conservation management decisions.

Daugherty et al. (1990 b) suggested that as the newly recognised species had previously gone unrecognised and because taxonomic classification is a primary determinant of management priorities for endangered species, this contributed to the loss of 25% (10 out of 40) of populations in the 19th century. From 37 populations of *S.p.punctatus,* 24 populations are distributed on 24 islands in Hauraki Gulf and Bay of Plenty plus 4 islands in Cook Strait and the other populations on Mokohinau group (3 islands), Motiti, Slipper, Shoe, Whenuakura, Whale and Somes Islands became extinct (Wettstein 1931; Mertens 1954; Williams & Given 1981; Crook 1970). The *S.p.reischeki* were present on Little Barrier Island in the Hauraki Gulf and became extinct in late 1970s (Wettstein 1943). *S.guntheri* are currently confined to only one island, North Brother Island in Cook Strait, and the other population from East Island became extinct in the

19th century (Buller 1878). Despite absolute protection of the species and its island habitats, any threat (introduced rats, cats, dogs, etc), could have significant impacts on survival of tuatara due to its incredibly slow reproductive rate (Gaze 2001). In particular, these threats are greater for populations on small islands that have poor adaptability to buffer the impact of change. Moreover, small populations on small islands are more susceptible to loss of genetic variation (Gaze 2001).

Daugherty et al. (1990 b) argue that the potential future extinction of this species is possible, without appropriate species-level recognition. Therefore, to increase the survival of this species a translocation programme was funded to remove the individuals of *S. guntheri* from the current habitat on Brothers Island to Titi and Matui Islands (Gaze 2001). Even recently, the suggested species status of *S. guntheri* has been underlined by the proposal that global climate change threatens this "rare species" with extinction (Mitchell et al. 2010).

Recently, Hay et al. (2010) have called the specific status of *S. guntheri* into question. Their study based on genetic data from microsatellite and mitochondrial D-loop of modern tuatara populations indicates a remarkably low variation within these populations (Hay et al. 2010). The study concludes that "*Sphenodon*" is best described as a single species with distinctive and important geographic variants (Hay et al. 2010).

5.2 Purpose and Scope of the Project

In most cases, the approach used to the determination of species status in tuatara (Daugherty et al. 1990 b) has been based on a particular level of genetic divergence (Ayala 1975). In addition, because this particular taxonomic decision was based on minimal divergence in isozyme allele frequencies and because the presumptive species are geographically separated, a major molecular study of these taxa was conducted as part of my research project.

Furthermore, given the recent study (Hay et al. 2010) based on a small region of the mtDNA genome and the difficulties in determining species status of geographically separated populations, I conducted a comprehensive study of the mitochondrial diversity among modern populations of tuatara to reassess this conclusion using a large number of entire mitochondrial genomes of the tuatara. Specifically, I sequenced 42 complete mitochondrial genomes of tuatara from both presumptive species from eight island groups (Table 5.1) to further explore the evidence that *S. guntheri* is specifically distinct.

Table 5.1: The mtDNA recovery length of modern tuatara sampled from eight island groups off the coast of New Zealand

Island Group	Island	Sample ID	mtDNA Recovery Length (bp)
		FT4244	15,154
		FT4251	15,201
		FT4253	15,110
		FT4254	14,650
		FT4850	15,200
		FT4859	15,202
		FT4889	15,193
		FT4906	15,058
		FT4911	15,202
		FT4913	15,157
Stephens	Stephens	FT4915	15,198
		FT4926	15,033
		FT4961	15,140
		FT5002	15,178
		FT5044	15,089
		FT5093	15,131
		FT5098	15,178
		FT5099	15,015
		FT5102	15,177
		FT4994	15,155
		FT5113	15,090
		CD1333	15,153
Poor Knights	Tawhiti Rahi	FT2522	15,193
Hen & Chickens	Hen	FT2409	15,061
		FT7200	15,169
		FT7205	15,087
		FT7613	14,743
Little Barrier	Little Barrier	FT7625	15,180
		FT7896	14,991
		FT7899	14,899
		FT7930	15,019
Cuvier	Cuvier	FT2917	15,170
	Green Mercury	FT2391	15,163
		FT2392	15,076
Mercury	Middle Mercury	FT2365	15,088
		FT2369	15,125
	Red Mercury	FT2654	15,210
	Stanley	FT2704	15,147
(Motunau)	Motunau	FT2139	14,933
		FT200	15,147
Brothers	North Brother	FT210	15,193
		FT205	15,040

5.3 Analytical Methods

I compiled a data set consisting of 42 complete modern tuatara mtDNA genome sequences (mean length of 14,650±15,210 bp), representing tuatara samples belonging to eight island groups around New Zealand (Table 5.1). The complete mitochondrial genomes were aligned to a tuatara genome obtained from Genbank (Accession NC_004815) using BioEdit software version 7.0 (Hall 1999). The approximately 15,000

bp of aligned sequences were used to infer the phylogenetic relationship among tuatara (Figure 5.1). A Neighbor-joining tree was constructed using a maximum composite likelihood model, implemented in the software MEGA (Tamura et al. 2007) using the complete site deletion option (Figure 5.1.a). I also examined the phylogenetic relationship among tuatara genomes using maximum parsimony criteria in conjunction with an heuristic search and stepwise addition options of PAUP (Swofford 2003) (Figure 5.1.b). For the maximum likelihood analysis I initially obtained the best model of evolution and other parameters using Modeltest (Posada & Crandell 1998) based on Akaike information criterion. These parameters were used to obtain the maximum likelihood tree generated with PAUP (Figure 5.1.c). Boot-strap re-sampling (50 replicates) was performed to judge the strength of support for each node of the phylogenetic tree. Data for the maximum likelihood analysis using nuclear copies of mitochondrial Cyt b gene were obtained from Hay et al. (2004).

5.4 Results

The data clearly illustrate that '*S. guntheri*' from Brother Island do not form a separate clade from the remaining *S. punctatus* sequences and in fact they cluster within the populations of *S. punctatus*. Other phylogenetic analyses of these data, including minimum evolution and maximum parsimony methods using the entire genomes or individual regions yielded similar results and the bootstrap analyses clearly does not support a distinct clade for the tuatara mitogenomes from Brother islands. Therefore it seems reasonable to conclude that '*S. guntheri*' cannot be a distinct species in a biological sense based on these new data.
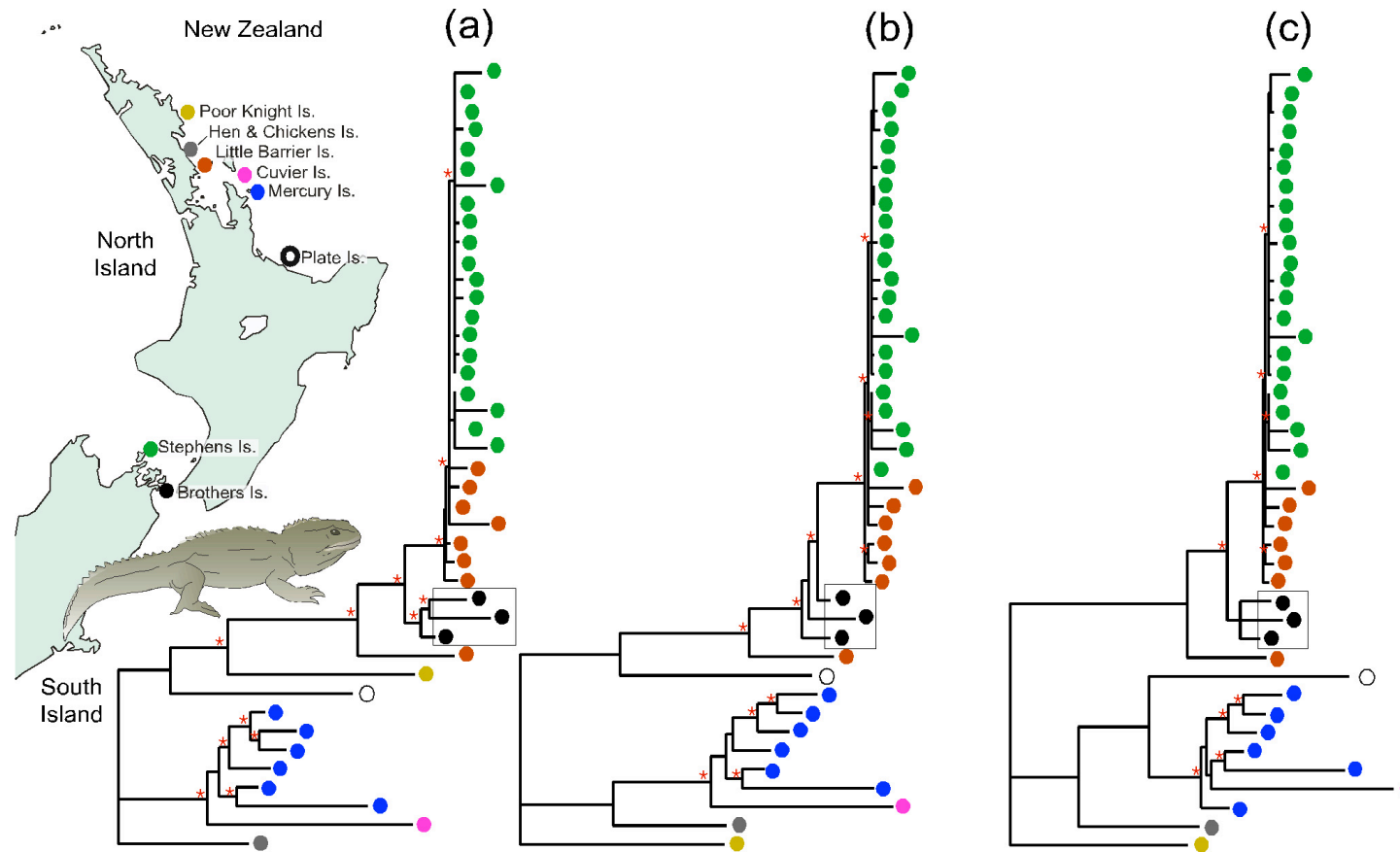
Figure 5.1: Phylogenetic relationship among modern tuatara populations from different islands off the coast of New Zealand derived from complete mtDNA genome sequences. The phylogenetic trees among the tuatara mtNA genomes were inferred based on **(a)** Neighbor-joining, **(b)** maximum parsimony and **(c)** maximum likelihood methods. Asterisks indicate the boot-strap values greater than 50 (replicates) indicating support for each node of the tree. Samples from tuatara individuals previously regarded as *S. punctatus* are shown (coloured circles), together with those from the putative second species *S. guntheri* from Brothers Island (black circles). The blank circles depict the tuatara genome from genebank (Accession NC_004815). The figure is drawn by the Author of the thesis.

Furthermore, I conducted a maximum likelihood analysis by constraining '*S. guntheri*' (samples from Brothers Island) and the remaining tuatara sequences as separate monophyletic groups. For this purpose Cyt b gene sequences were used together with their corresponding nuclear copies (Numt) as outgroups, as did Hay et al. (2004). This analysis revealed that constraining tuatara mtDNA genome sequences so that they belong to two distinct groups resulted in a log-likelihood score (1030.89) slightly less than that obtained (1031.91) when treating them as a single group, without any constraint.

5.4 Discussion

In 1990, the results and the conclusions regarding the new classification of *Sphenodon* reported by Daugherty et al. (1990 b) attracted immense international interest. In the same issue of *Nature* Robert May (1990) strongly supported this new classification. He remarked that because there were fewer than 300 individuals of this newly recognised species there were clear implications for its conservation management on North Brother Island.

The basis of Daugherty et al.'s (1990 b) conclusion was that *S. guntheri* populations were differentiated from other tuatara populations by differences in at least three of the 25 allozyme loci examined. They recommended that this population should be considered as a second species of tuatara. To support this finding, Daugherty et al. (1990 b) suggested that the recognition of two species of tuatara is consistent with both the evolutionary and phylogenetic species concepts (Frost & Hillis 1990). In addition, they suggested that the differentiation in allozymes and morphology of the proposed two species of tuatara was equivalent in magnitude to differences found between congeneric species pairs of New Zealand lizard (*Leiolopisma nigriplantare*). A clear pattern of genetic diversity coincided with considerable overlap in color variation among different 'populations' of *Leiolopisma nigriplantare* demonstrated that this endemic New Zealand lizard actually comprises of a group of four species whose differences in mate recognition are confirmed their species status by sympatry (Daugherty et al. 1990 a). Furthermore, Daugherty et al. (1990 a) indicated that the differentiation in allozymes found between these latter species is equivalent to that commonly found between congeneric species of non-avian vertebrates (Highton 1990).

The above rational can be questioned for a number of reasons. First, the allele

frequency data on which the genetic distance tree presented by Daugherty et al. (1990 b) was based were not reported in that original paper. Therefore, it was not possible to examine the data using other methods of tree construction in order to make a comparison. This is important, since, from a phylogenetic species perspective, the putative differentiation of *S. guntheri* needs to be robust under other phylogenetic analyses. Otherwise the evidence for specific status is not strong. Moreover, many authors have argued that, based on empirical evidence from a wide array of taxa, there appears to be no consistent relationship between genetic distance and species status (Lambert & Paterson 1982; Ferguson 2002). The practice of using genetic distance to infer species status is based on the assumption that, a particular level of genetic differentiation between two populations will inevitably lead to genetic incompatibility and 'reproductive isolation' (Sasa et al. 1998; Wu & Hollocher 1998). In contrast, in some cases it have been documented that taxa with high levels of genetic divergence still produce fertile hybrids e.g. for *Drosophila* (Zouros 1973, 1981), crickets (Howard 1983; Howard et al.1993), and frogs (Pyburn & Kennedy 1961; Mecham 1965).

On the other hand, in some species such as those belonging to the genus *Drosophila* (Hollocher et al. 1997 ab) and amphibia (Kuramoto 1984; Tilley et al.1990) pre-mating isolation mechanisms (e.g. behavioral and physical incompatibility) could prevent interbreeding between individuals of a population that is assigned as a single species. Therefore, an increase in genetic differences between sibling species does not necessarily result in reproductive isolation and speciation. Speciation is a distinctly different phenomenon that is driven by its own biological processes and genetic bases. There are a number of studies that show the genetic basis of post-mating isolation is more complex than pre-mating isolation (Zouros 1981; Coyne et al. 1998; Noor & Coyne 1996; Liu et al. 1996). Therefore, species with strong pre-mating isolation and weak post-mating isolation are likely to have minor genetic divergence since pre-mating isolation results from small genetic changes. Moreover, the reproductive mechanisms vary dramatically between taxa, therefore the predictive rule about degree of genetic divergence that lead to speciation is not realistic (Ferguson 2002). In general, the central fallacy behind the suggestion that species can be designated on the basis of any particular form of genetic differentiation is that it is clear that speciation is not necessarily a bi-product of any adaptive divergence of species themselves.

My comprehensive molecular data set from 42 complete mtDNA genomes of both presumptive species from a number of island groups indicate that *S. guntheri* is not a distinct species considering different level of genetic divergence, something that previous study (Daugherty et al. 1990 b) has failed to address. These findings support the tentative conclusions of a previous study based on a much smaller mitochondrial and microsatellite DNA dataset (Hay et al. 2010).

Finally, I agree with the suggestion by May (1990) that conservation biology and taxonomies are highly relevant because conservation priorities and managements are greatly influenced by a species taxonomic status. Taxonomic decisions need to be based on our understanding of the biological data, including genetics. However, it is important to examine those data carefully because poor taxonomic decisions can have negative consequences for rare and endangered species and can mislead management efforts.

Chapter Six

Estimating Evolutionary Rates in (*Sphenodon*: Reptilia) Mitochondrial DNA

6.1 Introduction

Tuatara have remained largely unchanged over long periods of evolution in terms of skeletal anatomy (Carroll 1988). In addition to the stable morphology, tuatara have some distinct physiological features that are uncommon among reptiles. They have low body temperature (they are active down to 5°C in the wild), slow rates of growth (50 cm in 35 years), a slow reproductive rate (at 2 to 5 - year intervals), a slow metabolism rate and long generation time (sexual maturity at 10 – 15 years) (Cree et al. 1990; Cree et al. 1991; Thompson & Daugherty 1998). All these traits suggest a slow rate of molecular evolution for this species over tens of million of year (Wu & WH 1985; Gillooly et al. 2005). In contrast to this view, Hay et al. (2008) provided evidence for relatively rapid molecular evolution in mitochondrial HVR regions of tuatara (Hay et al. 2008). This study was the first attempt to directly estimate the neutral evolutionary rates in a species with such unique physiological and behavioral features. Subsequently, Miller et al. (2008) argued that the high molecular rate reported in previous paper was an overestimate of the actual rate.

Miller et al. (2008) suggested a number of possible explanations for this high molecular rate in tuatara. These criticisms were the subject of a response by Subramanian et al. (2009 b) and they were reviewed in detail in chapter one of this thesis.

6.2 Purposes and Scope of the Project

Traditionally, rates of molecular evolution are estimated using comparative phylogenetic analyses of DNA sequences from different species, calibrated against points in the fossil record. Recent studies based on ancient DNA and pedigree methodologies suggest that rates of molecular evolution in vertebrates are much higher than those estimated using such phylogenetic/calibration approaches. Using ancient DNA technology, Hay et al. (2008) estimated the evolutionary rates for tuatara mtDNA relatively higher than other vertebrates. However, this estimate was based on HVR region of mtDNA, which is a small region of the entire mitochondrial genome.

It is been argued that the evolutionary rates for HVRs could not be used to infer the evolutionary rates for the entire genome since this region is evolving faster than the rest of the genome (Soares et al. 2009; Quinn TW 1992). It is been suggested that the evolutionary rates for synonymous regions of mitochondrial genome is an appropriate rate to infer the evolutionary rates for constrained sites such as tRNAs, rRNAs and non-synonymous nucleotide positions (Subramanian et al. 2009 a).

Therefore, this project aimed to recover the entire mtDNA of a large number of modern and ancient tuatara samples and then using an ancient DNA approach to estimate the evolutionary rates for the entire mitochondrial genome, as well as specific rates for rRNA, tRNA, synonymous and non-synosymous positions of the protein-coding regions.

6.2 Analytical Methods

6.2.1 Nucleotide Sequence Alignments, using CLUSTAL-W

The nucleotide sequences from rRNA, tRNA and D-loop regions of 42 modern and 11 ancient tuatara mtDNA were aligned in BioEdit software version 7.0 (Hall 1999) using CLUSTAL-W program. Using the alignment of amino acid sequences in MEGA software version 4.0 (Molecular Evolutionary Genetics Analysis) (Kumar et al. 2004;

Tamura et al. 2007) the cDNA alignments of the protein coding genes were created. Two ancient samples (SM1-SM28) were eliminated from the analytical data pool due to the lack of carbon dating information. Subsequently, these comprehensive data set comprising aligned sequences from 51 relatively complete tuatara mtDNA (recovery size ranges between 7,254-15,210 bp) were used for further analysis.

6.2.2 Constructing the Phylogenetic Relationship of Modern and Ancient Tuatara mtDNA, using sUPGMA, NJ and ME Methods

The phylogenetic relationship between modern and ancient tuatara populations derived by constructing the serial UPGMA tree (Drummond & Rodrigo 2000) using the software Pebble version 1.0 (Phylogenetics, Evolutionary Biology, and Bioinformatics in a molecular Environment) (Goode & Rodrigo 2004), which accounts the age of the serially sampled molecular sequences for the evolutionary analysis (Figure 6.2). Furthermore, using Neighbor-joining (NJ) (Saitou & Nei 1987) (Figure 6.3), sUPGMA (Figure 6.4), and minimum evolutionary methods (ME) (Figure 6.5) the phylogenetic relationship between modern and ancient populations of tuatara derived by MEGA software (Tamura et al. 2007).

6.2.3 Estimating the Molecular Evolutionary Rates in Tuatara mtDNA, using Baysian Statistics MCMC Method

Using complete mtDNA sequences from nine ancient and 42 modern tuatara samples, I estimated the overall mitochondrial genome evolutionary rates and specific rates for rRNA, tRNA, synonymous and non-synosymous positions of the protein-coding regions on tuatara mitochondrial genome. The evolutionary rates were estimated using the Baysian statistics based on Markov chain Monte Carlo (MCMC) (Drummond & Rambaut 2007) that accounts for various evolutionary parameters such as base composition, population size, rate and model of evolution implemented by the software BEAST version 1.5.3 (Baysian Evolutionary Analysis Sampling Trees (Drummond & Rambaut 2003). To model the pattern of nucleotide substitution the general time-reversible (GTR) and Hasegawa-Kishino-Yano (HKY) (Hasegawa et al. 1985) models were used. To account for the rate variation among different sites in the alignment I used the Gamma site heterogeneity model with six rate categories for the discrete approximation of the Gamma

distribution. In order to avoid any limitation due to the restricted prior assumptions wide uniform prior distributions were used as defaults. In addition, these wide uniform prior distributions enabled us to set up an upper and lower bound on parameters. The prior distribution of the Gamma shape parameter (alpha) that indicates the rate of variation among different sites modified from [0-1000] to [0-100]. The other priors for each parameter value were utilized as defaults. Both strict and relaxed molecular clock (exponential and lognormal) were used to estimate evolutionary rates under the constant and exponential population growth assumptions. For each analysis, posterior distributions of parameters were approximated by Markov chain Monte Carlo (MCMC) sampling. The MCMC was run for 10,000,000 generations with the initial 1,000,000 steps discarded as burn-in. Trees and model parameters were sampled every 1000 steps thereafter. The posterior distributions were examined using the software TRACER version 1.5 (Rambaut & Drummond, 2004) and the Effective Sample Sizes (ESS) for mean evolutionary rate, population size and posterior likelihood were found to be > 100 for all the models used. Moreover, convergence was assessed using TRACER (Rambaut & Drummond 2004). The results of this analysis are summarized in Table 6.7.

6.2.4 Estimating the Evolutionary Rates for Tuatara mtDNA, using sUPGMA Method

In addition to BEAST analysis, the evolutionary rates for synonymous, constrained sites (rRNA, tRNA, non-synonymous sites) and D-loop in mtDNA were estimated by sUPMGA method employed in the software Pebble. The HKY model of nucleotide substitution and Gamma site heterogeneity model with six rate categories was used. Moreover, Pebble was used to construct the sUPGMA tree for different regions of tuatara mtDNA of modern and ancient samples by accounting for the ages of the ancient samples (Figure 6.2 & 6.6-10). The results of this analysis are summarized in Table 6.8.

6.2.5 Estimating the Neutral Evolutionary Rates in Modern and Ancient Tuatara mtDNA

For estimating the neutral evolutionary rate, I initially analysed only the synonymous positions in protein-coding regions of modern and ancient tuatara mtDNA. Neutral evolutionary rates were estimated using the Baysian statistic-based Markov chain

Monte Carlo (MCMC) (Drummond & Rambaut 2007). To avoid any bias towards the method and the software, I also analysed the data by sUPGMA method implemented in software Pebble and estimated the synonymous rates of molecular evolution for tuatara mtDNA.

6.2.6 Estimating the Evolutionary Rates at Constrained Sites (rRNA, tRNA, Non-synonymous Sites) in Modern Tuatara mtDNA

For population data there is typically a low level of divergence in constrained sites. Therefore the rates of molecular evolution for such sites are prone to stochastic errors. However, we are able to estimate the average molecular rates of evolution at constrained sites using evolutionary rates for synonymous sites (μs). According to the neutral theory of molecular evolution, the rate of evolution (K) in constrained sites is equal to mutation rate (μ) times to the fraction of neutral positions (f0), (K= μ. f0) (Kimura 1983). The rate of evolution in synonymous sites (μs) is similar to mutation rate (μ) and the divergence at synonymous sites (ds) reflects the accumulation of neutral mutations over time (Yang & Nielsen 2008), therefore the evolutionary rates in constrained sites (Kx) could be estimated using the expression Kx= μs (dx/ds), where dx is the divergence in constrained site, and the ratio dx/ds is the fraction of neutral positions. The ratio of the mean pairwise divergences (dx/ds) for different regions of mtDNA was calculated by MEGA, under the complete deletion method that excludes the nucleotide positions that contained an alignment gap in one or more sequences. In addition, the rate variations among different sites were considered by using Gamma site heterogeneity model with six rate categories for the discrete approximation of the Gamma distribution.

6.2.7 Estimating the mtNA Evolutionary Rates, using PAML and MEGA

In addition to the rates were obtained by BEAST, the evolutionary rates for the entire mtDNA were estimated with a joint analysis by PAML44 (Phylogenetic Analysis by Maximum Likelihood) version 4.3 (Yang 2007) and MEGA software (Tamura et al. 2007). Using MEGA, the Neighbour joining phylogenetic tree were constructed based on complete modern and ancient tuatara mtDNA sequences. Then the outcome tree constructed by MEGA was used as a tree file in default control file (baseml.ctl) for baseml program in PAML software. The baseml program is for maximum likelihood analysis of

nucleotide sequences. Using Baseml, the name of seqfile, tree file and outfile were specified. There are various options for nucleotide substitution models in PAML. For this analysis the REV or GTR nucleotide substitution model (General-time-reversible) (Yang 1994; Zharkikh 1994) were chosen. The rateAncestor=1 was selected in baseml program while for the other parameter values the default were used. PAML-based analyses produce some phylogenetic trees that can be used as an ancestral tree for all the modern and ancient sequences. Then, an input file including all modern, ancient and ancestral sequences were used for further analysis by MEGA. In MEGA, three groups as modern, ancient and ancestral were defined in order to compute the pairwise distance between them. In addition to that the mean distance between these groups were also calculated under the maximum composite likelihood model. The mean difference between the ancestral sequence and all the modern sequences were calculated. Moreover, the pairwise distance between ancient sequences and the ancestral sequence were calculated in MEGA. By knowing the age of the ancient samples, the evolutionary rates for the complete mtDNA of tuatara were estimated according to this calculation: $\sum [(X-Y)/Z]/N$, which N is representing the number of ancient samples used for this calculation, X is the overall mean between the ancestral sequence and modern samples, Y is the genetic distance pairwise between ancestral sequence and ancient samples and Z is the age of ancient samples (years before present). The evolutionary rates of modern and ancient tuatara mtDNA estimated by this method are presented in Table 6.9.

6.2.8 Patterns of Nucleotide Changes in Modern and Ancient Tuatara Populations

One of the major characteristics of aDNA is that this DNA is usually modified as a result of oxidative damage or hydrolytic processes (Höss et al. 1996). Hydrolytic damage mostly results in C/G to T/A transitions. Unique to aDNA, the observation of C to T transition is higher than G to A (Brotherton et al. 2007; Briggs et al., 2007; Binladen et al. 2006; Gilbert et al. 2003; Hofreiter et al. 2001; Millar et al. 2008 a).

To make sure that the nucleotide changes observed in ancient DNA sequences are not the results of DNA damage or sequencing Numts, different pattern of nucleotide changes in both ancient and modern tuatara populations were examined. Using MEGA (Tamura et al. 2007) a comparison of modern and ancient nucleotide compositions,

nucleotide pair frequencies, numbers and the types of mutations were performed. The results of this comparison are presented in Table 6.10. In addition, in order to test for any such aDNA damage in the ancient samples used in this study, I constructed separate neighbor joining trees for modern and ancient mtDNA with Tamura-Nei distances (Tamura & Nei 1993) using the software MEGA (Tamura et al. 2007). Using the tree topology generated by MEGA, a maximum likelihood analysis was performed by PAML software. This was accomplished by using the program baseml, which is for maximum likelihood analysis of nucleotide sequences. In this program I performed the analysis under general time reversible (GTR or REV) model (Yang 1994; Zharkikh 1994) that assumes rate variation among nucleotide positions. The results of this analysis are summarized in Table 6.11.

6.3 Results

6.3.1 Summary Statistics

A total of 11 ancient and 42 modern tuatara mtDNA were amplified and sequenced during this study. The age of the ancient samples measured from 696±30 to 5191±35 years BP (before present, i.e., 1950) by radiocarbon ($_{14}$C) dating approach. The age distribution of ancient samples used for this study is shown (Figure 6.1). The complete mtDNA recovery size ranges between 7,254-15,210 bp in length for the ancient and modern samples (Table 5.1 & 6.1). Summary statistics were estimated using MEGA version 4.0 (Tamura et al. 2007). The final data set obtained from the generated sequences consists of 9024 protein coding, 2445 rRNA, 1434 tRNA, and 1757 D-loop positions. For all the coding and non-coding regions the pairwise distance estimations were performed by MEGA (Table 6.2-6). The average length of 14723.7 bp of modern and ancient mtDNA genome showed asymmetric base frequencies (π_A=33.3%, π_T= 26.8%, π_C=25.9%, π_G=14.0%). The maximum likelihood composite estimate of the pattern of nucleotide substitution showed a strong bias towards transition mutations (frequencies C↔T=0.30, G↔A=0.56) over transversion mutation (0.14). The transition/transversion rate ratios were calculated as k1 = 17.734 (purines), k2 = 8.411 (pyrimidines) and the overall transition/transversion bias (R) as 4.3, where R = [A*G*k1 + T*C*k2]/[(A+G)*(T+C)]. All positions containing gaps and missing data were eliminated from the dataset (Complete-

deletion option). There were a total of 3,736 positions in the final dataset. There were 477 variable sites, of which 268 were informative for parsimony and 209 singletons. Interpopulation Kimura-2 distances ranged from 0 to 0.01, with no particular trend among samples.
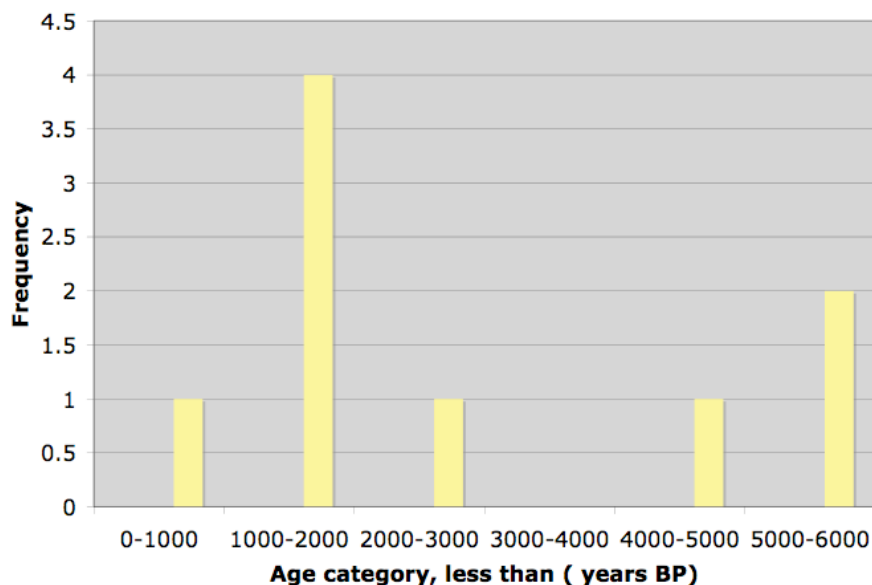
Figure 6.1: The frequency distribution of the age of nine ancient samples used for this study

Table 6.1: The mtDNA recovery length of ancient tuatara samples, used in this study

Sample ID, (Age yrBP)	mtDNA length (bp)	Sample ID, (Age yrBP)	mtDNA length (bp)
T31 (1,684 ± 35)	7,254	T110 (1,794 ± 35)	13,236
T91 (696 ± 35)	14,701	T102 (1,546 ± 35)	14,598
T87 (3,145 ± 35)	14,188	T151 (2,321 ± 35)	14,941
T106 (5,191 ± 35)	9,941	T60=T71	14,347
T82 (5,191 ± 35)	13,150	T62	12,621
T142 (1,189 ± 35)	14,637		

Table 6.2: The mean distance between tuatara populations (modern and ancient) based on protein coding regions of mtDNA. The distance (Kimura-2 parameter), calculated with software MEGA 4.0 with pairwise deletion option and rate heterogeneity among sites. Substitutions include both transition and transversion. The number of individuals for each island is shown with n.

Island	n	TRA	Hen	LBI	Cuv	Gmer	Mmer	Rmer	STA	BOP	STP	NBI	T31	T82	T87	T91	T102	T106	T110	T142	T151
TRA	1																				
Hen	1	0.0057																			
LBI	7	0.0098	0.0095																		
Cuv	1	0.008	0.0035	0.01																	
Gmer	2	0.0082	0.0037	0.0103	0.0013																
Mmer	2	0.0083	0.004	0.0111	0.0022	0.0017															
Rmer	1	0.0076	0.0034	0.0102	0.0018	0.0014	0.0017														
STA	1	0.0084	0.0041	0.0115	0.0023	0.0019	0.0007	0.0019													
BOP	1	0.0082	0.0057	0.0091	0.0046	0.0045	0.0049	0.0049	0.0051												
STP	22	0.0097	0.0096	0.0008	0.0101	0.0103	0.0112	0.0103	0.0117	0.009											
NBI	3	0.0094	0.0092	0.002	0.0098	0.01	0.0108	0.0098	0.0111	0.0087	0.0017										
T31	1	0.007	0.0081	0.0095	0.0091	0.0092	0.0093	0.0091	0.0099	0.0078	0.009	0.0091									
T82	1	0.0098	0.0096	0.0028	0.0102	0.0103	0.0111	0.01	0.0113	0.0093	0.0027	0.0027	0.0093								
T87	1	0.0086	0.0086	0.0032	0.0093	0.0091	0.01	0.0089	0.0102	0.0087	0.0031	0.0033	0.0085	0.0034							
T91	1	0.0084	0.008	0.0094	0.0078	0.0079	0.008	0.0074	0.0082	0.0074	0.0093	0.0091	0.0068	0.0088	0.0078						
T102	1	0.0091	0.0088	0.0073	0.0092	0.0094	0.0097	0.0087	0.01	0.0084	0.0074	0.0069	0.0087	0.0073	0.0062	0.0081					
T106	1	0.0086	0.0086	0.0027	0.0088	0.0087	0.0095	0.0082	0.0098	0.008	0.0023	0.0021	0.0079	0.002	0.0028	0.0074	0.0049				
T110	1	0.0093	0.0087	0.0061	0.0096	0.0098	0.0101	0.0092	0.0104	0.0085	0.006	0.0059	0.0082	0.0062	0.0061	0.0086	0.0022	0.0042			
T142	1	0.0094	0.0091	0.0035	0.01	0.0101	0.0106	0.0098	0.0109	0.0095	0.0035	0.0032	0.0082	0.0033	0.0034	0.0081	0.0068	0.0023	0.0066		
T151	1	0.0092	0.0089	0.0075	0.0094	0.0095	0.0098	0.0089	0.0101	0.0086	0.0076	0.0071	0.0089	0.0074	0.0063	0.0084	0.0012	0.0054	0.0027	0.0027	

Table 6.3: The mean distance between tuatara populations (modern and ancient) based tRNA regions on mtDNA. The distance (Kimura-2 parameter), calculated with software MEGA 4.0 with pairwise deletion option and rate heterogeneity among sites. Substitutions include both transition and transversion. The number of individuals for each island is shown with n.

Island	n	TRA	Hen	LBI	Cuv	Gmer	Mmer	Rmer	STA	BOP	STP	NBI	T31	T82	T87	T91	T102	T106	T110	T142	T151
TRA	1																				
Hen	1	0.0084																			
LBI	7	0.0084	0.007																		
Cuv	1	0.0113	0.0035	0.0083																	
Gmer	2	0.0109	0.0039	0.0101	0.0018																
Mmer	2	0.0092	0.0032	0.009	0.0025	0.0021															
Rmer	1	0.0099	0.0035	0.0097	0.0014	0.001	0.001														
STA	1	0.0099	0.0028	0.0097	0.0021	0.0017	0.0003	0.0007													
BOP	1	0.0084	0.0035	0.0095	0.0056	0.0053	0.0039	0.0042	0.0035												
STP	22	0.0085	0.007	0.0007	0.0085	0.0103	0.0092	0.0099	0.0099	0.0099											
NBI	3	0.0075	0.0061	0.0018	0.0075	0.0093	0.0082	0.0089	0.0089	0.0089	0.0014										
T31	1	0.0083	0.0062	0.0083	0.0104	0.0114	0.0124	0.0124	0.0124	0.0145	0.0083	0.0069									
T82	1	0.0066	0.0059	0.0039	0.0059	0.0078	0.0067	0.0074	0.0074	0.0074	0.0044	0.0034	0.0083								
T87	1	0.009	0.0082	0.0065	0.0098	0.0116	0.0105	0.0113	0.0112	0.0105	0.0066	0.0057	0.0041	0.0047							
T91	1	0.0092	0.0063	0.0105	0.0084	0.0081	0.0063	0.007	0.0063	0.0063	0.0106	0.0096	0.0083	0.0089	0.0097						
T102	1	0.0099	0.0035	0.0085	0.0056	0.006	0.0064	0.0056	0.0063	0.0063	0.0085	0.0075	0.0124	0.0074	0.0098	0.0085					
T106	1	0.0107	0.0068	0.0073	0.0077	0.0102	0.0087	0.0097	0.0087	0.0077	0.0078	0.0061	0.0094	0.0041	0.0084	0.0087	0.0107				
T110	1	0.0092	0.0042	0.0072	0.0064	0.0067	0.0071	0.0063	0.0071	0.0071	0.0071	0.0061	0.0104	0.0067	0.009	0.0078	0.0007	0.0098			
T142	1	0.0081	0.0051	0.0053	0.0073	0.0092	0.0074	0.0088	0.0081	0.0081	0.0051	0.0051	0.0021	0.0046	0.0078	0.0088	0.0073	0.0077	0.0066		
T151	1	0.0092	0.0042	0.0076	0.007	0.0074	0.007	0.007	0.007	0.007	0.0078	0.0068	0.0083	0.0066	0.009	0.0077	0.0028	0.0097	0.0021	0.0066	

Table 6.4: The mean distance between tuatara populations (modern and ancient) based rRNA regions on mtDNA. The distance (Kimura-2 parameter), calculated with software MEGA 4.0 with pairwise deletion option and rate heterogeneity among sites. Substitutions include both transition and transversion. The number of individuals for each island is shown with n.

Island	n	TRA	Hen	LBI	Cuv	Gmer	Mmer	Rmer	STA	BOP	STP	NBI	T31	T82	T87	T91	T102	T106	T110	T142	T151
TRA	1																				
Hen	1	0.0021																			
LBI	7	0.0022	0.0035																		
Cuv	1	0.0033	0.0046	0.004																	
Gmer	2	0.0033	0.0046	0.0039	0.0008																
Mmer	2	0.0031	0.004	0.0038	0.001	0.001															
Rmer	1	0.0041	0.0054	0.0047	0.0008	0.0008	0.001														
STA	1	0.0054	0.0066	0.0058	0.0029	0.0029	0.0023	0.0029													
BOP	1	0.0037	0.005	0.0028	0.0033	0.0025	0.0027	0.0033	0.0046												
STP	22	0.0017	0.0029	0.0006	0.0042	0.0041	0.004	0.005	0.0061	0.0029											
NBI	3	0.0032	0.0044	0.0014	0.004	0.0032	0.0034	0.004	0.0049	0.0023	0.0016										
T31	1	0.0072	0.0093	0.0088	0.0051	0.0051	0.0042	0.0041	0.0082	0.0073	0.0094	0.0072									
T82	1	0.0021	0.0037	0.0011	0.0016	0.0018	0.0013	0.0021	0.0037	0.0016	0.0011	0.0012	0.0047								
T87	1	0.0044	0.0058	0.0037	0.0044	0.0039	0.0038	0.0044	0.0057	0.0026	0.0036	0.0031	0.0072	0.0016							
T91	1	0.0039	0.0053	0.0042	0.0031	0.0031	0.0024	0.0031	0.0044	0.004	0.004	0.0038	0.0051	0.0016	0.0039						
T102	1	0.0069	0.0084	0.0068	0.0064	0.0059	0.0057	0.0064	0.0079	0.0054	0.0069	0.0061	0.0103	0.0047	0.0064	0.0069					
T106	1	0.0037	0.0037	0.0032	0.0037	0.003	0.0025	0.0036	0.003	0.0031	0.003	0.0026	0.005	0.0008	0.0024	0.0024	0.0064				
T110	1	0.0063	0.0069	0.0052	0.0068	0.0063	0.0062	0.0068	0.0083	0.0049	0.0054	0.0045	0.0082	0.0034	0.0052	0.0052	0.0016	0.005			
T142	1	0.0042	0.0056	0.0035	0.0032	0.0032	0.0026	0.0032	0.0046	0.0033	0.0033	0.0031	0.0061	0.0006	0.0028	0.0028	0.0057	0.0018	0.0044		
T151	1	0.0057	0.0071	0.006	0.0057	0.0053	0.0051	0.0057	0.007	0.0049	0.0058	0.0055	0.0103	0.0042	0.0057	0.0061	0.0024	0.0049	0.0031	0.0051	

Table 6.5: The mean distance between tuatara populations (modern and ancient) based on D-loop(s) regions of mtDNA. The distance (Kimura-2 parameter), calculated with software MEGA 4.0 with pairwise deletion option and rate heterogeneity among sites. Substitutions include both transition and transversion. The number of individuals for each island is shown with n.

Island	n	TRA	Hen	LBI	Cuv	Gmer	Mmer	Rmer	STA	BOP	STP	NBI	T31	T82	T87	T91	T102	T106	T110	T142	T151
TRA	1																				
Hen	1	0.0054																			
LBI	7	0.0073	0.0069																		
Cuv	1	0.0122	0.0061	0.0116																	
Gmer	2	0.012	0.0083	0.0124	0.0061																
Mmer	2	0.0099	0.0073	0.0075	0.0064	0.0061															
Rmer	1	0.0104	0.0085	0.0114	0.0069	0.0061	0.0058														
STA	1	0.0084	0.0056	0.0087	0.0036	0.0033	0.003	0.003													
BOP	1	0.008	0.0083	0.0029	0.0093	0.0107	0.0062	0.0093	0.0064												
STP	22	0.007	0.0066	0.0009	0.0116	0.0126	0.0075	0.0116	0.0089	0.0024											
NBI	3	0.0093	0.0092	0.0032	0.0131	0.015	0.01	0.0138	0.011	0.0052	0.0025										
T31	1	0	0	0.0022	0.0053	0.0035	0.0035	0.0052	0.0019	0.0018	0.0018	0.0035									
T82	1	0.0138	0.013	0.0081	0.019	0.0234	0.0164	0.0202	0.0161	0.0077	0.0088	0.0113	0.0095								
T87	1	0.0112	0.0103	0.008	0.0128	0.014	0.0123	0.0134	0.0101	0.0073	0.0082	0.0095	0.0054	0.0124							
T91	1	0.0111	0.0076	0.012	0.0125	0.0123	0.0138	0.0138	0.0115	0.0106	0.0119	0.0145	0.0017	0.0188	0.0119						
T102	1	0.0138	0.0125	0.0104	0.0173	0.0177	0.0145	0.0158	0.0129	0.0099	0.0105	0.0114	0.0104	0.0137	0.0119	0.0172					
T106	1	0.0157	0.0178	0.0086	0.0147	0.0188	0.0123	0.0179	0.013	0.0059	0.0089	0.0079	0.0058	0.0131	0.0103	0.0179	0.0179				
T110	1	0.0122	0.0098	0.01	0.0174	0.0189	0.0153	0.0153	0.0128	0.0075	0.0101	0.0118	0.0094	0.0115	0.0148	0.0173	0.003	0.0183			
T142	1	0.0066	0.009	0.0038	0.0073	0.0084	0.0053	0.008	0.0041	0.0036	0.004	0.0051	0.0087	0.0064	0.0045	0.0113	0.0093	0.0089	0.0082		
T151	1	0.0143	0.013	0.0111	0.0168	0.0173	0.0143	0.0155	0.0128	0.0099	0.0111	0.012	0.0122	0.0137	0.0134	0.0185	0.0007	0.019	0.003	0.0106	

Table 6.6: The mean distance between tuatara populations (modern and ancient) based complete mtDNA genome. The distance (Kimura-2 parameter), calculated with software MEGA 4.0 with pairwise deletion option and rate heterogeneity among sites. Substitutions include both transition and transversion. The number of individuals for each island is shown with n.

Island	*n*	TRA	Hen	LBI	Cuv	Gmer	Mmer	Rmer	STA	BOP	STP	NBI	T31	T82	T87	T91	T102	T106	T110	T142	T151
TRA	1																				
Hen	1	0.0057																			
LBI	7	0.0086	0.0085																		
Cuv	1	0.0082	0.0039	0.0095																	
Gmer	2	0.0082	0.0043	0.0099	0.0018																
Mmer	2	0.0079	0.0042	0.0097	0.0025	0.0021															
Rmer	1	0.0078	0.0043	0.01	0.0022	0.0018	0.0021														
STA	1	0.0081	0.0046	0.0105	0.0025	0.0021	0.0012	0.0021													
BOP	1	0.0078	0.0064	0.0075	0.0055	0.0054	0.0051	0.0056	0.0055												
STP	22	0.0084	0.0084	0.0008	0.0095	0.01	0.0098	0.0101	0.0107	0.0075											
NBI	3	0.0086	0.0086	0.002	0.0094	0.0098	0.0096	0.0098	0.0103	0.0075	0.0018										
T31	1	0.0071	0.008	0.0088	0.0088	0.0088	0.009	0.0089	0.0098	0.0082	0.0085	0.0084									
T82	1	0.0088	0.0088	0.003	0.0092	0.0099	0.0097	0.0096	0.0103	0.008	0.0031	0.0031	0.009								
T87	1	0.0086	0.0088	0.0042	0.0093	0.0094	0.0097	0.0094	0.0101	0.0078	0.0042	0.0043	0.008	0.004							
T91	1	0.0082	0.0075	0.0091	0.0078	0.0079	0.0078	0.0077	0.0081	0.0075	0.009	0.009	0.007	0.0082	0.0081						
T102	1	0.0096	0.009	0.0079	0.0096	0.0097	0.0097	0.0093	0.0099	0.0082	0.008	0.0075	0.0098	0.0074	0.0072	0.009					
T106	1	0.0089	0.0089	0.0038	0.0088	0.0092	0.009	0.009	0.0093	0.0071	0.0037	0.0032	0.008	0.0026	0.0043	0.0078	0.007				
T110	1	0.0097	0.0087	0.0067	0.01	0.0103	0.0102	0.0098	0.0106	0.0081	0.0066	0.0065	0.0091	0.0064	0.0071	0.009	0.0023	0.0064			
T142	1	0.0086	0.0087	0.0038	0.0089	0.0092	0.0089	0.009	0.0094	0.0081	0.0038	0.0037	0.0079	0.0031	0.0041	0.0079	0.0072	0.0034	0.0069		
T151	1	0.0097	0.0092	0.0081	0.0098	0.01	0.0099	0.0095	0.0101	0.0084	0.0081	0.0078	0.01	0.0075	0.0075	0.0093	0.0016	0.0073	0.0031	0.0075	

6.3.2 The Phylogenetic Relationship of Modern and Ancient Tuatara Mitochondrial Genome

Using Pebble and MEGA software, the phylogenetic relationship between modern and ancient tuatara mtDNA were examined by sUPGMA, NJ and ME methods. The sUPGMA tree, show the separation of northern and all Cook Strait tuatara with some mixing from Little Barrier tuatara with Cook Strait ones. In addition, within Cook Strait, North Brother tuatara individuals form a group. The structure of the northern tuatara follows island groups but with some mixing between Cuvier and Mercury Islands groups. The northland ancient samples (T31, T91) are grouped with modern northern tuatara and the southland ancient samples are grouped with Cook Strait and Little Barrier Island samples. Although T87 geographically belongs to eastern north island, phylogenetically it groups with samples from western southland. The sUPGMA, NJ and ME trees constructed by MEGA show similar structure pattern among modern and ancient tuatara individuals.
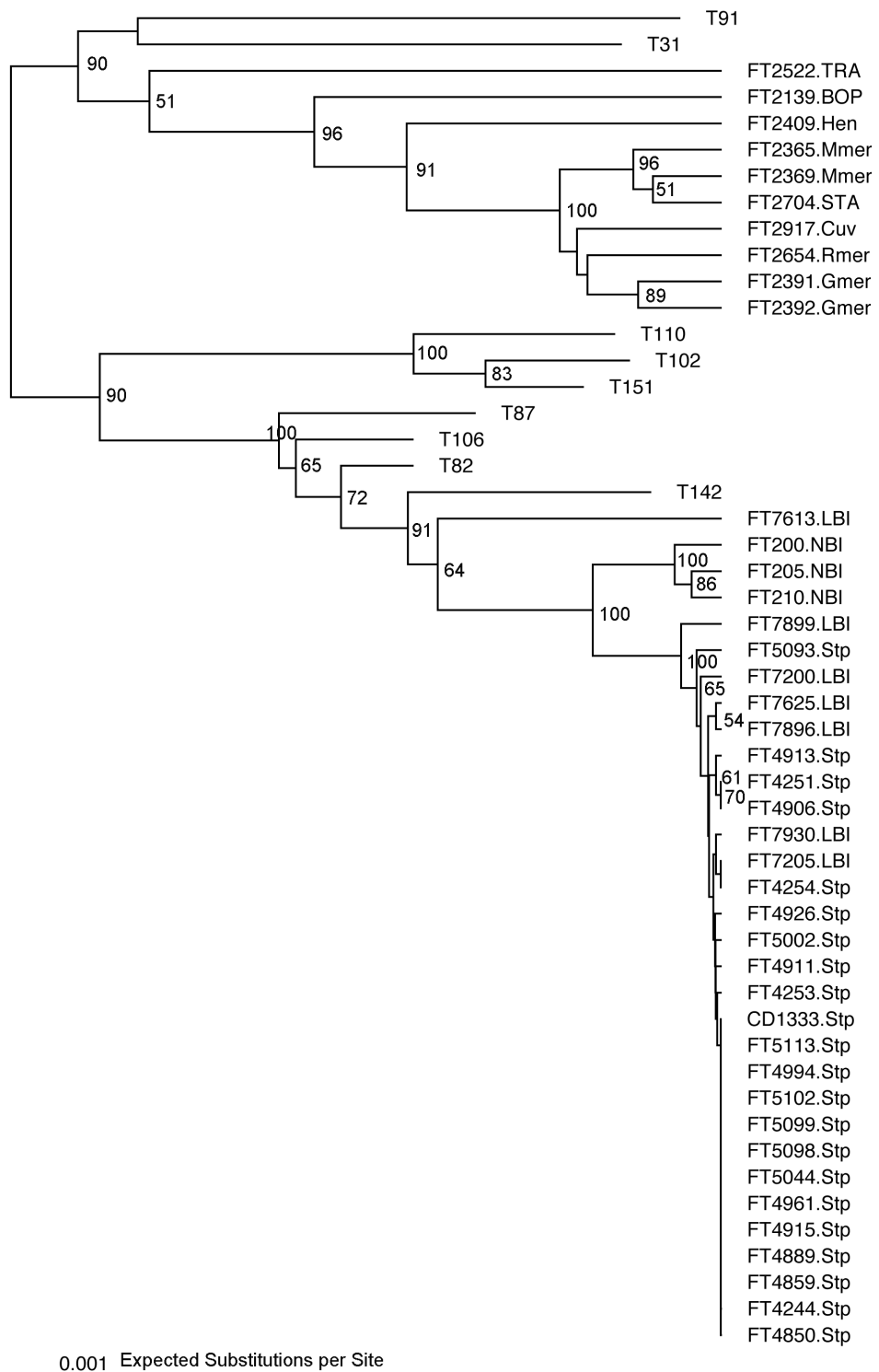
Figure 6.2: Serial UPGMA tree based on complete mtDNA of modern and ancient tuatara samples. The phylogeny tree was constructed, using the software Pebble under HKY and Gamma site heterogeneity model. The geographical locations of the samples are depicted on the map (Figure 4.3). Support for nodes on branches are bootstrap confidence levels (BCL) ×1,000 replications.
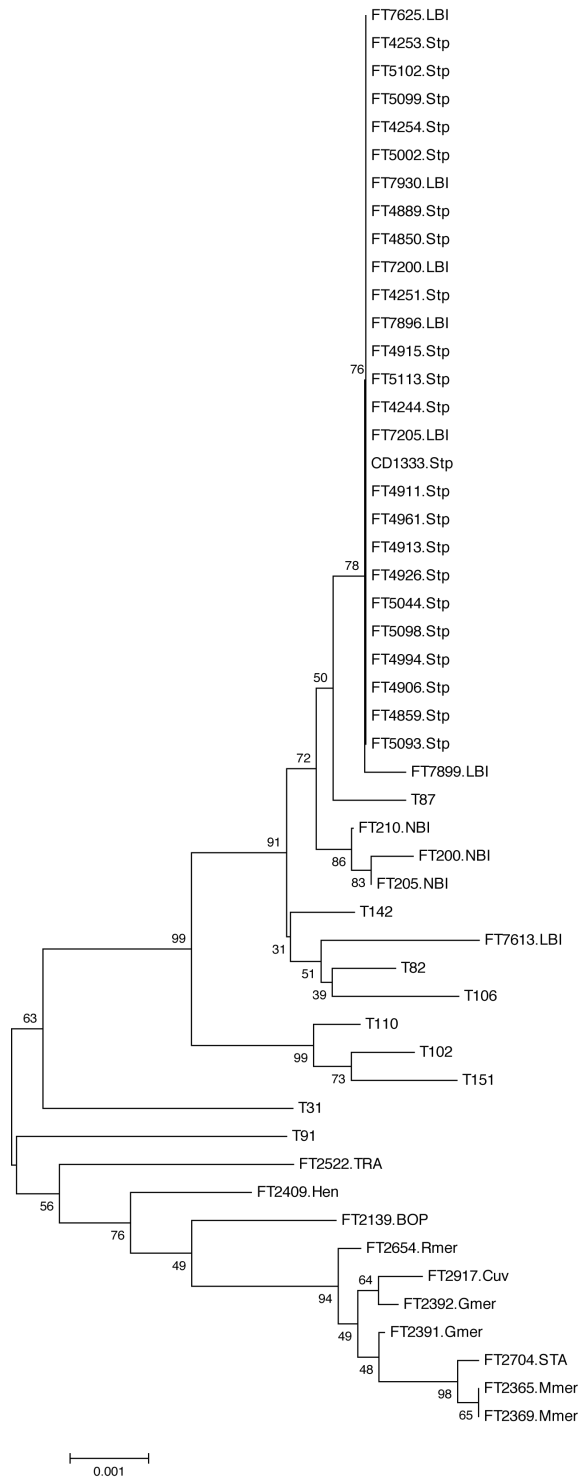
Figure 6.3: The bootstrap Neighbor-joining tree based on complete mtDNA sequences from modern and ancient tuatara populations, constructed by the software MEGA (Tamura et al. 2007). The geographical locations of the samples are depicted on the map (Figure 4.3). Support for nodes on branches are bootstrap confidence levels (BCL) ×1,000 replications.
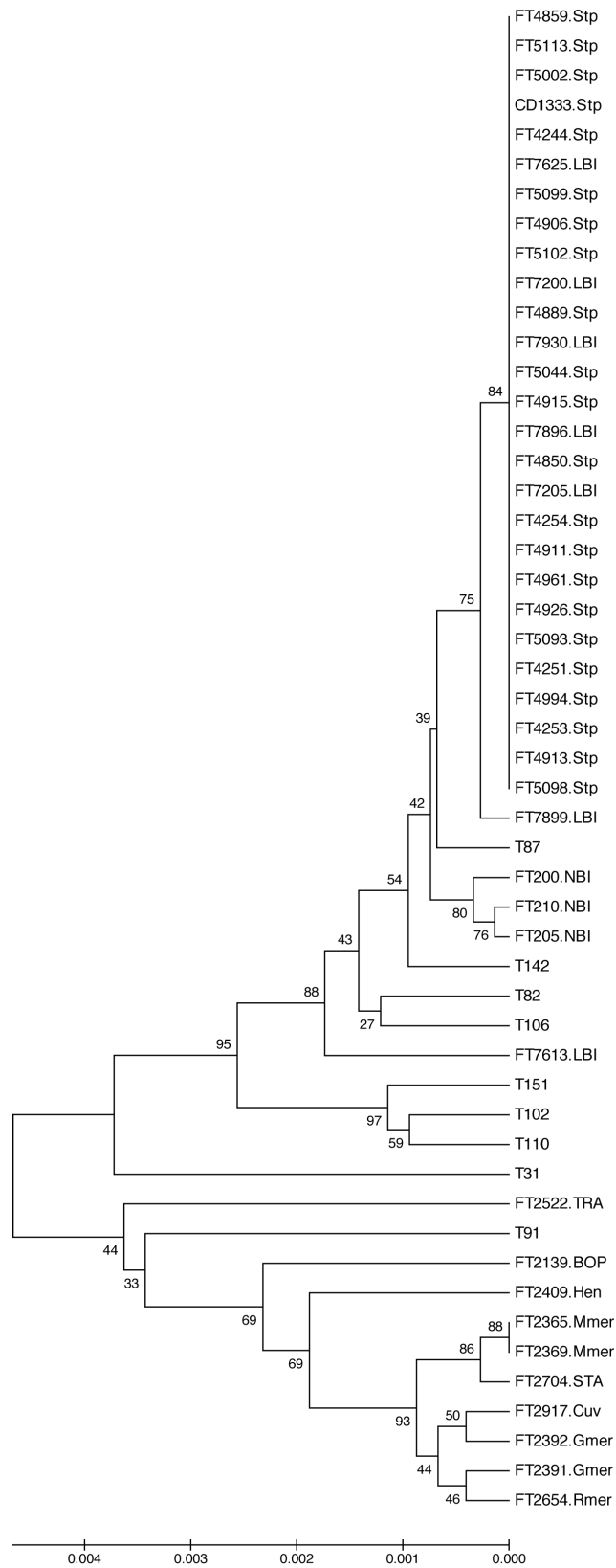
Figure 6.4: The bootstrap sUPGMA tree based on complete mtDNA sequences from modern and ancient tuatara populations, constructed by the software MEGA (Tamura et al. 2007). The geographical locations of the samples are depicted on the map (Figure 4.3). Support for nodes on branches are bootstrap confidence levels (BCL) ×1,000 replications.
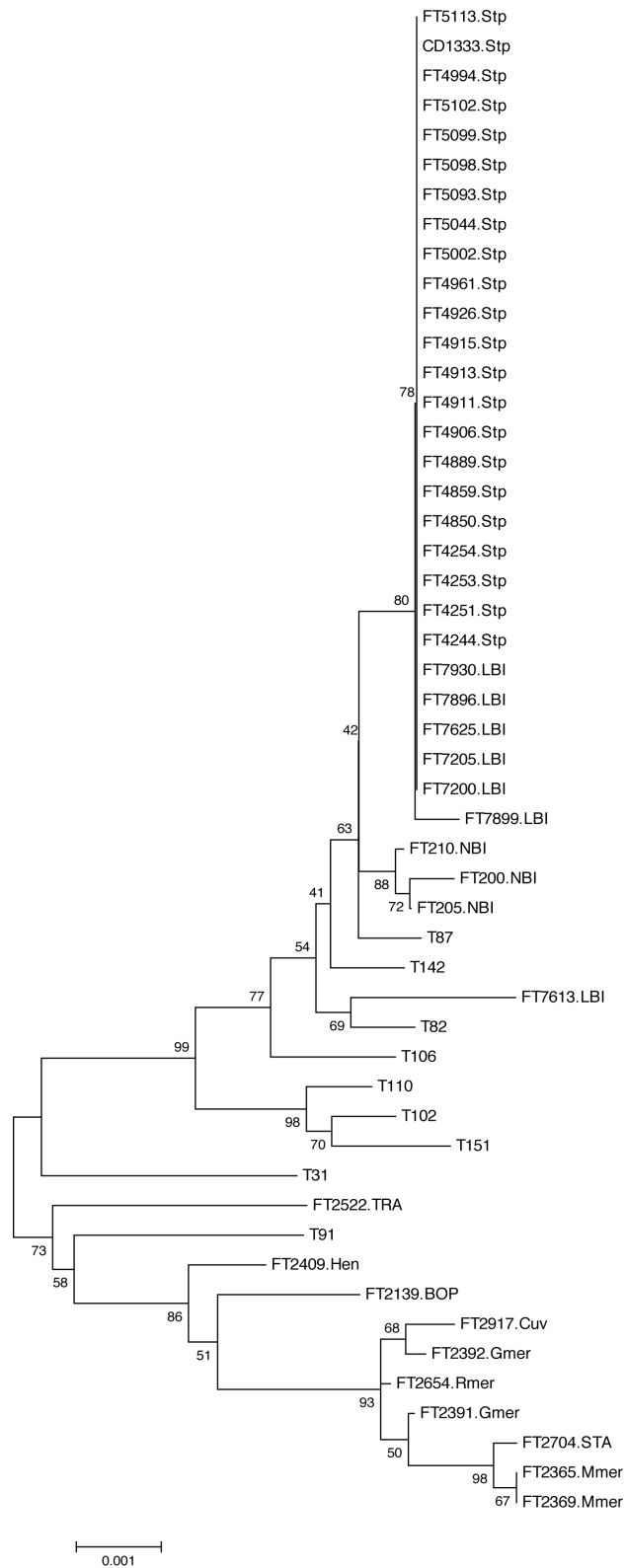
Figure 6.5: The bootstrap minimum evolutionary tree based on complete mtDNA sequences from modern and ancient tuatara populations, constructed by the software MEGA (Tamura et al. 2007). The geographical locations of the samples are depicted on the map (Figure 4.3). Support for nodes on branches are bootstrap confidence levels (BCL) ×1,000 replications.

6.3.3 Rates of Molecular Evolution in Tuatara Mitochondrial Genome Estimated, using Baysian Statistical MCMC Method

Using BEAST, the evolutionary rates for different regions of mtDNA were estimated based on Baysian statistical MCMC method as detailed in section 6.2.3. In this analysis different molecular clock (strict and uncorrelated relaxed lognormal/exponential), growth models (constant and exponential) and nucleotide substitution models (GTR and HKY) were used. According to these estimates, the third nucleotide position in protein coding regions has the highest molecular rates while rRNA is evolving with the slowest rates in comparison to the other regions. There is no significant difference between the rates estimated considering GTR and HKY nucleotide substitution models. However, introducing the exponential model of rate change into the analysis results in almost three times higher rates compared to lognormal model of change. In summary, the evolutionary rates estimated were: 0.33 s/s/Myr (HPD: 95% CI: 0.19 - 0.47) for third codon positions (synonymous), 0.30 s/s/Myr (HPD: 95% CI: 0.16 - 0.44) for D-loop, 0.19 s/s/Myr (HPD: 95% CI: 0.10 - 0.28) for tRNA, 0.15 s/s/My (HPD: 95% CI: 0.09 - 0.22) for protein coding (non-synonymous) and 0.13 s/s/Myr (HPD: 95% CI: 0.07 - 0.19) for rRNA regions of tuatara mtDNA under constant population size, uncorrelated log normal relaxed clock and GTR model of nucleotide substitution. The overall mtDNA evolutionary rate was estimated as 0.17 s/s/Myr (HPD: 95% CI: 0.10 - 0.25) with the same models.

In addition, I estimated the root heights of the most recent common ancestor (μt_{MRCA}, substitutions per site) of tuatara using complete mitochondrial genome. This parameter represents the total height of the tree that was estimated 35,391 years (35 Ky) (95% HPD 22,462 – 50,323) for modern and ancient tuatara mitochondrial genome, using Baysian statistic-based MCMC under constant population size and GTR model of nucleotide substitution.

Table 6.7: Evolutionary rate estimates (s/s/Myr) for different regions of mtDNA, using different molecular clock and growth models.

| Molecular clock model | Constant population growth | | | | | | | | | | | |
| | Third position | | Coding region | | tRNA | | rRNA | | D-loop | | mtDNA | |
	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval
HKY												
Strict clock model	0.31	0.17 - 0.45	0.14	0.08 - 0.20	0.17	0.09 - 0.26	0.12	0.06 - 0.17	0.28	0.15 - 0.42	0.16	0.10 - 0.23
Uncorrelated exponential relaxed clock	0.97	0.43 - 1.75	0.44	0.20 - 0.80	0.54	0.22 - 0.99	0.37	0.16 - 0.68	0.85	0.35 - 1.54	0.51	0.23 - 0.91
Uncorrelated log normal relaxed clock	0.34	0.19 - 0.51	0.15	0.09 - 0.23	0.19	0.10 - 0.29	0.13	0.07 - 0.19	0.31	0.17 - 0.46	0.18	0.10 - 0.26
GTR												
Strict clock model	0.30	0.17 - 0.44	0.14	0.07 - 0.20	0.17	0.08 - 0.26	0.11	0.06 - 0.17	0.28	0.15 - 0.42	0.16	0.09 - 0.23
Uncorrelated exponential relaxed clock	1.11	0.66 - 1.58	0.51	0.31 - 0.73	0.63	0.34 - 0.94	0.42	0.23 - 0.62	0.99	0.57 - 1.45	0.59	0.35 - 0.81
Uncorrelated log normal relaxed clock	0.33	0.19 - 0.47	0.15	0.09 - 0.22	0.19	0.10 - 0.28	0.13	0.07 - 0.19	0.30	0.16 - 0.44	0.17	0.10 - 0.25
Molecular clock model	Exponential population growth											
	Third position		Coding region		tRNA		rRNA		D-loop		mtDNA	
	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval	Mean	HPD interval
HKY												
Strict clock model	0.34	0.21 - 0.49	0.15	0.09 - 0.22	0.19	0.10 - 0.28	0.13	0.07 - 0.18	0.31	0.17 - 0.45	0.18	0.11 - 0.25
Uncorrelated exponential relaxed clock	1.14	0.69 - 1.62	0.52	0.31 - 0.72	0.62	0.34 - 0.92	0.43	0.24 - 0.63	0.99	0.57 - 1.44	0.59	0.36 - 0.82
Uncorrelated log normal relaxed clock	0.38	0.24 - 0.53	0.17	0.11 - 0.24	0.21	0.12 - 0.31	0.14	0.08 - 0.21	0.34	0.21 - 0.49	0.20	0.13 - 0.28
GTR												
Strict clock model	0.33	0.20 - 0.47	0.15	0.09 - 0.21	0.19	0.11 - 0.28	0.12	0.07 - 0.18	0.31	0.18 - 0.44	0.18	0.11 - 0.24
Uncorrelated exponential relaxed clock	1.09	0.59 - 1.62	0.50	0.27 - 0.74	0.62	0.31 - 0.96	0.42	0.21 - 0.65	0.97	0.49 - 1.48	0.57	0.30 - 0.84
Uncorrelated log normal relaxed clock	0.36	0.22 - 0.50	0.16	0.10 - 0.23	0.20	0.11 - 0.30	0.14	0.08 - 0.20	0.32	0.19 - 0.46	0.19	0.12 - 0.26

6.3.4 Rates of Molecular Evolution for Synonymous and Constraint Sites of Tuatara mtDNA, Estimated by sUPGMA method

Considering the high rates obtained in section 6.3.3 and the fact that BEAST tends to have an upward bias on its rate estimates (Drummond per comm, Emerson 2007) when there is not enough data in data set, the time frame (5,000 years) is reasonably short for a long-lived vertebrate, I employed the other software and statistical methods in order to estimate the evolutionary rates. Using software Pebble, the evolutionary rates were estimated 0.39 s/s/Myr (HPD: 95% CI: 0.0 - 2.95) for synonymous, 0.20 s/s/Myr (HPD: 95% CI: 0.0 - 1.52) for non-synonymous, 0.26 s/s/Myr (HPD: 95% CI: 0.0 - 1.56) for tRNA, 0.19 s/s/Myr (HPD: 95% CI: 0.0 - 1.49) for rRNA, 1.44 s/s/Myr for D-loop (HPD: 95% CI: 0.0 - 3.55) and 0.40 s/s/Myr (HPD: 95% CI: 0.0 - 2.17) for the entire tuatara genome using the sUPGMA method and HKY model of nucleotide substitution (Table 6.8). Furthermore, the sUPGMA tree for different regions of tuatara mtDNA was constructed with software Pebble by accounting for the ages of the ancient samples (Figure 6.6-10). All these trees clearly show a similar pattern to the trees constructed by Pebble and MEGA in 6.3.2 for the entire mtDNA (Figure 6.2-5).

Table 6.8: The evolutionary rates for different regions of mtDNA are estimated by sUPGMA distance method, using Pebble software.

Substitution Model HKY+Gamma		
	Evolutionary rate	HPD 95% Confidence Interval
Synonymous positions	0.39	0.0 - 2.95
Non-synonymous positions	0.20	0.0 - 1.52
tRNAs	0.26	0.0 - 1.55
rRNAs	0.19	0.0 - 1.49
D-loop	1.44	0.0 - 3.55
HVRs	2.61	0.0 - 7.24
mtDNA	0.40	0.0 - 2.17
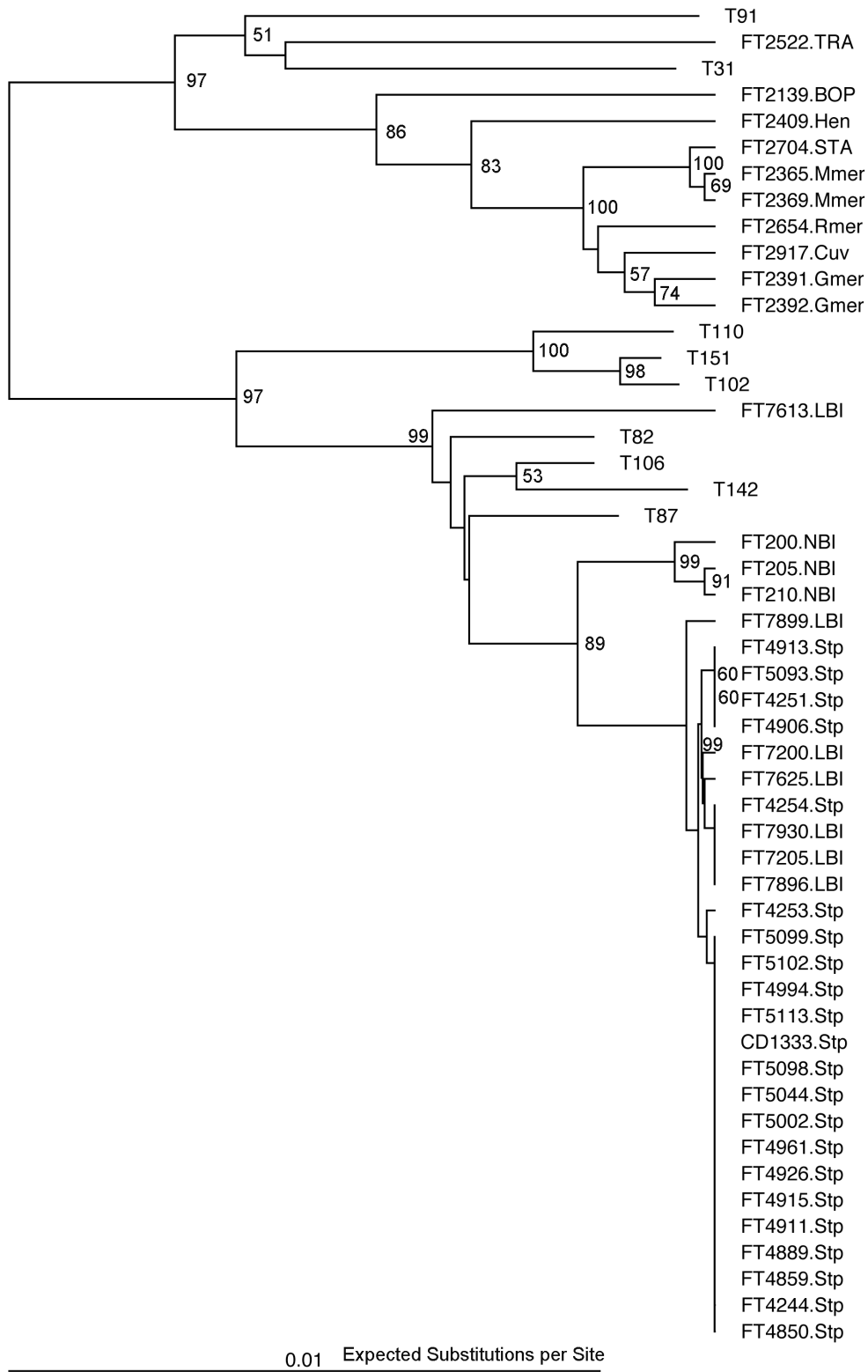
Figure 6.6: The sUPGMA tree for synonymous regions of modern and ancient tuatara mtDNA constructed by Pebble under HKY model of nucleotide substitution and Gamma site heterogeneity model.
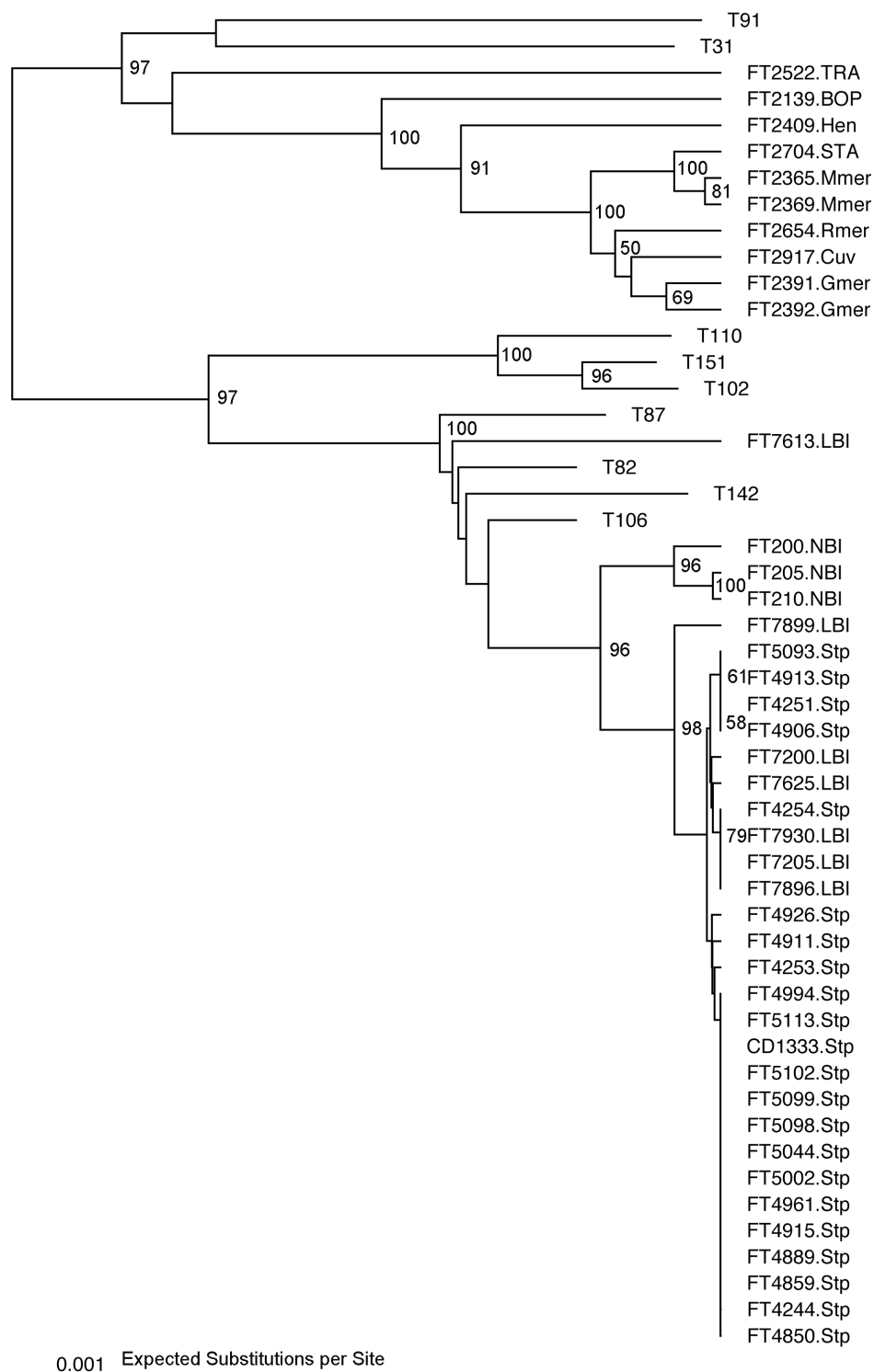
Figure 6.7: The sUPGMA tree for non-synonymous regions of modern and ancient tuatara mtDNA constructed by Pebble under HKY model of nucleotide substitution and Gamma site heterogeneity model.
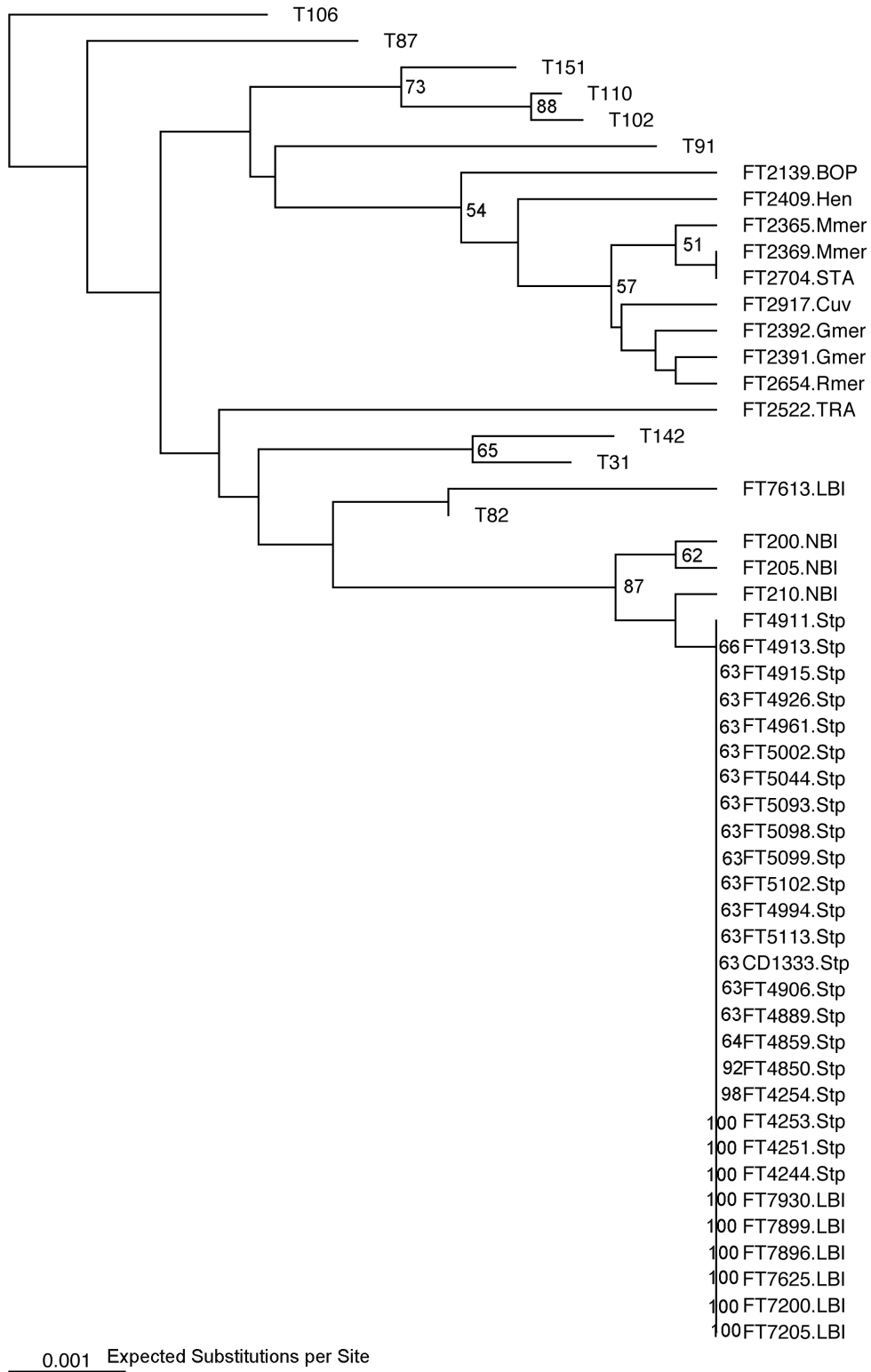
Figure 6.8: The sUPGMA tree for tRNA regions of modern and ancient tuatara mtDNA constructed by Pebble under HKY model of nucleotide substitution and Gamma site heterogeneity model.
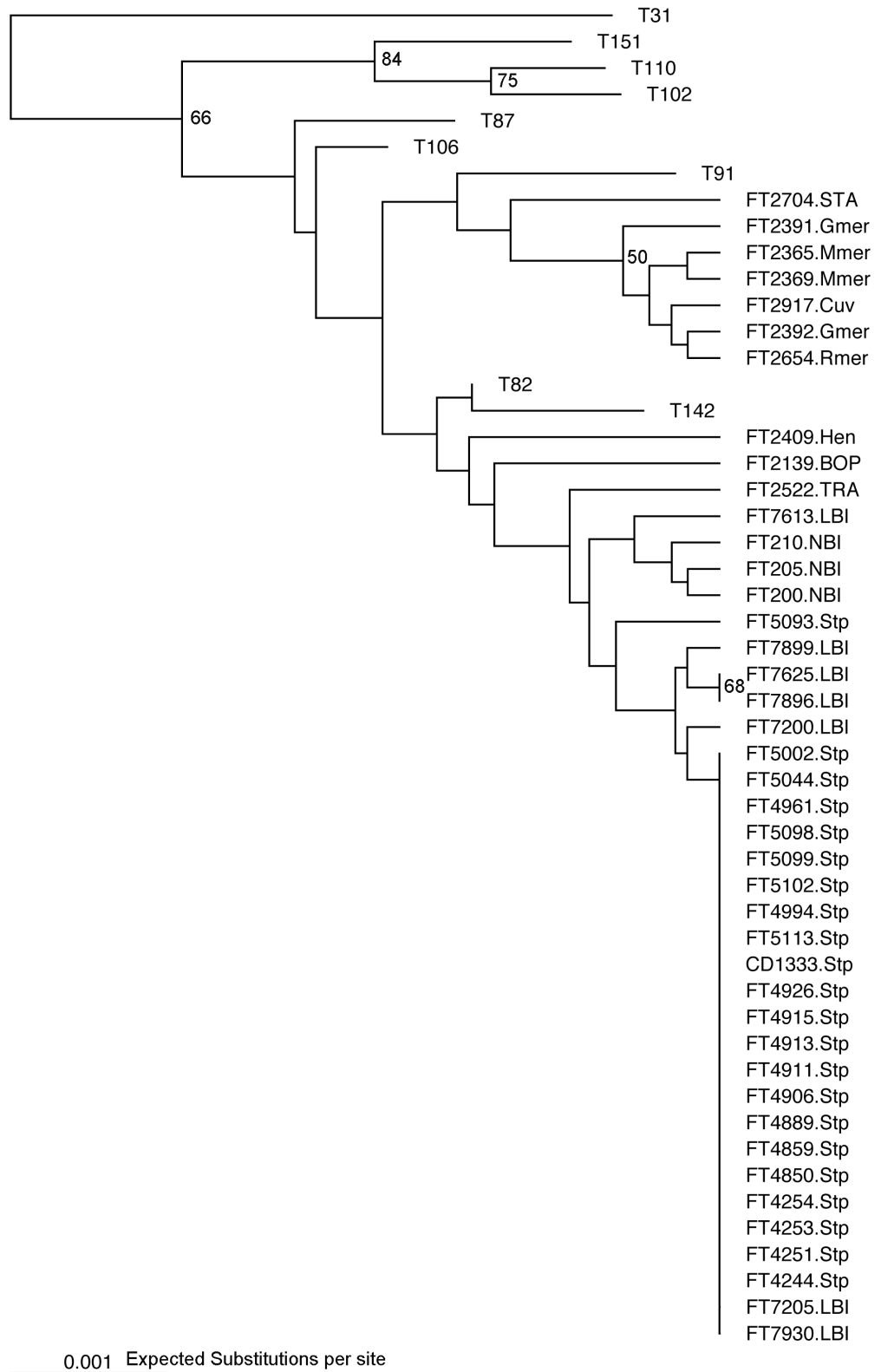
Figure 6.9: The sUPGMA tree for rRNA regions of modern and ancient tuatara mtDNA constructed by Pebble under HKY model of nucleotide substitution and Gamma site heterogeneity model.
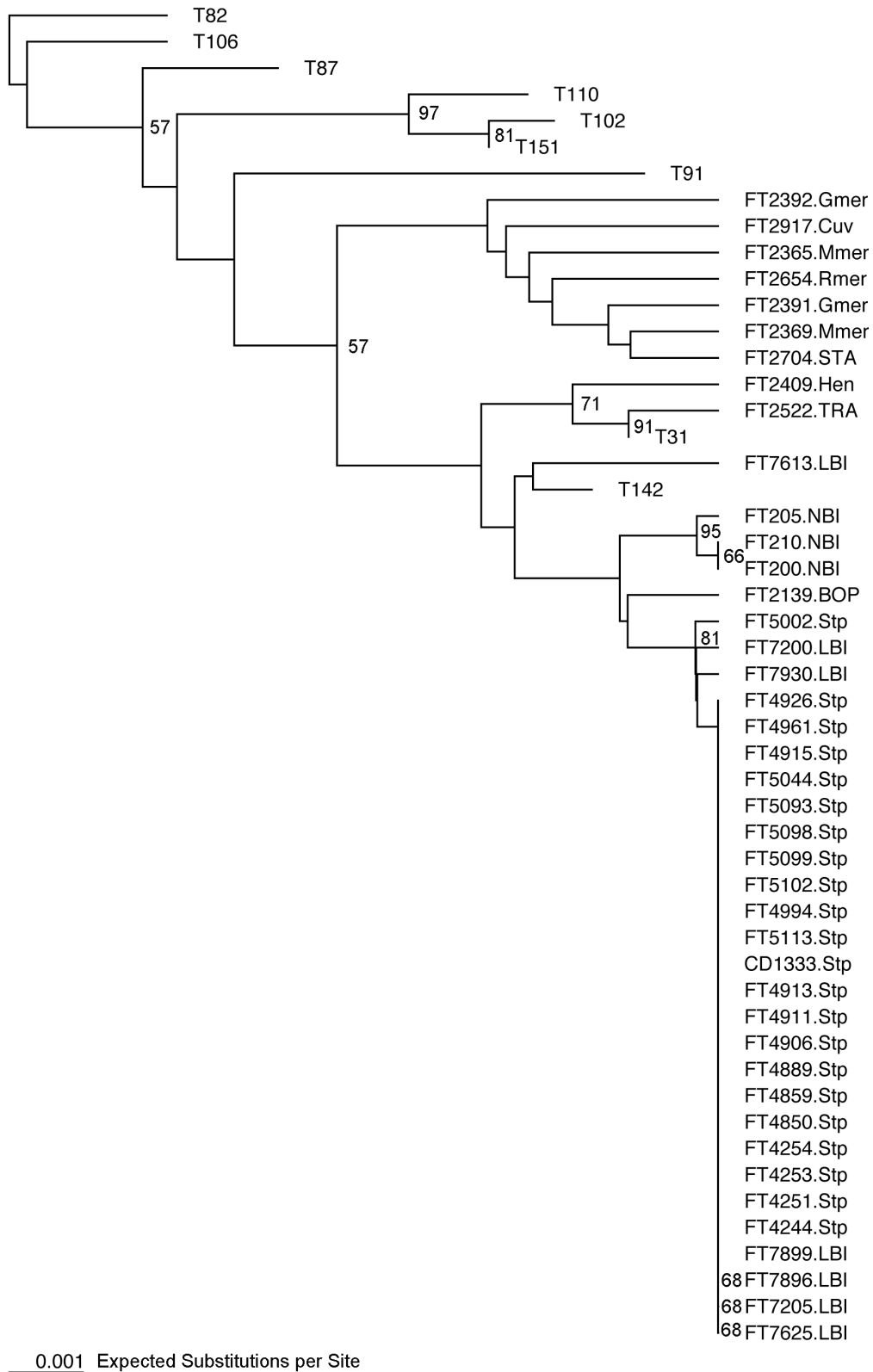
Figure 6.10: The sUPGMA tree for D-loop regions of modern and ancient tuatara mtDNA constructed by Pebble under HKY model of nucleotide substitution and Gamma site heterogeneity model.

6.3.5 Rates of Neutral Molecular Evolution in Tuatara mtDNA, Estimated from Modern and Ancient Samples, using Baysian Statistical MCMC and sUPGMA Method

The synonymous sites are suitable candidate for estimating the neutral molecular rates in mtDNA as they are supposedly free from selection and therefore are known to evolve neutrally (Yang & Nielsen 2008). The MCMC analysis shows that the mtDNA rate of evolution for synonymous sites (μs) is 0.33 s/s/Myr (HPD: 95% CI: 0.19 - 0.47) under GTR model, uncorrelated lognormal relaxed clock and constant population size assumption (Table 6.6). To avoid any bias towards the method and the software, I also analysed the data using sUPGMA method implemented in software Pebble and estimated the synonymous rates of molecular evolution 0.39 s/s/Myr (CI 0.0 - 2.95). This rate is slightly higher than the rate estimated using the Bayesian approach but not significantly different as it is within the range of HPD.

6.3.6 Inferring Tempo of Evolution at Constrained Sites (rRNA, tRNA, Non-synonymous Sites) in Modern Tuatara mtDNA

As detailed in section 6.2.6, I estimated the evolutionary rates in non-synonymous sites for tuatara mtDNA about 0.058 s/s/Myr (CI 0.033 - 0.083). The non-synonymous rate of evolution in tuatara mtDNA is about 5.8 times slower than the synonymous rates (0.33 s/s/Myr, CI 0.19 - 0.47). The evolutionary rates for tRNA and rRNA region of mtDNA were estimated 0.165 (CI 0.095 - 0.235) and 0.062 s/s/Myr (CI 0.035 - 0.088) respectively. These results indicate that the tempo of the evolution for tRNA is 2.6 times faster than rRNA. Furthermore, the evolutionary rate for D-loop region was estimated about 0.077 s/s/Myr (CI 0.044 - 0.109). The rate of evolution for complete tuatara genomes is 0.128 s/s/Myr (CI 0.073 - 0.182), which indicates the fast evolving pace of this genome in tuatara.

6.3.7 Rates of Molecular Evolution in Tuatara mtDNA, Estimated with joint Analysis by PAML and MEGA

As detailed in section 6.2.7, using joint analysis with PAML and MEGA, I estimated the overall rates of molecular evolution for tuatara mtDNA about 0.15 s/s/Myr

(Table 6.9). This rate is similar to the one obtained with BEAST analysis (0.17 s/s/Myr) in 6.3.3.

Table 6.9: The evolutionary rates for mtDNA sequences of modern and ancient tuatara samples were estimated by MEGA and PAML joint analysis. X is the overall mean between the ancestral sequence and modern samples. Y is the genetic distance pairwise between ancestral sequence and ancient samples. Z is the age of ancient samples (years before present). The evolutionary rates estimated according to this equation: (X-Y)/Z.

Ancient Sample	X	Y	(X-Y)	(Z)	Evolutionary Rates (s/s/Myr)
T31	0.00465	0.00349	0.00116	1684 BP	0.688
T82	0.00465	0.00457	0.00008	5191	0.015
T87	0.00465	0.00457	0.00008	4145	0.019
T91	0.00465	0.00457	0.00008	696	0.114
T102	0.00465	0.00457	0.00008	1546	0.051
T106	0.00465	0.00484	-0.00019	5191	-0.036
T110	0.00465	0.00430	0.00035	1794	0.195
T142	0.00465	0.00403	0.00062	1189	0.521
T151	0.00465	0.00511	-0.00046	2321	-0.198
					Average: 0.152

6.3.8 A Comparison of Nucleotide Composition and type of Mutation among Modern and Ancient Tuatara Populations

As detailed in section 6.2.8, a comparison of modern and ancient nucleotide compositions, nucleotide pair frequencies, numbers and the types of mutations were performed by MEGA. The results of this comparison indicate that there are no differences in nucleotide frequency, nucleotide pair frequency, numbers and types of transitions, transversions and types of mutations (Table 6.10). In addition, I attempted to determine if DNA sequences from the ancient samples used in this study had undergone post mortem damage as a result in substitution from C to T. The summarized results in Table 6.11 enable us to recognise and classify the orientation of all the nucleotide changes occurred in ancient and modern mtDNA of tuatara populations. The data clearly shows that there is no significant variation due to ancient DNA damage. If there is significant damage in ancient genome then the proportion of C/G to T/A change is higher compare to the modern sequences but the data show that these rates are similar in modern and ancient genomes.

Table 6.10: Ancient and modern tuatara complete mtDNA sequence characteristics: nucleotide composition, and pairwise number of identical pairs (ii), transitions (si), transversions (sv) and all possible nucleotide pairs, using MEGA.

	Ancient	Modern	All
Sample size	9	42	51
***Nucleotide Frequencies (%)**			
T	26.6	26.9	26.8
C	26.1	25.8	25.9
A	33.2	33.3	33.3
G	14.0	14.0	14.0
Total bp (avg.)	12961.3	15101.3	14723.7
****Nucleotide Pair Frequencies**			
Identical pairs (ii)	11,214	14,938	14,194
Transitional pairs (si)	59	49	56
Transversional pairs (sv)	14	11	13
R= si/sv	4.2	4.4	4.3
TT	2,975	4,013	3,804
TC	15	12	13
TA	3	2	3
TG	1	0	1
CT	18	14	16
CC	2,938	3,860	3,677
CA	3	2	3
CG	1	1	1
AT	3	3	3
AC	2	2	2
AA	3,732	4,975	4,727
AG	13	12	14
GT	0	0	0
GC	1	1	1
GA	14	12	13
GG	1,570	2,090	1,986
Total (avg.)	11287.8	14998.6	14263.0
Identical proportion	0.99	0.99	0.99
Transitions proportion	0.80	0.81	0.81
Transversion proportion	0.19	0.18	0.18

*All frequencies are given in percent
**All frequencies are averages (rounded) over all taxa.

Table 6.11: Summary of the types and number of nucleotide changes along branches using maximum likelihood method implemented in PAML (Yang 1997). The direction of changes is from the reconstructed ancestral sequences to ancient and modern mtDNA sequences.

Nucleotide change			
From	**To**	**Ancient**	**Modern**
A	C	0	10
A	G	25	60
A	T	4	20
C	A	11	8
C	G	0	4
C	T	21	65
G	A	19	62
G	C	1	2
G	T	0	4
T	A	4	10
T	C	12	48
T	G	0	3
Total		97	296
CT proportion		0.22	0.22
GA proportion		0.20	0.21
(CT+GA) proportion		0.42	0.43
(CT+GA)/(AG+TC)		1.08	1.18

6.4 Discussion

The estimation of rates of molecular evolution is one of the most important and controversial issues in evolutionary biology. There are various methods for estimating the evolutionary rates such as phylogenetic, ancient DNA and pedigree approaches. Since tuatara are not closely related to any modern taxa, estimating the evolutionary rates using the phylogenetic/calibration method would not provide a reliable rates estimate. Therefore, in order to estimate the molecular evolutionary changes over period of time, the ancient and pedigree methods are applicable. This study is the first attempt to estimate molecular evolutionary rates for the entire mtDNA from a comprehensive data set consist of 51 modern and ancient tuatara samples. Using ancient DNA approach, the evolutionary rates for complete tuatara mtDNA were estimated about 0.17 s/s/Myr (HPD: 95% CI: 0.10 - 0.25) by Baysian statistic MCMC implemented in BEAST. The rates of neutral molecular evolution for D-loop region were estimated 0.30 s/s/Myr (HPD: 95% CI: 0.16 - 0.44), which is 1.7 times faster than the entire genome. The rate for D-loop region is perfectly comparable to some published rate estimates of ancient mtDNA, using the same joint analysis of ancient and modern DNA by MCMC approach in BEAST (Table 6.12).

Table 6.12: A number of published rate estimates from studies of ancient D-loop regions of mtDNA, using Bayesian method.

Species	Region	Rate (Substitution/Site/Myr)	Oldest tip (kyr)	Reference
Tuatara (*Sphenodon*)	D-loop	0.30 (0.16-0.44)	5	This study
Tuatara (*Sphenodon*)	D-loop	1.56 (0.83-2.34)	9	Hay et al. (2008)
Adelie Penguin (*pygoscelis adeliae*)	D-loop	0.96 (0.53-1.43)	6	Lambert et al. (2002)
Adelie Penguin (*pygoscelis adeliae*)	D-loop	0.86 (0.53-1.17)	6	Millar et al. (2008 b)
Aurochs *(Bos primigenius)*	D-loop	0.69 (0.15-1.30)	12	Edwards et al. (2007)
Mappin's moa (*Pachyornis mappini*)	D-loop	0.67 (0.01-2.09)	6	Ho et al. (2007 a)
Bison (*Bison bison*)	D-loop	0.32 (0.23-0.41)	60	Shapiro et al. (2004)
Brown bear (*Ursus arctos*)	D-loop	0.30 (0.13-0.48)	59	Saarma et al. (2007)
Cave bear (*Ursus spelaeus*)	D-loop	0.26 (0.10-0.53)	80	Saarma et al. (2007)
Cave lion (*Panthera leo spelaea*)	D-loop	0.20 (0.03-0.40)	62	Ho et al. (2007 a)
Ox (*Bos Taurus*)	D-loop	0.13 (0.00-0.41)	8	Ho et al. (2007 a)
Horse (*Equus caballus*)	D-loop	0.11 (0.02-0.31)	28	Ho et al. (2007 a)

In general, it has been argued that estimates of rates based on ancient DNA may generally be upwardly biased (Emerson 2007). This could be, for instance, as a result of reduction in population size over the time period studied. In such a case, this would result in higher rates of fixation of even deleterious mutations and therefore result in an increase in evolutionary rates estimates (Emerson 2007). Although, that is possible, since the high molecular rates have been observed in ancient DNA studies using a broad range of organisms such a primates, artiodactyls, rodents, birds, reptiles and mammals (Ho et al. 2005 b; Lambert et al. 2002; Hay et al. 2008; Edwards et al. 2007; Ho et al. 2007 a; Shapiro et al. 2004; Saarma et al. 2007), it is unlikely that all this species went through an universal population decline. Although tuatara population substantially declined since human arrival in New Zealand around 730 years ago (Duncan et al. 2002), only one out of nine samples in our study postdate this decline. Therefore, the recent demographic history of tuatara cannot explain the rates estimated for tuatara.
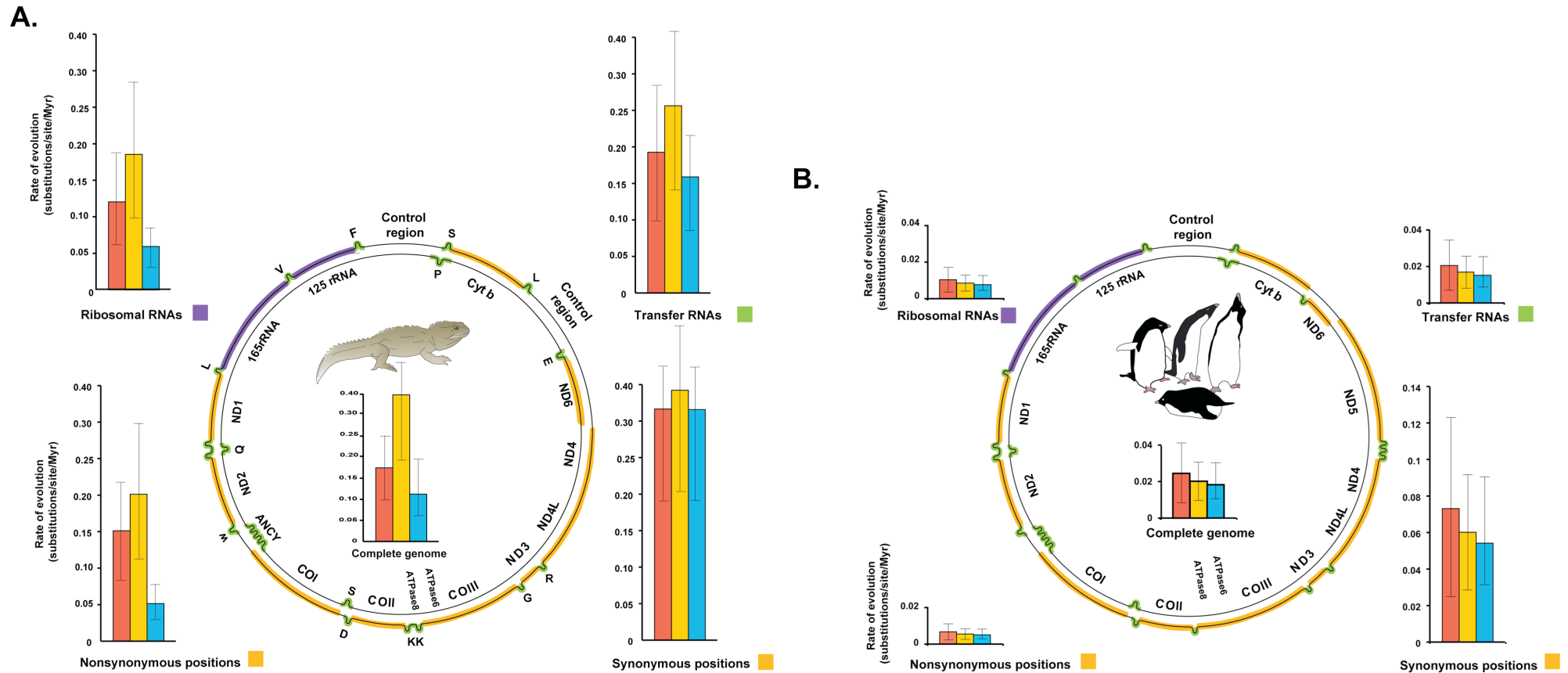
Figure 6.11: A comparison of mitogenomic rates of evolution between tuatara **(A)** and Adélie penguins **(B). A.** Using mitochondrial genomes of 42 modern and nine ancient tuatara, the evolutionary rates for synonymous and constrained sites (rRNA, tRNA, non-synonymouse) were estimated by Baysian statistics MCMC (red column) and maximum likelihood approach (yellow column). In addition, the evolutionary rates in different mitochondrial regions (Kx) are estimated as Kx= μs (dx/ ds) (blue column), where dx is the divergence in constrained site (e.g. tRNA) calculated by MEGA, and the ratio dx/ds is the fraction of neural positions. The error bars show 95% HPD intervals. **B.** In this study, the evolutionary rates estimated for various regions of tuatara mitochondrial genome is about 5 - 10 times higher than those reported for Adélie penguins by Subramanian et al. (2009 a). The figure is drawn by the Author of the thesis.

The evolutionary rate estimated for D-loop region of tuatara in this study (0.30 s/s/Myr, HPD: 95% CI: 0.16 - 0.44) is not significantly different from rates estimated for the same region in horse, ox, cave lion, cave bear, brown bear and bison (Table 6.12). However, the evolutionary rates estimated by Hay et al. (2008) for the same region of tuatara mtDNA is higher than the rates estimated in this study. On the other hand, the high rate reported (1.56 s/s/Myr) by Hay et al. (2008) is not significantly different than the rates reported for penguins, aurochs and moa (Table 6.12).

Initially, the evolutionary rates were separately estimated for different regions of tuatara mtDNA under constant population size and GTR model of nucleotide substitution. According to this analysis, the evolutionary rate for D-loop region of tuatara mtDNA was estimated about 1.85 s/s/Myr (HPD: 95% CI: 0.77 - 3.06). This rate is similar to the rate obtained by Hay et al. (2008). However, since D-loop and the coding sequences are all on the same linked genome (the mitochondrion), it could be argued that they should be analysed together. From the estimation of root height in these analyses, it was clear that they do not provide consistent signals. The root heights should be comparable because these data sets should have evolved on the same time tree. In order to achieve this point, the data were partitioned into different genes (tRNA, rRNA, D-loop, coding region) and the substitution model and clock model for these regions were unlinked, while the time tree were linked. This method estimates the rates for different partitions while forcing them to evolve down the same tree. Using this approach, the evolutionary rates for D-loop came down to 0.30 s/s/Myr as reported in Table 6.7. In general, the rates estimated for tRNA, rRNA and coding regions seem to be relatively high in comparison to those for the D-loop region of mtDNA. Moreover, the evolutionary rates estimated here for various regions of tuatara genome are almost 5 - 10 times greater than those estimated for Adélie penguins (*Pygoscelis adeliae*) (Subramanian et al. 2009 a) (Figure 6.11) using similar analytical methods. Conservatively, these results indicate that the overall evolutionary rate for tuatara mtDNA is likely to be high, however the methods employed in this study are based on simple models of evolution. It has been shown that when sequences evolve in a constant rate, various models of rate change (strict and relaxed clock) are able to infer rates more accurately compare to when sequences evolve under autocorrelated/uncorrelated rate change. However, if sequences evolved under a model of autocorrelated rate change, rates were accurately estimated, using relaxed-clock models

(lognormal and exponential) while when the sequences evolved under a model of uncorrelated rate change, Baysian inference using an exponential rate model is suggested. Therefore, Baysian inferences using an exponential model of rate change are suggested as a conservative approach to estimate the molecular evolutionary rates and divergence time (Ho et al. 2005 a). In addition, it has been argued that when nucleotide differences between samples arise from genetic drift, DNA damage, migration, selection, bottleneck or population subdivision rather than mutation, they can bias the point rate estimates and divergence date (Navascues et al. 2010; Navascues & Emerson 2009). The statistical approaches and software for the joint analysis of ancient and modern DNA may be further developed in future however, this probably affects the rate estimates of all the vertebrates and the high evolutionary rates reported for tuatara with such a stable morphology may not be dramatically changed.

Chapter Seven

Mitochondrial DNA Variant Discovery in Tuatara Using Next-Generation DNA Sequencing

7.1 Introduction

The average vertebrate cell contains literally thousands copies of the mitochondrial genome. These copies are found within mitochondria with typically hundreds of mitochondria within each cell (Figure 2.3). And then, nested within each of these mitochondria are multiple copies of mitochondrial genomes. This nested pattern, like Russian dolls (Figure 1.1), together with a typically high mutation rate in this genome result in the presence of different mtDNA variants within individuals (heteroplasmy). Heteroplasmy can persist over generations and is potentially difficult to detect, depending on the frequency of the rare variants and the sensitivity of DNA sequencing technologies (Millar et al. 2008 b).

Traditionally, the Sanger sequencing method has been used to detect nucleotide variants in mitochondrial genome (Sanger et al. 1977; Schrijver et al. 2009). However, this method is neither particularly sensitive nor specific to the detection of heteroplasmic

variants in mitochondrial genome (Hartmann et al. 2009). In addition, limitations including the time required to manually generate and inspect electropherograms for heteroplasmic variants typically prevents the routine use of complete mtDNA by Sanger sequencing method (Millar et al. 2008 a). Recently, advances in digital imaging systems and bioinformatics have lead to introduction of next generation sequencing methods. Currently, there are four next-generation sequencing methods that have been widely used (Millar et al. 2008 a): The Roche 454 Genome Sequencer FLX system (Indianapolis, IN, USA), The Solexa (Illumina) Genome Analyser (San Diego, CA, USA), the Applied Biosystems SOLiD system (Foster City, CA, USA) and the HeliScope True Single Molecule Sequencing System (Cambridge, MA, USA). These methods in part at least overcome some of these limitations. The next-generation sequencing machines have the capacity of massive sequencing and are capable of generating sequences in less time and overall costs compared to traditional methods (Meyer et al. 2007, 2008; Voelkerding et al. 2009). In comparison, the most advanced conventional sequencers produce at best 70 kb sequences per run, whereas the 454 sequencing machine is able to generate 100 Gb per run (Millar et al. 2008 a). Therefore, a number of studies (He et al. 2010; Tang et al. 2010; Zaragoza et al. 2010) have focused on the detection of heteroplasmies, using this technology.

The higher the frequency of the heteroplasmic variant and the larger the number of generations that it persists, the higher the probability it will be detected by any sequencing method. A heteroplasmy could persist in a maternal line of descent for many generations until it is eventually lost or fixed in a population of mitochondria within an individual. The larger the number of mitochondrial genomes that pass through the inheritance bottleneck, the longer the heteroplasmy is likely to persist over generations. However, this is dependent upon the distribution of the mitochondrial variants within and among mitochondria and upon their inheritance patterns. Since several mutant variants of a heteroplasmic genotype are likely to coexist in a cell and perhaps even in an individual or a population of individuals, it is difficult to discriminate between rare variants and background signals in DNA trace data generated using Sanger DNA sequencing methods. In Sanger sequencing, if the threshold detection level is set very low, the noise signals will erroneously be scored as heteroplasmic variants and this would consequently overestimate the mutation rates. Millar et al. (2008 b) suggested the use of 23% detection threshold frequency for Sanger

Sequencing method. They indicated that when the threshold detection level is set lower than 23% the number of detected heteroplasmic sites slightly increase but false positives are also included. According to this model most heteroplasmies do not reach this level and most of them that reach this level do not go to fixation within the individual, let alone the population or species.

Using a pedigree approach and a recently developed mathematical model (Hendy et al. 2009), Millar et al. (2008 b) estimated the observed rate of heteroplasmies (μ_0) of the HVR-I region of Adélie penguin mtDNA at about 54.9 s/s/Mys (95% HPD 41.2 - 68.6) (Figure 1.3). Moreover, they estimated the mutation rate (μ) at about 0.55 s/s/Mys (95% HPD 0.29 - 0.88) after accounting the inter-generational persistence of heteroplasmies. Therefore, they concluded the short-term mutation (μ) (0.55 s/s/Mys) and long-term evolutionary rates (k) (0.86 s/s/Mys) are similar in Adélie penguins over the 37,000 years time period examined. In addition, they suggested the artificially high estimates of mutation rates from pedigree data, previously reported by many studies, have been subject to a range of errors, e.g. earlier less sensitive DNA sequencing technologies. Furthermore, as reviewed in chapter one, they argued that previous studies (Parsons et al. 1997; Sigurğardóttir et al. 2000; Howell et al. 2003; Santos et al. 2005) have treated heteroplasmies in varying ways. Some studies have simply ignored heteroplasmic sites, others have included them and still others have estimated their individual probabilities of fixation and then included them on that basis.

7.2 Purposes and Scope of the Project

The goal of this component of my study was: (i) to explore the sensitivity of 454 DNA sequencing technology to detect nucleotide variations in tuatara mtDNA; and (ii) to investigate if the heteroplasmic variants in tuatara mtDNA could be partially or completely responsible for the high molecular evolutionary rate estimated for this species. This study is the first attempt to use the most recent sequencing technology to generate the nucleotide sequences of tuatara mitochondrial genome and determine the level of heteroplasmy.

7.3 Analytical Methods

The samples were prepared as detailed in 4.2.5 (chapter four) and sequenced at University of Otago High-Throughput DNA Sequencing Unit using a GS FLX, using Titanium chemistry. After sequencing, a large number of 454 sequence reads (1 - 2 Mb per sample) were sorted according to sample-specific tags and the remaining sequences discarded. The output was saved as FASTA and sff (standard flow format) files for analysis by FLX assembly software using the tuatara genome sequence from GenBank (Accession NC_004815) as a reference.

7.4 Results

7.4.1 Analysis of Heteroplasmic mtDNA Variants

Here, I have explored the possibility that heteroplasmic variants were at least one of the factors influencing of the high evolutionary rates estimated in tuatara. For this purpose, a total number of eight mtDNA from tuatara adults from Stephen Island were sequenced using Roche 454 DNA next generation DNA sequencing technology. The numbers of enriched beads for constructed libraries ranged from 2,200-152,400 (Table 4.6). One sample (FT4253) was discarded from the analytical data pool due to the low number of enriched beads (2,200) obtained for its library. However, the sequence of this mitochondrial genome was obtained, using Sanger sequencing. The data from the remaining seven genomes were screened for possible heteroplasmic variants.

It has been documented that, in theory, as few as 1 heteroplasmic variant per 10,000 mitochondrial genome could be detected by massively parallel 454 sequencing approach (He et al. 2010). However, errors that have accumulated during PCR and/or sequencing limit the sensitivity and accuracy of the approach. 454 sequences data show that a large number of sites (avg. 2000 per genome) across the mitochondrial genome possess two different nucleotides at the same position. The majority of these nucleotide differences are likely to represent sequencing/PCR errors or alignment irregularities arising from homopolymer runs in the sequence (Clark et al. 2001). In such cases, individual sequence reads will over/underestimate the number of bases in a homopolymer run. This causes

slippage in the alignment, which shows as differences at nucleotide positions flanking the homopolymer and that potentially mask a real mutation at or adjacent to the region.

Using gsAssembler/newbler (454/Roche) software, the true variants were identified from all types of sequencing/PCR errors and the nucleotide positions with variation compared to reference sequence were separately detected in each tuatara genome. Among these positions, those with 100% nucleotide variation in comparison with reference sequence were defined as homoplasmic sites in that particular sample and the remaining ones (yellow colored) were considered as potential heteroplasmies (Appendix B. Tables B1-B7). In summary, the total numbers of 28 potential heteroplasmic positions were identified among all these seven tuatara genomes in comparison to Genbank reference sequence (Table 7.1). The majority of these heteroplasmic variants (21 out of 28) were detected in protein coding regions of the genome and a small number in rRNAs, tRNAs and the control region (Table 7.1, Figure 7.1).

When the alignments of the 42 modern tuarara mitochondrial genomes were examined a number of variations (polymorphisms) were found in exactly the same positions where heteroplasmies were detected (last two columns in Table 7.1). Since these 42 genomes were sequenced using the Sanger method it is very likely that one of the variant was preferentially amplified by PCR. This would falsely result in defining a heteroplasmy as a true polymorphic variant. Such a potential error would elevate any estimate of molecular evolutionary rates. Therefore to remove this bias I excluded the heteroplasmic positions and re-estimated the rates, using Baysian MCMC method.

Eliminating the heteroplasmic positions, the evolutionary rates were re-estimated as: 0.19 s/s/Mys (95% HPD 0.08 - 0.32) for D-loop, 0.11 s/s/Mys (95% HPD 0.04 - 0.19) for tRNA, 0.09 s/s/Mys (95% HPD 0.04 - 0.15) for protein coding (non-synonymous) and 0.07 s/s/Mys (95% HPD 0.02 - 0.11) for rRNA regions of tuatara mtDNA, under constant population size, uncorrelated log normal relaxed clock and GTR model of nucleotide substitution. The overall mtDNA evolutionary rate was estimated about 0.10 s/s/Mys (95% HPD 0.04 - 0.17) with the same models (Table 7.2).

According to these data and using Baysian MCMC method, removing the heteroplasmic sites potentially reduced the rate of molecular evolution by approximately 40% [(0.17 - 0.10)/0.17= 0.41] compared to those originally estimated in chapter six (Table 7.2, Figure 7.1). In general, the results above suggest that a generally high mutation

rate observed in this study could have resulted due to the misidentification of heteroplasmies as true polymorphisms.

Table 7.2: A comparison of the mutation rates for different regions of mtDNA before and after removing the heteroplasmic variants, using Baysian MCMC method.

Evolutionary Rates, Gamma, GTR model		
	Original Rate (95% CI)	New Rate (95% CI)
Non-synonymous positions	0.15 (0.09 - 0.22)	0.09 (0.04 - 0.15)
tRNAs	0.19 (0.10 - 0.28)	0.11 (0.04 - 0.19)
rRNAs	0.13 (0.07 - 0.19)	0.07 (0.02 - 0.11)
D-loop	0.30 (0.16 - 0.44)	0.19 (0.08 - 0.32)
mtDNA	0.17 (0.10 - 0.25)	0.10 (0.04 - 0.17)

In addition, I estimated the root heights of the most recent common ancestor (µtMRCA, substitutions per site) of tuatara using complete mitochondrial genome data. The total height of the tree was estimated 57,153 years (57 Ky) (95% HPD 27,112 – 94,076) including both the modern and ancient tuatara mitochondrial genomes, using Baysian statistic-based MCMC under constant population size and GTR model of nucleotide substitution. This coalescent time pre-dates that estimated in chapter six (35 Ky) due to reduced rates of molecular evolution resulting from the elimination of the heteroplasmic sites across the genome.

Table 7.1: Summary of potential heteroplasmic variants recorded among seven tuatara complete mitochondrial genomes, mapped to NCBI Reference (NC_004815). The first column shows the tuatara individuals that represent the heteroplasmic variations. The positions of heteroplasmic variants in the reference and variant nucleotides are shown. If the number of samples that show heteroplamies is > 1 then the total variation and total depth of sequencing coverage were the average estimates. The total depth is the average number of mapped sequences (non-redundant and redundant sequences) at each base position. The polymorphism column indicates the heteroplasmic sites that are polymorphic among the 42 aligned sequences. The polymorphic samples are the individuals among 42 aligned modern tuatara sequences that represent this variation.

Potential Heteroplasmies									
Sample ID	Position in Ref	Position on mtDNA Genes	Reference Bases	Variation Bases	Number of Samples	Total Variation Percent	Total Depth	Polymorphism	Polymorphic Sample
4241	1571	ND-2	G	T	1	25.8	31	G	-
4246	2063	ND-2	C	T	1	95.2	21	C/T	FT7613
4246, 4250	3916	COX-I	C	T	2	91	39	C/T	FT7613
4246,4250	4114	COX-I	T	C	2	78.6	50	T/C	Many
4249	4127	COX-I	C	T	1	15.2	46	C	-
4249	4128	COX-I	A	G	1	15.2	46	A	-
4246	4713	COX-II	C	G	1	91.3	23	C/G	Many
4246	5018	tRNA-Lys	G	A	1	95.7	46	G/A	FT7613
4241	5021	tRNA-Lys	T	C	1	23.4	47	T/C	Many
4246	6402	COX-III	C	T	1	97.4	39	C/T	Many
4250	7325	ND4-L	A	G	1	91.1	45	A/G	FT7613

4241,4246, 4250, 4251, 4252	7605	ND-4	A	G	5	98.1	79	G/A	FT4911
4250	7690	ND-4	A	T	1	96.9	98	A/T	FT7613
4249	8030	ND-4	A	G	1	11.7	60	A	-
4246	8105	ND-4	G	A	1	81.7	60	G/A	Many
4246	8269	ND-4	A	C	2	85.9	51	A/C	Many
4246	8596	ND-4	T	C	1	98.7	79	T/C	Many
4246	8689	ND-4	A	C	1	68.5	73	A/C	Many
4251	9197	ND-6	T	A	1	13.5	37	T	-
4250	9205	ND-6	G	A	1	98.4	64	G/A	Many
4246	9250	ND-6	C	A	1	93.5	46	C/A	FT7613
4250	9298	ND-6	A	G	1	80.6	31	A/G	Many
4246	11265	Cyt-B	G	A	1	88.9	27	G/A	FT7613
4246,4250	12587	D-loop	G	A	2	94.7	35	G/A	Many
4246	12709	12S-rRNA	G	A	1	98.9	89	G/A	Many
4246	12860	12S-rRNA	T	C	1	98.2	57	T/C	Many
4249,4252	13092	12S-rRNA	T	A	2	83.7	30	A/T	FT205
4244	14515	16S-rRNA	A	G	1	97.1	35	G/A	Many
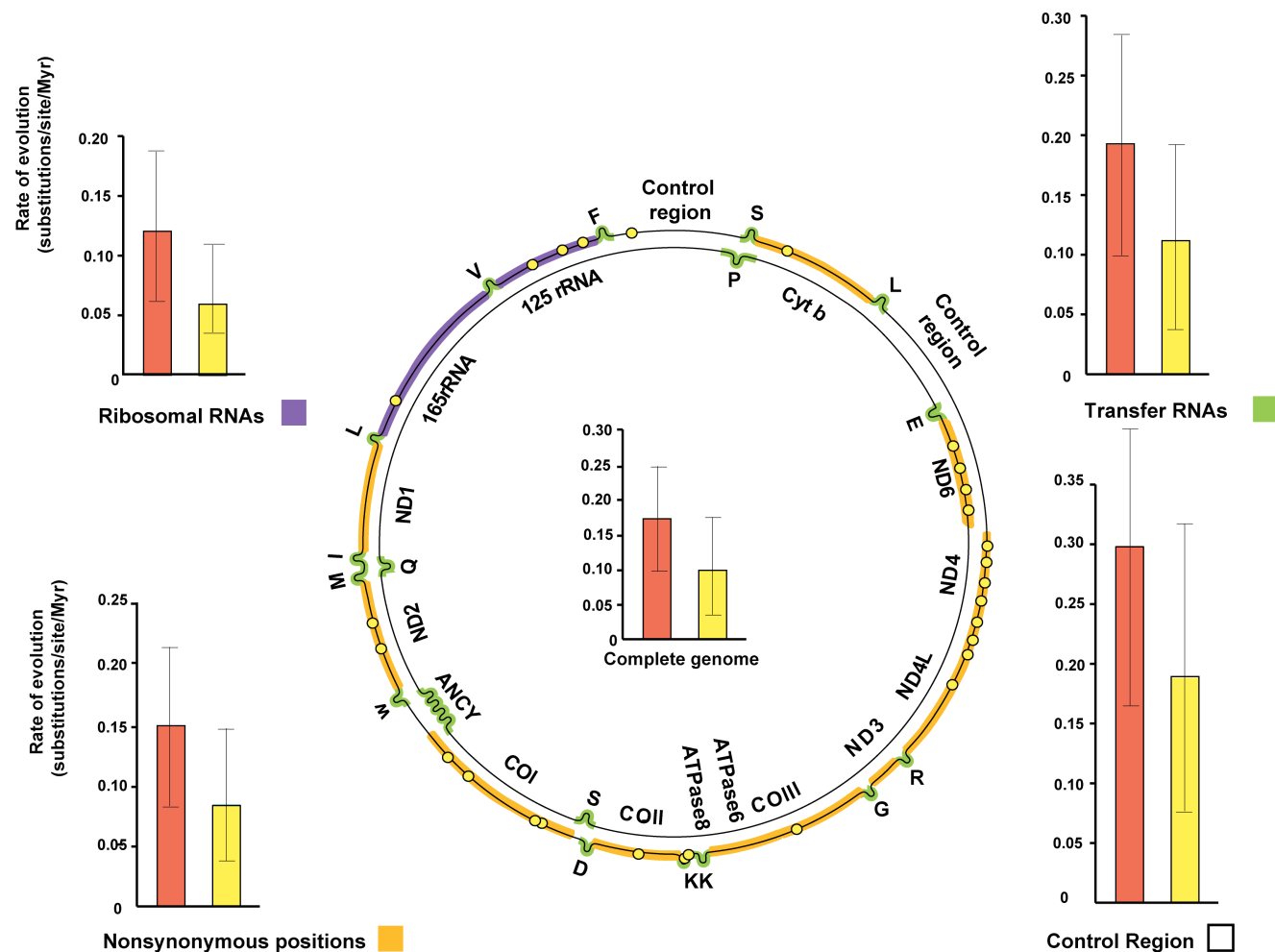
Figure 7.1: A comparison of molecular rates of evolution for different regions of tuatara mtDNA before (orange color) and after (yellow color) removing the heteroplasmic variants, using Baysian MCMC method. The positions of heteroplasmic variants are depicted on the genome by yellow circles. The figure is drawn by the Author of the thesis.

7.5 Discussion

The use of Roche 454 sequencing provides the opportunity to simultaneously obtain the complete nucleotide sequence of tuatara genomes and to detect potential heteroplasmic positions in the mtDNA. The results of this study indicate that the 454 DNA sequencing method is more sensitive to low levels of mtDNA heteroplasmy than Sanger sequencing. This is partially explained by high DNA sequence coverage achieved by next generation methods.

Using Roche 454 sequencing, the numbers of 28 potential heteroplasmic variants were detected among seven tuatara mitochondrial genomes. This suggests that there is a high level of heteroplasmy in tuatara mtDNA. However, the number of mutations that can be detected is largely dependent on DNA sequencing coverage at variant sites. A 20X coverage for detecting homoplasmy and at least 200X for heteroplasmies where the rarer variant is present at a frequency of greater than 10% has been suggested (Zaragoza et al. 2010; Meyer et al. 2008; Vasta et al. 2009). Moreover, it has been suggested that detecting the heteroplasmic variants with lower frequencies would require even deeper sequencing coverage, estimated 1,500 X for ≥ 5% (Tang et al. 2010) and 15,000 X to 100,000 X for as low as 2% (He et al. 2010; Thomas et al. 2006). The average 50X coverage obtained in this study would be expected to detect only a fraction of the real heteroplasmic sites. Hence the molecular rates reported above are almost certainly higher than the actual rates.

This study supports the use of 454 DNA sequencing technology in order to detect heteroplasmic variants that otherwise would be missed using the conventional Sanger sequencing. I also suggest that a generally high mutation rate, manifested by the detection of a high level of heteroplasmic nucleotide sites, is at least partially responsible for the tuatara's high rate of molecular evolution. Supporting this, removing the potential heteroplasmic sites from 42 aligned mitochondrial genomes reduced the mtDNA evolutionary rates by about 40%. However, the number of heteroplasmic variants reported here is based on screening only seven genomes (with about 50X sequencing coverage) and on an average it is four heteroplasmies per genome. Therefore extrapolating this for 51 tuatara mitochondrial genomes (used to estimate the rates of evolution) results in an estimated total of 204 heteroplasmies, assuming all are present in independent positions. This suggests that the actual rate of evolution would be much less than that reported here if

all the heteroplasmies were excluded. Therefore, in order to achieve a more accurate estimate of molecular evolutionary changes, the detection of all heteroplasmic variations is required.

In future, as the performance and output from next-generation DNA sequencing technologies continue to improve, it may be possible to sequence a larger number of genomes with higher coverage at a lower cost. This would provide more accurate data for mutation detection. These will be important developments that will help us to insignificantly improve estimates of molecular evolutionary rates estimated using different methods.

Chapter Eight

Discussion and Conclusion

8.1 Correlation between Molecular and Morphological Evolution

This study concerns a fundamental and important question in evolutionary biology, whether and/or to what degree organismic form and molecular evolution are coupled. The relationship between rates of anatomical/phenotypic evolution and genetic change has been the subject of debate for decades. Early evolutionary biologists such as Lamarck (1809) and Darwin (1859) were concerned about the evolution of biological form in various species and of course the mechanism driving these changes. Simpson (1944) focused on the tempo and mode of evolution in lineages that have undergone very little morphological changes (e.g., horseshoe crabs, the coelacanth Latimeria, and crocodiles) compared to mammalian lineages (e.g., humans and elephants) that have evolved rapidly at the level of morphology. With these examples in mind, Simpson (1944) argued that the rates of phenotypic evolution and genetic changes are correlated. However, at the time that Simpson wrote we had a very poor understanding of the nature and extent of genetic

changes in these species and others. Moreover, it is fair to comment that at that time, and perhaps even now, the traditional expectation has been that species that have remained relatively unchanged anatomically over long periods of evolutionary time should be expected to exhibit slow rates of molecular change. This view is consistent with the idea that changes in both anatomy, as well as molecular organisation, are driven by the environment which is external to the organism.
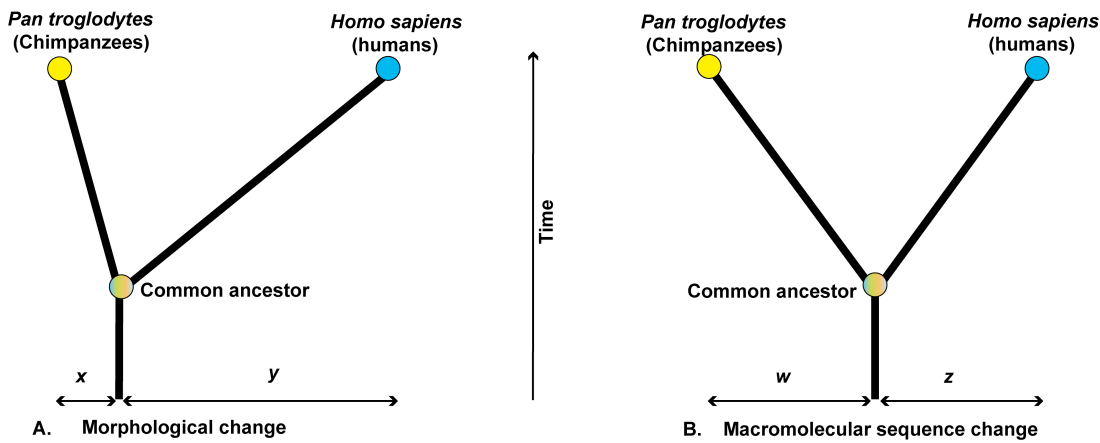
Figure 8.1: The contrast between morphological and molecular evolution in human and chimpanzee lineages since their divergence from a common ancestor. **A.** Since this evolutionary split from a common ancestor, more morphological change has resulted in humans (y) compared to the chimpanzee lineage (x). **B.** In contrast, the levels of protein and DNA evolution in human and chimpanzee lineages remain remarkably similar (Simpson 1970; King & Wilson 1975). The figure is re-drawn by the Author of the thesis.

Over the years some studies have suggested the opposite, namely that rates of molecular and morphological evolution might be decoupled. Mary-Clair King and Allan Wilson (King & Wilson 1975) suggested that the low level of molecular divergence at the protein level observed in human and chimpanzee lineages could not account for differences in anatomical and behavioural features in these lineages (Figure 8.1). Supporting this view, a range of other studies have reported no association between molecular and morphological evolution in various vertebrate taxa such as birds (Prager & Wilson 1975), mammals (Wilson et al. 1974 b), frogs (Wilson et al. 1974 a; Wallace et al. 1971; Wallace et al. 1973; Cherry et al. 1978) and tuatara (Subramanian et al. 2009 b). Therefore, it has been proposed that the evolutionary changes in anatomy and way of life (behavior and

physiology) are more often based on changes in gene regulatory systems than on changes on protein sequences (Ohno 1972; King & Wilson 1975; Wilson et al. 1974 ab; Prager & Wilson 1975; Miquerol et al. 2000; Bromham et al. 2002; Enard et al. 2002; Cáceres et al. 2003; Clark et al. 2003; Khaitovich et al. 2004; Wienholds et al. 2005; Morton et al. 2008; Huang et al. 2009; Nowick et al. 2009). A recent study reporting the draft genome of Neanderthals and comparing this genome to that of humans and chimpanzees has illustrated that Neanderthals are not genetically very distinct from modern humans (Green et al. 2010). For example, Green et al. (2010) recorded only 78 amino acid changes between Neanderthals and humans (Green et al. 2010). This suggests that a small number of changes in proteins that have taken place in the past few hundred thousand years of human evolution could not be entirely responsible about morphological differences between humans, Neanderthals and chimpanzees.

In contrast and supporting Simpson's earlier view, a number of studies have reported a level of congruence between morphological and molecular patterns of differentiation (Bousquet et al. 1992 ab; Smith et al. 1992; Savard 1993; Omland 1994, 1997; Polly 2001, 2003). Bousquet et al. (1992 a) reconstructed the phylogenetic relationships among an ancient family of woody dicots (the birch family) based on morphological characters and molecular data (rbcL gene in chloroplast). These phylogenies strongly concluded that there is an association between the rates of molecular and morphological evolution in that group. Moreover, Polly et al. (2001, 2003) came to the same conclusion based on comparison of morphological (molar shape) and molecular data (cytochrome b) in North American Marmota (Rodentia: Sciuridae). In agreement, Omland (1997), presented evidence on a wide range of taxa from carnivores to dwarf dandelions that rates of molecular and morphological evolution are usually coupled. However, Bromham et al. (2002) re-examined the relationship between morphological and molecular rates, using 13 published phylogenetic datasets for vertebrate taxa, including both DNA sequence and morphological data. They found no evidence of an association between molecular and morphological rates of change and concluded that Omland's (1997) result is probably due to methodological bias rather than being indicative of an underlying evolutionary phenomenon. They even argued further, that neither molecular evolutionary theory, nor experimental studies or observations from molecular phylogenies support a

close association between molecular and morphological rates of change (Bromham & Hendy (2000).

Carroll (2005) argued that we should distinguish between the evolution of anatomy and physiology. He argued that morphological evolution such as changes in size, shape, number, or colour patterns of morphological characters are fundamentally different from changes in chemistry of physiological processes. There is sufficient evidence that evolution of protein coding sequences has an influence in some physiological differences between species such as animal vision (Yokoyama 2002), respiration (Jessen et al. 1991), digestive metabolism (Zhang et al. 2002), and host defense (Hughes 2002). However, the amount of direct evidence explaining the relationship between coding, non-coding and regulatory sequences and morphological evolution is rather limited. Based on theoretical and empirical evidences, Carroll (2005) argued that when considering the evolution of form, changes in regulatory sequences should be considered as the primary hypothesis.

In order to understand the relationship between molecular and morphological evolution it is important to identify the rates at which mutations are accumulated in different lineages during the course of evolution. This topic plays an important role in neutralist/selections debate of molecular evolution (Kimura 1983; Gillespie 1991). It has been suggested that the hypothesis of the molecular clock (Zuckerkandl & Pauling 1962, 1965), that there are constant rates of molecular changes over time and across lineages, is hardly met since many species evolve faster than others, in terms of molecular changes (Bromham & Penny 2003). In contrast, Gillooly et al. (2005) suggested that there is a single molecular clock but that it 'ticks' at a constant nucleotide substitution rate per unit of mass-specific metabolic energy, rather than time. Understanding the reasons for these differences in molecular evolutionary rates across various lineages and linking them to species biology and life history traits (physiology and morphology) is a central challenge for evolutionary biologists. There are various theories linking the evolution of DNA to several physiological traits.

8.2 The Rate of DNA Evolution: Effects of Physiology

As reviewed in chapter two, several hypotheses have been proposed to explain observed differences in rates of molecular evolution. Britten 1986 suggested that the more efficient DNA repair mechanism in primates could explain the slow molecular evolution in

this lineage. In addition, increases in mutation rates have been clearly demonstrated in genetically altered mice and genomic regions controlling several human diseases. These have a deficient proof reading ability in their mtDNA polymerase enzyme (Trifunovic et al. 2004). Considering the relatively high molecular evolutionary rates reported for tuatara, perhaps understanding the biology of mtDNA such as mechanisms that control the DNA replication, transcription and translation might contribute to our understanding of the fast evolving pace of change of this genome. Although the efficiency and the accuracy of the DNA replication enzymes in tuatara mtDNA could not be tested with the data available from this project, this would be an appropriate subject of further investigation in order to explain the high molecular evolutionary rates in tuatara.

A generation time effect has been proposed (Wu & Li 1985; Bromham et al. 1996; Martin & Palumbi 1993; Mooers & Harvey 1994) according to the assumption that in species with shorter generation times there is a larger number of DNA replication per unit time and consequently a greater chance of replication error per unit time. The metabolic rate hypothesis (Martin & Palumbi 1993; Nabholz et al. 2008 a; Gillooly et al. 2005; Allan et al. 2006; Wright et al. 2006) proposes that increased metabolic rates in organism results in elevated levels of DNA-damaging metabolites and consequently increased mutation rates. Moreover, the inverse correlation between mutation rates and body size (Martin & Palumbi 1993; Gillooly et al. 2005; Estabrook et al. 2007) and/or longevity (Barja & Herrero 2000; Kujoth et al. 2007; Nabholz et al. 2008 a; Galtier et al. 2009) has also been proposed.

It is unlikely that a single factor could entirely explain variation in rates of molecular evolution and any correlation with phenotypic changes. For example, Gillooly et al. (2007) suggested that body size and temperature, through their effect on metabolic rates, may play an important role in both genotypic and phenotypic evolution. In addition, Rottenberg (2006, 2007 ab) have proposed a hypothesis linking the mitochondrial protein evolutionary rates to species longevity in mammals and birds. Other contributing factors such as base pair composition, transcription and the efficiency of DNA repair may also influence rates of protein evolution (Pal et al. 2006). Gillooly et al. (2007) indicated that the evolutionary rates of protein coding genes such as NADH, cytb and haemoglobin are independent of body size, and thus presumably of effective population size (Lynch & Conery 2003). The effective population size is expected to influence the evolution of the

genes under positive selection (Ohta 1973; Lynch & Conery 2003; Keightley et al. 2005). Therefore they suggested that neutral processes, which in turn are strongly influenced by individual metabolic rate, mainly control the rates of mitochondrial protein evolution. However, the extent to which amino acid changes lead to morphological changes remains unclear. Some studies indicate that such changes often modify the structure and the function of proteins and consequently lead to morphological evolution (Bloom et al. 2005; Ma et al. 2006). When a gene is mutated so that its protein can no longer carry out its normal structure and function, a disorder can result. Some examples are sickle cell anemia, cystic fibrosis and Huntington's disease (Pauling et al. 1949; Andersen 1938; MacDonald et al. 1993). While such changes may not have any significant effect on the fitness of individuals in the short term, they may nevertheless affect the genotype of individuals and eventually lead to phenotypic evolution in the longer term (Eyre-Walker & Keightley 2007).

The results of this study, showing a relatively high molecular evolutionary rates in tuatara mtDNA, is in contrast to stable morphology of this species over the long periods of evolutionary time. This is accompanied by slow rates of growth, a slow metabolic rate, low body temperature, long generation time and a slow reproductive rate that are characteristic of the species. This result suggests that life history and physiological traits are not the only parameters affecting the rates of molecular evolution.

8.3 Energy and the Tempo of Evolution

The Evolutionary speed hypothesis (ESH) proposed by Rench (1959) and extended by Rohde (1978, 1992) directly links the genetic responses of animals to environmental change. This hypothesis could potentially explain patterns of biodiversity at the global level, with a decrease in species numbers along the climate gradient from the tropics to the poles. The cause of this global pattern has been the subject of much speculation. Rohde's hypothesis suggests a link between higher temperature and solar radiation in lower latitude and high rates of speciation in equatorial regions, via two major processes. Warmer ecological environments characterized by high levels of solar radiation could have a direct mutagenic effect on DNA. In turn, this might be expected to result in higher mutation rates and consequently faster rates of molecular evolution. Increased mutation rates are argued to lead to rapid 'reproductive isolation' between populations and eventually to the

promotion of rapid speciation (Rohde 1978, 1992; Wright et al. 2006). This would then be expected to lead to increased levels of divergence between species, at the molecular level. Second, it has been hypothesized that higher temperatures might increase individual growth rates, decreasing the generation times, increasing the speed at which the selection operates and elevating rates of speciation (Rohde 1978, 1992) (Figure 8.2). A shorter generation times have been associated with faster molecular evolution (Martin & Palumbi 1993, Bromham et al. 1996; Bromham 2002), this pathway could also potentially promoting rapid molecular evolution and speciation towards lower latitudes.

An alternative explanation for faster rates of DNA evolution in warmer climates is that, because of putatively smaller population size at low latitudes, warmer climate species might be subject to greater rates of nearly neutral genetic drifts (Stevens 1989; Ohta 1992). However this theory is not supported by recent empirical studies (Wright et al. 2006, 2009; Gillman et al. 2009).

There is a growing body of evidence that support the Rohde's hypothesis for plants, marine foraminifera, terrestrial mammals, birds, amphibians and fishes (Gillman et al. 2010; Davies et al. 2004; Wright et al. 2006; Allen et al. 2006; Gillman et al. 2009; Bleiweiss 1998; Wright et al. 2010; Wright et al. 2011). However this positive relationship between the rate of genetic evolution and ambient temperature was not found among microbial thermophiles and mesophiles (Drake 2009; Swami 2009). It has been demonstrated that in extremely high temperature, thermophiles show a reduction in genetic substitution rates. This suggests that the slower genetic evolution in microbial thermophiles is a result of mutation control mechanisms in extreme thermal regimes.

Furthermore, Bromham and Cardillo (2003), using congeneric pairs of birds have detected no significant difference in evolutionary rates between high and low latitude avian species (Bromham & Cardillo 2003). Wright and his colleagues (2006) argued that the power of this study was affected by including species pairs with overlapping distributions (up to 100%) when the seasonal migration is taken into analysis (Wright et al. 2006). They suggested that species that experience tropical or subtropical winter climates cannot be defined as temperate in terms of their temperature regime. They explained that this feature is more related to their incursions and migration into higher latitudes in summer and low latitudes in winter (Wright et al. 2006). Therefore, the result of study by Bromham and Cardillo (2003) could not potentially reject the correlation between molecular evolution

and ambient temperature because these congeneric bird species are migrating between high and low latitude seasonally.

High temperature and solar radiation

Elevated mutation rate

Faster growth of individuals

Faster molecular evolution

Shorter generations

Faster selection

Higher speciation rate
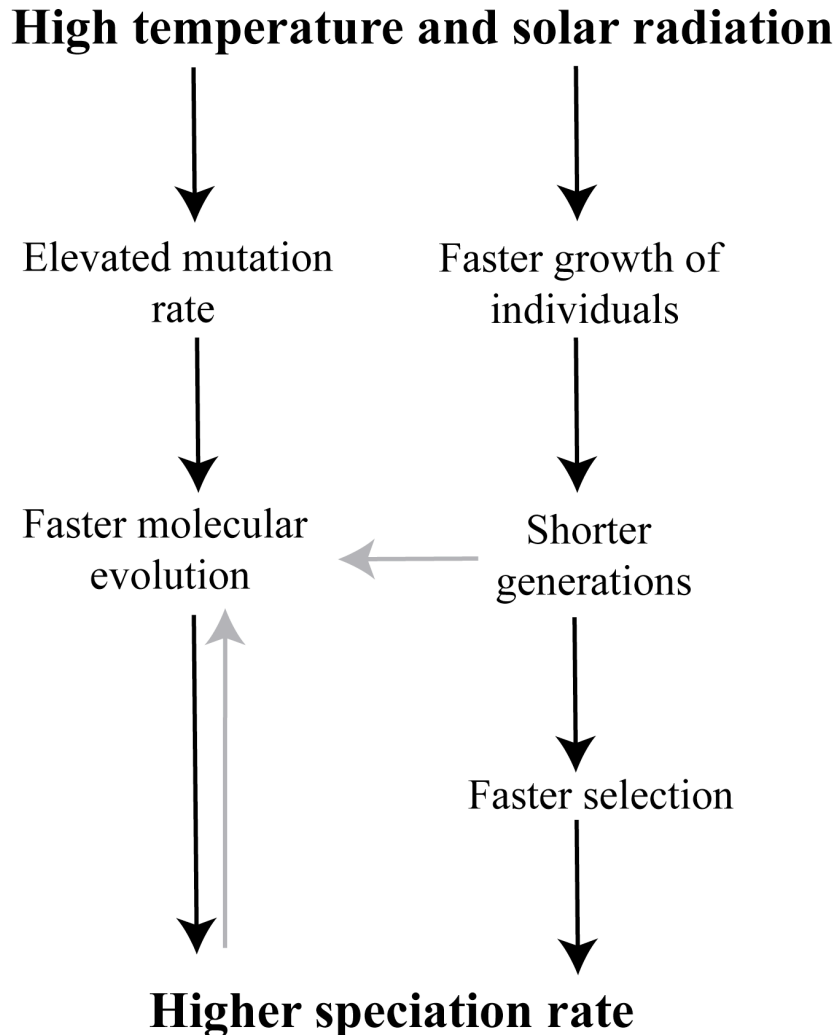
Figure 8.2: The climate-speciation hypothesis (Rohde, 1978, 1992) links higher temperatures and solar radiation levels at low latitudes with higher speciation rate, via two pathways. First, higher temperatures may increase mutation rates, and subsequently the rate of molecular evolution. The latter might increase the rate of speciation. Secondly, higher temperatures and more solar radiation might increase growth rates; decrease generation time and potentially increasing the speciation rates. The gray lines indicate the links between generation time, evolutionary rates and speciation that were not explicitly discussed by Rohde. The figure is re-drawn by the Author of the thesis.

The alternative hypothesis to ESH is the tropical conservatism hypothesis (TCH) that assumes the rates of molecular evolution and diversification of clades are independent of latitudes and elevations (Wiens et al. 2006). Consistent with the THC, and in

contradiction to the ESH, Wiens et al. (2006) found no evidence that diversification rates (number of species) for Hylidae (tree frogs) were greater in clades that are located in lower latitudes. Wright et al. (2010) argued that these results suffer from limitations in latitudinal breadth of clade distribution (Algar et al. 2009; Wright et al. 2010). They explained that seven out of the 11 clades examined by Wiens et al. (2006) span the equator to north and south and extended into temperate latitudes. There is only one clade in their study with entirely extra-tropical distribution and this clade actually shows the lowest diversification rate they reported. Another limitation for this study is the method used by Wiens et al. (2006). They assume that rates of DNA divergence can be used to estimate time since divergence and to subsequently estimate the speciation rates. The latter was then compared across latitudes. This approach assumes that the rate of molecular evolution is not related to changes in latitude and climate. However, Wright et al. (2010) tested for rate heterogeneity in molecular DNA evolution in Hylidae. In contrast, they suggested that species of Hylidae in warmer climates have undergone faster DNA evolution than species in cooler climates. Therefore, the slower rates of molecular evolution in Hylidae species in cooler climate indicates that the age of cooler temperature Hylidae clades have been underestimated by Wiens et al. (2006) and subsequently the diversification rates overestimated.

8.4 Diversification and Molecular Evolutionary Rates

Diversification is the net result of addition of species by speciation and elimination of species by extinction in the environment. A positive correlation between diversification rate and nucleotide substitution rate is reported over a range of taxa (Webster et al. 2003), including angiosperms (Lancaster 2010), birds (Lanfear et al. 2010) and plants (Davies et al. 2004; Barraclough & Savolainen 2001; Jobson & Albert 2002).

There are different hypothesis to explain the correlation between the rates of molecular evolution and net diversification. Higher rates of molecular evolution might increase the net diversification via increasing speciation rates (Barraclough et al. 1996; Venditti & Pagel 2009; Pagel et al. 2006) and/or decreasing extinction rates (Franklin & Frankham 1998).

The speciation process is often associated with factors that increase the molecular evolutionary rates, such as adaptation to new environments and temporary reductions in population size (Venditti & Pagel 2009). Since a comprehensive molecular data based on

mtDNA and microsatellites clearly suggest that there is only a single species of tuatara (This thesis, Hay et al. 2010), speciation is unlikely to be a cause of high molecular evolutionary rates detected in this iconic taxon. The other possibility is that diversification might increase owing to a decrease in extinction rates as a result of the active maintenance of genetic variation in populations (Franklin & Frankham 1998). If this is the case, the higher mutation rates could be associated with lower extinction rates and ultimately might lead to higher diversification rates. The high evolutionary rates reported for tuatara mtDNA might be one of the mechanisms for maintaining the genetic variation in tuatara populations in order to escape from extinction. However, that is a speculation since the sample size and the evolutionary time frame in this study is not long enough.

8.5 Population Size and Molecular Evolutionary Rates

The nearly neutral theory of molecular evolution predicts the higher molecular evolution in smaller population (Ohta 1972). However, the alternative theory proposed by Fisher (1930) suggests that the molecular evolution should proceed faster in larger populations as a result of greater effect of natural selection in larger populations.

A recent comprehensive molecular study of a global dataset of cytb gene sequences belonging to seven orders of birds indicated that species found on larger landmasses have higher rates of molecular evolution (Wright et al. 2009). This suggests that confining species to a limited refugia decreases the genetic variation and consequently decrease the rate and effectiveness of adaptive evolution in response to changing environment e.g. global warming. This is potentially important for conservation biology. Similarly, Hawks et al. (2007) suggested that there is a positive correlation between human population size and adaptive evolution.

In the case of tuatara, small island populations may have suffered a sever population bottleneck, associated with island colonization, and a long-term reduction in population size. However, due to difficulties in producing adequate molecular data representing both ancient (mainland populations) and modern (island populations) tuatara populations, the relationship between evolutionary rates and population size is not clear. Further empirical examination of this relationship might partially explain the relatively high molecular evolutionary rates estimated for tuatara mtDNA.

8.6 Conclusion

In this study, the complete mitochondrial genomes of nine ancient and 42 contemporary tuatara (*Sphenodon,* Reptilia) samples were examined in order to examine the relationship between rates of molecular and morphological evolution in tuatara. The molecular rates for tuatara reported in this study generally support previous suggestions (Hay et al. 2008; Subramanian et al. 2009 b) that this species is characterized by high rates of molecular change in the HVR regions of the mitochondrial genome level. In comparison, the molecular evolutionary rates for tuatara reported in this study are about 5 - 10 times higher than those estimated for Adélie penguins based on similar analytical methods (Subramanian et al. 2009 a) (Figure 6.11). In this study, I explored that the high level of heteroplasmic variants detected in mitochondrial genome were at least one of the factors that contribute to the high molecular rates estimated in tuatara. Whilst these rates are still likely to represent overestimates, we can be at least confident that the species is characterized by high molecular rates generally.

In conclusion, although tuatara have remained largely physically unchanged over tens of million of years, they are generally evolving fast in the level of mtDNA evolution. This is not surprising, since the processes that govern skeletal morphology are very different to the processes that govern molecular changes.

Gillespie 1991 suggested that there is an association between adaptation and rates of molecular evolution. Selection for a particular adaptation could result in nucleotide changes in genes associated with that particular trait and thereby increases the molecular rates. For example, rapidly evolving viruses show high rates of molecular evolution (Shackelton et al. 2005; Duffy & Holmes 2008, 2009). However, a measurable phenotypic change is likely to affect only a handful of nucleotide sites in one or a few coding or regulatory gene (Majerus & Mundy 2003; Fondon & Garner 2004; Galant & Carroll 2002; Wittkopp et al. 2002; Sucena et al. 2003; Durbin et al. 2003). Furthermore, many of the observed molecular changes are unlinked to adaptive evolution and occur in some nucleotide sites that do not affect the protein and RNA products and have no effect on fitness of individual (Kimura 1983). These findings support the hypothesis that rates of neutral molecular and phenotypic evolution are decoupled (King & Wilson 1975). However, the association between molecular and morphological changes in tuatara deserves further empirical investigation especially in the area of nuclear genome.

Reference List

Allen AP, Gillooly JF, Savage VM, Brown JH. Kinetic effects of temperature on rates of genetic divergence and speciation. Proc. Natl. Acad. Sci. U.S. A. 2006, 103: 9130–9135

Algar AC, Kerr JT, Currie DJ. Evolutionary constraints on regional faunas: whom, but not how many. Ecol Lett. 2008, [Epub ahead of print]

Andersen DH. Cystic fibrosis of the pancreas and its relation to celiac disease: a clinical and pathological study. Am J Dis Child 1938, 56:344-399

Anderson WA. Structure and fate of the paternal mitochondrion during early embryogenesis of Paracentrotus lividus. J Ultrastruct Res. 1968, 24: 311-321

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. Nature. 1981, 290:457-465

Apesteguía S and Novas FE. Large Cretaceous Sphenodontian from Patagonia provides insight into Lepidosaur evolution in Gondwana. Nature, 2003, 425:609-612

Aris-Brosou S and Yang Z. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Syst Biol. 2002, 51:703–714

Armour JA, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, Kidd KK, Bertranpetit J, Pääbo S, Jeffreys AJ. Minisatellite diversity supports a recent African origin for modern humans. Nat Genet. 1996, 13:154–160

Attardi G, Albring M, Amalric F, Gelfand R, Griffith J, Lynch D, Merkel C, Murphy W, Ojala D. In: Bücher Th, Neupert W, Sebald W, Werner S, Editors, Genetics and Biogenesis of Chloroplasts and Mitochondria. North-Holland, Amsterdam, 1976, pp: 573-585

Avise JC, Lansman RA, Shade RO. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus Peromyscus. Genetics. 1979, 92:279-295

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst.1987, 18:489–522

Avise JC, Bowen BW, Lamb T, Meylan AB, Bermingham E. Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Mol Biol Evol. 1992, 9:457–473

Ayala FJ. Genetic differentiation during the speciation process. Evol Biol. 1975, 8:1–78

Ballard JWO and Whitlock MC. The incomplete natural history of mitochondria. Mol Ecol. 2004, 13:729–744

Barja G and Herrero A. Oxidative damage to mitochondrial DNA is inversely related to maximum life span in the heart and brain of mammals. FASEB J. 2000, 14:312–318

Barraclough TG, Harvey PH, Nee S. Rate of rbcL gene sequence evolution and species diversification in flowering plants (angiosperms). Proc R Soc Lond Ser B Biol Sci.1996, 263:589-591

Barraclough TG and Savolainen V. Evolutionary rates and species diversity in flowering plants. Evolution. 2001, 55:677–683

Battey J and Clayton DA. The transcription map of mouse mitochondrial DNA. Cell. 1978, 14:143-156

Bazin E, Glemin S, Galtier N. Population size does not influence mitochondrial genetic diversity in animals. Science, 2006, 312:570–572

Bendall KE, Macaulay VA, Baker JR, Sykes BC. Heteroplasmic point mutations in the human mtDNA control region. Am J Hum Genet. 1996, 59:1276–1287

Benton MJ. Reptilia. In: Benton MJ, Editor, The Fossil Record 2. Chapman and Hall, London, 1993, pp: 681-715

Benton MJ. Vertebrate Paleaontology (2nd edn). Blackwell Science, Oxford. 2000

Binladen J, Wiuf C, Gilbert MT, Bunce M, Barnett R, Larson G, Greenwood AD, Haile J, Ho SY, Hansen AJ, Willerslev E. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Genetics. 2006.172:733-741

Bleiweiss R. Slow rate of molecular evolution in high-elevation hummingbirds. Proc. Natl. Acad. Sci. U.S.A. 1998, 95:612-616.

Blin N and Stafford DW. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res. 1976, 3:2303-2308

Bloom JD, Silberg JJ, Wilke CO, Drummond DA, Adami C, Arnold FH. Thermodynamic prediction of protein neutrality. Proc. Natl. Acad. Sci. U.S.A. 2005, 102:606–611

Bousquet J, Strauss SH, Li P. Complete congruence between morphological and rbcL-based phylogenies in birches and related species (Betulaceae). Mol Biol Evol. 1992 a, 9:1076–1088

Bousquet J, Strauss SH, Doerksen AH, Price RL. 1992b. Extensive variation in evolutionary rate of rbcL gene sequences among seed plants. Proc Nat Acad Sci U S A. 1992 b. 89: 7844-7848

Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S. Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. U.S.A. 2007, 104:14616-21

Britten RJ. Rates of DNA sequence evolution differ between taxonomic groups. Science. 1986, 231:1393-1398

Brotherton P, Endicott P, Sanchez JJ, Beaumont M, Barnett R, Austin J, Cooper A. Novel high-resolution characterization of ancient DNA reveals C > U type base modification events as the sole cause of post mortem miscoding lesions. Nucleic Acids Res. 2007, 35:5717-5728

Bromham L, Rambaut A, Harvey PH. Determinants of rate variation in mammalian DNA sequence evolution. J Mol Evol. 1996, 43:610–621

Bromham L and Hendy MD. Can fast early rates reconcile molecular dates to the Cambrian explosion? Proc. R. Soc. Lond. B. 2000, 267: 1041-1047

Bromham L. Molecular clock in reptiles: life history influences rate of molecular evolution. Mol. Biol. Evol. 2002, 19:203-209

Bromham L, Woolfit M, Lee MSY, Rambaut A. Testing the relationship between morphological and molecular rates of changes along phylogenies. Evolution. 2002, 56:1921–1930

Bromham L, Eyre-Walker A, Smith NH, Maynard Smith J. Mitochondrial Steve: paternal inheritance of mitochondria in humans. Trends Ecol Evol. 2003, 18:2–4

Bromham L and Penny D. The modern molecular clock. Nat Rev Genet. 2003, 4:216–224

Bromham L and Cardillo M. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. J Evol Biol. 2003, 16:200-207

Brown WM and Vinograd J. Restriction endonuclease cleavage maps of animal mitochondrial DNAs. Proc. Natl. Acad. Sci. U.S. A. 1974, 71:4617-4621

Brown WM, George M Jr, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. U.S. A. 1979, 76:1967-1971

Brown WM. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc. Natl. Acad. Sci. U.S. A. 1980, 77:3605–3609

Brown WM. The mitochondrial genome of animals. In: Macintyre RJ, Editor, Molecular Evolutionary Genetics. Plenum, New York, 1985, pp: 95–130

Brown DT, Samuels DC, Michael EM, Turnbull DM, Chinnery PF. Random genetic drift determines the level of mutant mtDNA in human primary oocytes. Am J Hum Genet. 2001, 68:533–536

Buller WL. Notice of a new variety of tuatara lizard (Sphenodon) from East Cape Island. Transactions and Proceedings of the New Zealand Institute. 1878, 10:220-221

Buller WL. Illustration of Darwinism. Transactions and Proceedings of the New Zealand Institute. 1894, 11:349–351

Cáceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, Lockhart DJ, Preuss TM, Barlow C. Elevated gene expression levels distinguish human from non-human primate brains. Proc. Natl. Acad. Sci. U.S. A. 2003, 100:13030–13035

Cann RL, Stoneking M, WilsonAC. Mitochondrial DNA and human evolution. Nature. 1987, 325:31–36

Carroll RL. Vertebrate Paleontology and Evolution. Freeman WH and Company, London, 1988

Carroll SB. Evolution at two levels: on genes and form. PLoS Biol. 2005, 3: e245 doi:10.1371/journal.pbio.0030245, Review

Cherry LM, Case SM, Wilson AC. Frog perspective on the morphological difference between humans and chimpanzees. Science, 1978, 200: 209–211

Chomyn A, Mariottini P, Cleeter MW, Ragan CI, Matsuno-Yagi A, Hatefi Y, Doolittle RF, Attardi G. Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature. 1985, 314:592-597

Chomyn A, Cleeter MW, Ragan CI, Riley M, Doolittle RF, Attardi G. URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit. Science. 1986, 234:614-618

Clarke LA, Rebelo CS, Gonc alves J, Boavida MG, Jordan P. PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. Mol Pathol. 2001, 54: 351–353

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, Ferriera S, Wang G, Zheng X, White TJ, Sninsky JJ, Adams MD, Cargill M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. Science. 2003, 302:1960-1963

Clayton DA, Doda JN, Friedberg EC. The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. Proc. Natl. Acad. Sci. U.S. A. 1974, 71:2777-81

Cooper A and Penny D. Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. Science. 1997, 275:1109–1113

Cooper A and Poinar HN. Ancient DNA: do it right or not at all. Science. 2000, 289:1139

Coyne JA, Simeonidisa S, Rooneya P. Relative Paucity of Genes Causing Inviability in Hybrids Between Drosophila melanogaster and D. simulans. Genetics. 1998, 150: 1091–1103

Cree A, Guillette LJ Jr, Cockrem JF, Brown MA, Chambers GK. Absence of daily cycles in plasma sex steroids in male and female tuatara (Sphenodon punctatus), and the effects of acute capture stress on females. Gen Comp Endocrinol. 1990, 79:103– 113

Cree A, John FC, Mark AB, Peter RW, Louis JG, Donald GN, Geoffrey KC. Laparoscopy, radiography, and blood analyses as techniques for identifying the reproductive condition of female tuatara. Herpetologica. 1991, 47:238–249

Cree A and Butler D. Tuatara recovery plan. Threatened species recovery plan 9. Department of Conservation, Wellington, 1993

Cree A. Low and annual reproductive output in female reptiles from New Zealand. New Zealand Journal of Zoology. 1994, 21:351–372

Cree A, Thompson MB, Daugherty CH. Tuatara sex determination. Nature. 1995, 375: 543

Crook IG. Tuatara. Wildlife-A Review. 1970, 2:16–19

Crook IG. The tuatara, Sphenodon punctatus Gray, on islands with and without populations of the Polynesian rat, Rattus exulans (Peale). Proceedings of the New Zealand Ecological Society. 1973, 20:115–120

Crook IG. The tuatara. In: Kuschel G, Editor, Biogeography and Ecology in New Zealand. Junk, The Hague, 1975, pp: 331–352

Cutler DJ. Estimating divergence times in the presence of an over dispersed molecular clock. Mol Biol Evol. 2000, 17:1647–1660

Drake JW. Avoiding dangerous missense: thermophiles display especially low mutation rates. PLoS Genet. 2009, 5:e1000520

Darwin C. The origin of species by means of natural selection. John Murray, London, 1859

Daugherty CH, Patterson GB, Thorn CJ, French DC. Differentiation of the Members of the New Zealand Leiolopisma nigriplantare Species Complex (Lacertilia: Scincidae). Herpetological Monographs. 1990 a. 4:61-76

Daugherty CH, Cree A, Hay JM, Thompson MB. Neglected taxonomy and continuing extinctions of tuatara (Sphenodon). Nature. 1990 b, 347: 177–179

Davies TJ, Savolainen V, Chase MW, Moat J, Barraclough TG. Environmental energy and evolutionary rates in flowering plants. Proc Biol Sci. 2004, 271:2195-2200

Dawid IB, Klukas CK, Ohi S, Ramirez JL, Upholt WB. In: Saccone C and Kroon AM, Editors, The Genetic Function of Mitochondrial DNA. North-Holland, Amsterdam. 1976, pp: 3-13

Denver DR, Morris K, Lynch M, Vassilieva LL, Thomas WK. High direct estimate of the mutation rate in the mitochondrial genome of Caenorhabditis elegans. Science. 2000, 289:2342-2344

Drummond A and Rodrigo AG. Reconstructing genealogies of serial samples under the assumption of a molecular clock using serial-sample UPGMA. Mol Biol Evol. 2000, 17: 1807-1815

Drummond AJ and Rambaut A. BEAST. University of Oxford. Oxford. 2003

Drummond AJ and Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007, 7: 214

Duffy S and Holmes EC. Phylogenetic evidence for rapid rates of molecular evolution in the single-stranded DNA begomovirus tomato yellow leaf curl virus. J Virol. 2008, 82:957-65

Duffy S and Holmes EC. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. J. Gen. Virol. 2009, 90: 1539-1547

Durbin ML, Lundy KE, Morrell PL, Torres-Martinez CL, Clegg MT. Genes that determine flower color: The role of regulatory changes in the evolution of phenotypic adaptations. Mol Phylogenet Evol. 2003, 29: 507–518

Duncan RP, Blackburn TM, Worthy TH. Prehistoric bird extinctions and human hunting. Proc Biol Sci. 2002, 269:517-521

Edwards CJ, Bollongino R, Scheu A. Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. Proc Biol Sci. 2007, 274:1377–1385

Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Pääbo S. Molecular evolution of FOXP2, a gene involved in speech and language. Nature. 2002, 418:869-872

Emerson BC. Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates. Syst Biol. 2007, 56:337-345

Endicott P and Ho SY. A Bayesian evaluation of human mitochondrial substitution rates. Am J Hum Genet. 2008, 82:895-902

Eo SH and DeWoody JA. Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. Proc Biol Sci. 2010, published a head of print

Estabrook GF, Smith GR, Dowling TE. Body mass and temperature influence rates of mitochondrial DNA evolution in North American cyprinid fish. Evolution. 2007 61:1176-1187

Excoffier L. Human demographic history: refining the recent African origin model. Curr Opin Genet Dev. 2002,12:675-682

Eyre-Walker A, Smith N, Smith J. How clonal are human mitochondria? Proc Biol Sci.1999, 266:477–483

Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. Nat Rev Genet. 2007, 8:610-8, Review

Ferguson JWH. On the use of genetic divergence for identifying species. Biol J Linn Soc. 2002, 75:509–516

Filosto M, Mancuso M, Vives-Bauza C, Vila MR, Shanske S, Hirano M, Andreu AL, DiMauro S. Lack of paternal inheritance of muscle mitochondrial DNA in sporadic mitochondrial myopathies. Ann Neurol. 2003, 54:524–526

Fisher RA. The genetical theory of neutral selection: a complete variorum edition. Oxford Uni. Press, Oxford. 1930

Fleischer RC, McIntosh CE, Tarr CL. Evolution on a volcanic conveyor belt: using phylogeo- graphic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Mol Ecol. 1998, 7: 533-545

Fondon JW and Garner HR. Molecular origins of rapid and continuous morphological evolution. Proc. Natl. Acad. Sci. U.S.A. 2004, 101: 18058–18063.

Franklin IR and Frankham R. How large must populations be to retain evolutionary potential? Anim Conserv. 1998, 1:69–70

Fraser NC. The osteology and relationships of Clevosaurus (Reptilia: Sphenodontida). Phil Trans R Soc Lond. 1988, 321:125-178

Frost DR and Hillis DM. Species in concept and practice: Herpetological considerations. Herpetologica. 1990. 46:87–104

Galant R and Carroll SB. Evolution of a novel transcriptional repression domain in the Ultrabithorax Hox protein in insects. Nature. 2002, 415: 848–849

Galtier N, Enard D, Radondy Y, Bazin E, Belkhir K. Mutation hot spots in mammalian mitochondrial DNA. Genome Res. 2006, 16:215–222

Galtier N, Blier PU, Nabholz B. Inverse relationship between longevity and evolutionary rate of mitochondrial proteins in mammals and birds. Mitochondrion. 2009, 9:51-57

Garcia-Moreno J. Is there a universal mtDNA clock for birds? J Avian Biol. 2004, 35:465-468

Gaze P. Tuatara recovery plan. Threatened species recovery plan 47. Department of Conservation, Wellington, 2001

Gilbert MT, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, Lynnerup N, Cooper A. Characterization of genetic miscoding lesions caused by postmortem damage. Am J Hum Genet. 2003, 72:48-61

Gillespie JH. The Causes of Molecular Evolution. Oxford University Press, Oxford, 1991

Gillman LN, Keeling DJ, Ross HA, Wright SD. Latitude, elevation and the tempo of molecular evolution in mammals. Proc Biol Sci. 2009, 276:3353-9

Gillman LN, Keeling DJ, Gardner RC, Wright SD. Faster evolution of highly conserved DNA in tropical plants. J Evol Biol. 2010, 23:1327-1330

Gillooly JF, Allen AP, West GB, Brown JH. The rate of DNA evolution: effects of body size and temperature on the molecular clock. Proc. Natl. Acad. Sci. U.S.A. 2005, 102:140–145

Gillooly JF, McCoy MW, Allen AP. Effects of metabolic rate on protein evolution. Biol Lett. 2007, 3:655-659

Goddard JM and Wolstenholme DR. Origin and direction of replication in mitochondrial DNA molecules from Drosophila melanogaster. Proc. Natl. Acad. Sci. U.S.A. 1978, 75:3886-3890

Goode M and Rodrigo AG. Using PEBBLE for the evolutionary analysis of serially sampled molecular sequences. Curr Protoc Bioinformatics. 2004, Chapter 6:Unit 6.8

Gray JE. Note on a peculiar structure in the head of an agama. Zoological Miscellany. 1831, 1:13-14

Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, Du L, Egholm M, Rothberg JM, Paunovic M, Pääbo S. Analysis of one million base pairs of Neanderthal DNA. Nature. 2006, 444:330–336

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, et al. A draft sequence of the Neandertal genome. Science. 2010, 328:710-22

Groombridge B. The IUCN Amphibia-Reptilia Red Data Book. IUCN, Gland, Switzerland, 1982

Grossman LI, Watson R, Vinograd J. The presence of ribonucleotides in mature closed-circular mitochondrial DNA. Proc. Natl. Acad. Sci. U.S.A. 1973, 70:3339-3343

Guo J, Du LX, Ma YH, Guan WJ, Li HB, Zhao QJ, Li X, Rao SQ. A novel maternal lineage revealed in sheep (Ovis aries). Anim Genet. 2005, 36:331–336

Günther A. Contribution to the anatomy of Hatteria (Rhynchocephalus Owen). Phil Trans of the R Soc of London. 1867, 157: 595-629

Gyllensten U, Wharton D, Josefsson A, Wilson AC. Paternal inheritance of mitochondrial DNA in mice. Nature. 1991, 352: 255–257

Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 1999, 41:95-98

Hallidy T and Adler K. The New Encyclopaedia of Reptiles and Amphibians. Oxford University Press, Oxford, 2002

Harper GL, Maclean N, Goulson D. Analysis of museum specimens suggests extreme genetic drift in the adonis blue butterfly (Polyommatus bellargus). Biological Journal of the Linnean Society. 2006, 88, 447–452

Harrison RG. Mitochondrial DNA as a genetic marker in population and evolutionary genetics. Trends Ecol Evol. 1989, 4:6–11

Hartmann A, Thieme M, Nanduri LK, Stempfl T, Moehle C, Kivisild T, Oefner PJ. Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. Hum Mutat. 2009, 30:115-22

Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 1985, 22:160-174

Hasegawa M, Di Rienzo A, Kocher TD, WilsonAC. Toward a more accurate time scale for the human mitochondrial DNA tree. J Mol Evol. 1993, 37:347–354

Hatefi Y. The Mitochondrial Electron Transport and Oxidative Phosphorylation System. Ann Rev Biochem. 1985, 54:1015-1069

Hauswirth WW and Laipis PJ. Mitochondrial DNA polymorphism in a maternal lineage of Holstein cows. Proc. Natl. Acad. Sci. U.S.A. 1982, 79:4686–4690

Hawks J, Wang ET, Cochran GM, Harpending HC, Moyzis RK. Recent acceleration of human adaptive evolution. Proc. Natl. Acad. Sci. U.S.A. 2007, 104:20753-20758

Hay JM, Daugherty CH, Cree A, Maxson LR. Low genetic divergence obscures phylogeny among populations of Sphenodon, remnant of an ancient reptile lineage. Mol Phylogenet Evol. 2003, 29:1-19

Hay JM, Sarre SD, Daugherty CH. Nuclear mitochondrial pseudogenes as molecular outgroups for phylogenetically isolated taxa: a case study in Sphenodon. Heredity. 2004, 93:468-475

Hay JM, Subramanian S, Millar CD, Mohandesan E, Lambert DM. Rapid molecular evolution in a living fossil. Trends Genet. 2008, 24:106-109

Hay JM, Sarre SD, Lambert DM, Allendorf FW, Daugherty CH. Genetic diversity and taxonomy: a reassessment of species designation in tuatara (Sphenodon: Reptilia). Conserv Genet. 2010, 11:1063–1081

He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, Diaz LA Jr, Kinzler KW, Vogelstein B, Papadopoulos N. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells.Nature. 2010, 464:610-4

Hecht NB, Liem H, Kleene KC, Distel RJ, Ho SM. Maternal inheritance of the mouse mitochondrial genome is not mediated by a loss or gross alteration of the paternal mitochondrial DNA or by methylation of the oocyte mitochondrial DNA. Dev Biol. 1984, 102:452-461

Hendy MD, Woodhams MD, Dodd A. Modelling mitochondrial site polymorphisms to infer the number of segregating units and mutation rate. Biol Lett. 2009 5:397-400

Heyer E, Zietkiewicz E, Rochowski A, Yotova V, Puymirat J, Labuda D. Phylogenetic and familial estimates of mitochondrial substitution rates: study of control region mutations in deep-rooting pedigrees. Am J Hum Genet. 2001, 69:1113-1126

Highton R. Taxonomic treatment of genetically differentiated populations. Herpetologica. 1990. 46:114–121.

Hofreiter M, Jaenicke V, Serre D, Haeseler Av A, Pääbo S. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 2001, 29:4793-4799

Hollocher H, Ting CT, Wu ML, Wu C-I. Incipient speciation by sexual isolation in Drosophila melanogaster: extensive genetic divergence without reinforcement. Genetics. 1997a, 147: 1191–1201

Hollocher H, Ting CT, Wu C-I. Incipient speciation by sexual isolation in Drosophila melanogaster: variation in mating preference and correlation between sexes. Evolution. 1997b, 51: 1175–1181

Höss M, Jaruga P, Zastawny TH, Dizdaroglu M, Pääbo S. DNA damage and DNA sequence retrieval from ancient tissues. Nucleic Acids Res. 1996, 24:1304–1307

Hovers E, Ilani S, Bar-Yosef O, Vandermeersch B. An early case of color symbolism–Ochre use by modern humans in Qafzeh cave. Curr Anthropol. 2003, 44:491–522

Ho SY, Phillips MJ, Drummond AJ, Cooper A. Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. Mol Biol Evol. 2005 a, 22:1355-1363

Ho SY, Phillips MJ, Cooper A, Drummond AJ. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol Biol Evol. 2005 b, 22:1561-1568

Ho SY and Larson G. Molecular clocks: when times are a-changin'. Trends Genet. 2006, 22: 79-83

Ho SY, Shapiro B, Phillips MJ, Cooper A, Drummond AJ. Evidence for time dependency of molecular rate estimates. Syst Biol. 2007 a, 56: 515-522

Ho SY, Kolokotronis SO, Allaby RG. Elevated substitution rates estimated from ancient DNA sequences. Biol Lett. 2007 b, 3:702-705

Ho SY. Calibrating molecular estimates of substitution rates and divergence times in birds. Journal of Avian Biology. 2007, 38: 409-414

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc. Natl. Acad. Sci. U. S. A. 1995, 92:532-536

Howard DJ. Electrophoretic survey of eastern North American Allonemobius (Orthoptera: Gryllidae): evolutionary relationships and the discovery of three new species. Annals of the Entomological Society of America. 1983, 76: 1014– 1021

Howard DJ, Waring GL, Tibbets CA, Gregory PG. Survival of hybrids in a mosaic hybrid zone. Evolution. 1993, 47: 789–800

Howell N, Kubacka I, Mackey DA. How rapidly does the human mitochondrial genome evolve? Am J Hum Genet. 1996,59:501-509

Howell N, Smejkal CB, Mackey DA, Chinnery PF, Turnbull DM, Herrnstadt C. The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. Am J Hum Genet. 2003, 72:659-670

Huang B, Li W, Zhao B, Xia C, Liang R, Ruan K, Jing N, Jin Y. MicroRNA expression profiling during neural differentiation of mouse embryonic carcinoma P19 cells. Acta Biochim Biophys Sin. 2009, 41: 231-236

Huelsenbeck JP, Larget B, Swofford DL. A compound Poisson process for relaxing the molecular clock. Genetics. 2000, 154:1879–1892

Hughes S and Mouchiroud D. High evolutionary rates in nuclear genes of squamates. J. Mol Evol. 2001, 53:70–76

Hughes AL. Natural selection and the diversification of vertebrae immune effectors. Immunol Rev. 2002, 190: 161–168

Hurst GDD and Jiggins FM. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc Biol Sci. 2005, 272:1525–1534

Ingman M, Kaessmann H, Pääbo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. Nature. 2000, 408:708–713

IUCN. IUCN Red List Categories and Criteria: Version 3.1 IUCN Species Survival Commission. IUCN, Gland, Switzerland and Cambridge, UK, 2001

Jansen T, Forster P, Levine MA, Oelke H, Hurles M, Renfrew C, Weber J, Olek K. Mitochondrial DNA and the origins of the domestic horse. Proc. Natl. Acad. Sci. U.S. A. 2002, 99:10905–10910

Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllensten U. Mitochondrial mutation rate revisited: hot spots and polymorphism. Nat Genet. 1998, 18:109-110

Jenuth JP, Peterson AC, Fu K, Shoubridge EA. Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. Nat

Genet. 1996, 14:146–151

Jessen TH, Weber RE, Fermi G, Tame J, Braunitzer G. Adaptation of bird hemoglobins to high altitudes: Demonstration of molecular mechanism by protein engineering. Proc. Natl. Acad. Sci. U.S.A. 1991, 88: 6519–6522

Jobson RW and Albert VA. Molecular rates parallel diversification contrasts between carnivorous plant sister lineages. Cladistics-Int J Willi Hennig Soc. 2002, 18:127-136

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL. Radiation of human mitochondria DNA types analysed by restriction endonuclease cleavage patterns. J Mol Evol. 1983, 19: 255–271

Jones R. The Fifth Continent: Problems concerning The Human Colonization of Australia. Annual Review of Anthropology. 1979, 8: 445-466

Kaneda H, Hayashi J, Takahama S, Taya C, Lindahl KF, Yonekawa H. Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. Proc. Natl. Acad. Sci. U.S.A. 1995, 92:4542–4546

Kasamatsu H and Vinograd J. Replication of circular DNA in eukaryotic cells. Annu Rev Biochem. 1974, 43:695-719.

Keightley PD, Lercher MJ, Eyre-Walker A. Evidence for widespread degradation of gene control regions in hominid genomes. PLoS Biol. 2005, 3:e42 (doi:10. 1371/journal.pbio.0030042)

Khaitovich P, Muetzel B, She X, Lachmann M, Hellmann I, Dietzsch J, Steigele S, Do HH, Weiss G, Enard W, Heissig F, Arendt T, Nieselt-Struwe K, Eichler EE, Pääbo S. Regional patterns of gene expression in human and chimpanzee brains. Genome Res. 2004, 14:1462-1473

Kimura M. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, 1983

Kim KI, Lee JH, Li K, Zhang YP, Lee SS, Gongora J, Moran C. Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop

sequence polymorphism. Anim Genet. 2002, 33:19–25

King MC and Wilson AC. Evolution at two levels in humans and chimpanzees. Science. 1975, 188:107–116

King M. The Penguin History of New Zealand. Penguin Books, Auckland, 2003

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K, Passarino G, Underhill PA, Scharfe C, Torroni A, Scozzari R, Modiano D, Coppa A, de Knijff P, et al. The role of selection in the evolution of human mitochondrial genomes. Genetics. 2006, 172:373-387

Koehler CM, Lindberg GL, Brown DR, Beitz DC, Freeman AE, Mayfield JE, Myers AM. Replacement of bovine mitochondrial DNA by a sequence variant within one generation. Genetics. 1991, 129:247–255

Kondo R, Satta Y, Matsuura ET, Ishiwa H, Takahata N, Chigusa SI. Incomplete maternal transmission of mitochondrial DNA in Drosophila. Genetics. 1990, 126:657– 663

Krause J, Dear PH, Pollack JL, Slatkin M, Spriggs H, Barnes I, Lister AM, Ebersberger I, Pääbo S, Hofreiter M. Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. Nature. 2006, 439:724-727

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. Neanderthal DNA sequences and the origin of modern humans. Cell. 1997, 90:19-30

Kroon AM, Pepe G, Bakker H, Holtrop M, Bollen JE, Van Bruggen EF, Cantatore P, Terpstra P, Saccone C. The restriction fragment map of rat-liver mitochondrial DNA: a reconsideration. Biochim Biophys Acta. 1977, 478:128-145

Kujoth GC, Bradshaw P, Haroon S, Prolla T. The role of mitochondrial DNA mutations in mammalian aging. PLoS Genet. 2007, 3:161–173

Kumar S and Subramanian S. Mutation rates in mammalian genomes. Proc. Natl. Acad. Sci. U.S.A. 2002, 99:803–808

Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform. 2004, 5: 150-163

Kumar S. Molecular clocks: four decades of evolution. Nat Rev Genet. 2005, 6:654–662

Kumar S, Filipski A, Swarna V, Walker A, Hedges SB. Placing confidence limits on the molecular age of the human-chimpanzee divergence. Proc. Natl. Acad. Sci. U.S.A. 2005, 102:18842–18847

Kuramoto M. Systematic implications of hybridization experiments with some Eurasian treefrogs (genus Hyla). Copeia. 1984,1984: 609–616

Kvist L, Martens J, Nazarenko AA, Orell M. Paternal leakage of mitochondrial DNA in the great tit (Parus major). Mol Biol Evol. 2003, 20:243–247

Lamarck JB. Zoological Philosophy: An Expostion with Regard to the Natural History of Animals. Univ. of Chicago Press, Chicago, (1984 (1809)), pp. 414

Lambert DM and Paterson HE. Morphological resemblance and its relationship to genetic distance measures. Evolutionary Theory. 1982, 5:291-300

Lambert DM, Ritchie PA, Millar CD, Holland B, Drummond AJ, Baroni C. Rates of evolution in ancient DNA from Adélie penguins. Science. 2002, 295: 2270-2273

Lamsman RA and Clayton DA. Selective nicking of mammalian mitochondrial DNA in vivo: photosensitization by incorporation of 5-bromodeoxyuridine. J Mol Biol. 1975, 99:761-776

Lancaster LT. Molecular evolutionary rates predict both extinction and speciation in temperate angiosperm lineages. BMC Evol Biol. 2010, 10:162

Lanfear R, Ho SY, Love D, Bromham L. Mutation rate is linked to diversification in birds. Proc Natl Acad Sci U S A. 2010, 107:20423-20428

Li WH, Tanimura M, Sharp P. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J Mol Evol. 1987, 25:330–342

Lima-de-Faria A. Molecular evolution and organization of the chromosome. Elsevier, Amsterdam. New York. Oxford, 1983

Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J Mol Evol. 1994, 39:174-190

Lovette IJ. Mitochondrial dating and mixed support for the '2% rule' in birds. The Auk. 2004, 121:1–6

Liu J, Mercer JM, Stam LF, Gibson GC, Zeng ZB, Laurie CC. Genetic analysis of a morphological shape difference in the male genitalia of Drosophila simulans and D. mauritiana. Genetics. 1996, 142: 1129–1145

Lynch M and Conery JS. The origins of genome complexity. Science. 2003, 302: 1401–1404 (doi:10.1126/ science.1089370)

Lynch M, Koskella B, Schaack S. Mutation pressure and the evolution of organelle genomic architecture. Science. 2006, 311:1727–1730

Ma B, Guo J, Zhang H. Direct correlation between proteins' folding rates and their amino acid compositions: an ab initio folding rate prediction. Proteins. 2006, 65:362–372

Macaulay VA, Richards MB, Forster P, Bendall KE, Watson E, Sykes B, Bandelt HJ. mtDNA mutation rates--no need to panic. Am J Hum Genet. 1997, 61:983-990

Macavoy ES, McGibbon LM, Sainsbury JP, Lawrence H, Wilson CA, Daugherty CH, Chambers GK. Genetic variation in island populations of tuatara (Sphenodon spp) inferred from microsatellite markers. Conservation Genetics. 2006, 8:305-318

MacDonald ME, Ambrose1 CM, Duyao1 MP, Myers RH, Lin C, Srinidhi L, Barnes G, Taylor SA, James M, et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 1993, 72: 971–983

Macreadie IG, Novitski CE, Maxwell RJ, John U, Ooi BG, McMullen GL, Lukins HB, Linnane AW, Nagley P. Biogenesis of mitochondria: the mitochondrial gene (aap1) coding for mitochondrial ATPase subunit 8 in Saccharomyces cerevisiae. Nucleic Acids Res. 1983, 11:4435-4451

Majerus MEN and Mundy NI. Mammalian melanism: Natural selection in black and white. Trends Genet. 2003, 19: 585–588

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005, 437:376–380

Marko PB. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. Mol Biol Evol. 2002, 19: 2005–2021

Martin AP, Naylor G, Palumbi S. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature.1992, 357:153-155

Martin AP and Palumbi SR. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. U.S.A. 1993, 90:4087–4091

Martin AP. Metabolic rate and directional nucleotide substitution in animal mitochondrial DNA. Mol Biol Evol.1995, 12:1124-1131

Martinez-Cruz B, Godoy JA, Negro JJ. Population fragmentation leads to spatial and temporal genetic structure in the endangered Spanish imperial eagle. Molecular Ecology, 2007, 16:477–486

May RM. Taxonomy as destiny. Nature. 1990. 347:129-130

McLeod BA and White BN. Tracking mtDNA Heteroplasmy through Multiple Generations in the North Atlantic Right Whale (Eubalaena glacialis). J Hered. 2010, 101:235-239

Mecham JS. Genetic relationships and reproductive isolation in southeastern frogs of the genera Pseudacris and Hyla. American Midlands Naturalist. 1965, 74: 269–308

Mellars P. The impossible coincidence: a single-species model for the origins of modern human behavior in Europe. Evol Anthropol. 2005,14, 12–27

Melton T, Clifford S, Martinson J, Batzer M, Stoneking M. Genetic evidence for the

proto-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal tribes. Am J Hum Genet. 1998, 63:1807–1823

Mertens R and Wermuth H. Zool Jahrb. 1954, 83:1–413

Meusel MS and Moritz RF. Transfer of paternal mitochondrial DNA during fertilization of honeybee (Apis mellifera L.) eggs. Curr Genet. 1993, 24:539-543

Meyer S, Weiss G, von Haeseler A. Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. Genetics. 1999, 152:1103-1110

Meyer M, Stenzel U, Myles S, Prüfer K, Hofreiter M. Targeted high-throughput sequencing of tagged nucleic acid samples. Nucleic Acids Res. 2007 35: e97 doi:10.1093/nar/gkm566.

Meyer M, Stenzel U, Hofreiter M. Parallel tagged sequencing on the 454 platform. Nat Protoc. 2008, 3: 267–278

Michaels GS, Hauswirth WW, Laipis PJ. Mitochondrial DNA copy number in bovine oocytes and somatic cells. Dev Biol. 1982, 94:246-251

Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G. Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. Science. 1999, 286:774-779

Mitchell NJ, Allendorf FW, Keall SN, Daugherty CH, Nelson NJ. Demographic effects of temperature-dependent sex determination: will tuatara survive global warming? Global Change Biology. 2010, 16:60–72

Millar CD, Huynen L, Subramanian S, Mohandesan E, Lambert DM. New developments in ancient genomics. Trends Ecol Evol. 2008 a, 23:386-393, Review

Millar CD, Dodd A, Anderson J, Gibb GC, Ritchie PA, Baroni C, Woodhams MD, Hendy MD, Lambert DM. Mutation and evolutionary rates in adélie penguins from the antarctic. PLoS Genet. 2008 b, 4:e1000209

Miller HC, Belov K, Daugherty CH. Characterization of MHC class II genes from an ancient reptile lineage, Sphenodon (tuatara). Immunogenetics. 2005, 57: 883-891

Miller HC, Belov K, Daugherty CH; SMBE Tri-National Young Investigators. Proceedings of the SMBE Tri-National Young Investigators' Workshop 2005. MHC Class I genes in the Tuatara (Sphenodon spp.): evolution of the MHC in an ancient reptilian order. Mol Biol Evol. 2006, 23:949-956

Miller HC, Moore JA, Allendorf FW, Daugherty CH. The evolutionary rate of tuatara revisited. Trends Genet. 2008, 25:13-5

Mindell DP, Knight A, Baer C, Huddleston CJ. Slow rates of molecular evolution in birds and the metabolic rate and body temperature hypotheses. Mol Biol Evol. 1996, 13: 422–426

Miquerol L, Langille BL, Nagy A. Embryonic development is disrupted by modest increases in vascular endothelial growth factor gene expression. Development. 2000, 127:3941-3946

Mooers AO and Harvey P. Metabolic rate, generation time, and the rate of molecular evolution in birds. Mol Phylogenet Evol. 1994, 3:344–350

Morton SU, Scherz PJ, Cordes KR, Ivey KN. Stainier DYR, Srivastava D. microRNA-138 modulates cardiac patterning during embryonic development. Proc. Natl. Acad. Sci. U.S. A. 2008, 105:17830-17835

Moses MJ. Spermatogenesis in crayfish (Procambarus clarkii). II. Description of stages. J Biophys Biochem Cytol. 1961, 10: 301-333

Nabholz B, Glémin S, Galtier N. Strong variations of mitochondrial mutation rate across mammals--the longevity hypothesis. Mol Biol Evol. 2008 a, 25:120-130

Nabholz B, Mauffrey JF, Bazin E, Galtier N, Glemin S. Determination of mitochondrial genetic diversity in mammals. Genetics. 2008 b, 178:351-361

Navascues M, Emerson BC. Elevated substitution rate estimates from ancient DNA: model violation and bias of Bayesian methods. Mol Ecol, 2009, 18:4390–4397

Navascues M, Depaulis F, Emerson BC. Combining contemporary and ancient DNA in population genetic and phylogeographical studies. Molecular Ecology Resources. 2010. 9999

Nei M. Selectionism and neutralism in molecular evolution. Mol Biol Evol. 2005, 22:2318-2342

Nelson NJ, Keall SN, Brown D, Daugherty CH. Establishing a new wild population of tuatara (Sphenodon guntheri). Conserv Biol. 2002, 16:887-894

Newman AK. Notes on the physiology and anatomy of the tuatara (Sphenodon guntheri). Transactions and Proceedings of the New Zealand Institute. 1878, 10: 222–239

Newman DG. Tuatara. Endangered New Zealand Wildlife Series. John McIndoe, Limited. Dunedin, 1987

Newman DG, Watson D, McFadden I. Egg production by tuatara on Lady Alice and Stephens Island. New Zealand Journal of Zoology. 1994, 21: 387–398

Noor MAF, Coyne JA. Genetics of a difference in cuticu- lar hydrocarbons between Drosophila pseudo-obscura and D. persimilis. Genetical Research (Cambridge). 1996, 68: 117–123

Nowick K, Gernat T, Almaas E, Stubbs L. Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. Proc. Natl. Acad. Sci. U.S. A. 2009, 106:22358-63

Nunn GB and Stanley SE. Body size effects and rates of cytochrome b evolution in tube-nosed seabirds. Mol. Biol. Evol. 1998, 15:1360-1371

Ohno S. Argument for genetic simplicity of man and other mammals. Journal of Human Evolution. 1972, 1:651

Ohta T. Population size and the rate of evolution. J. Mol. Evol. 1972, 1:305-314

Ohta T. Slightly deleterious mutant substitutions in evolution. Nature. 1973, 246:96–98 (doi:10.1038/246096a0)

Ohta T. An examination of the generation-time effect on molecular evolution. Proc. Natl. Acad. Sci. U.S. A. 1993, 90:10676–10680

Ohta T. The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst. 1992, 23:263-286

Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981, 290:470-474

Olivo PD, Van deWalle MJ, Laipis PJ, Hauswirth WW. Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial D-loop. Nature 1983, 306:400–402

Omland KE. 1994. Character congruence between a molecular and a morphological phylogeny for dabbling ducks (Anas). Syst Biol. 1994, 43:369-386

Omland KE. Correlated rates of molecular and morphological evolution. Evolution. 1997, 51:1381–1393

Osawa S, Jukes TH, Watanabe K, Muto A. Recent evidence for evolution of the genetic code. Microbiol Rev. 1992, 56:229-264, Review

Päckert M, Martens J, Tietze DT, Dietzen C, Wink M, Kvist L. Calibration of a molecular clock in tits (Paridae)--do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? Mol Phylogenet Evol. 2007, 44:1-14

Pagel M, Venditti C, Meade A. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. Science. 2006, 314:119–121

Pakendorf B and Stoneking M. Mitochondrial DNA and human evolution. Annu Rev Genomics Hum Genet. 2005, 6:165-183

Pal C, Papp B, Lercher MJ. An integrated view of protein evolution. Nat. Rev. Genet. 2006, 7:337–348 (doi:10.1038/nrg1838)

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM. A high observed substitution

rate in the human mitochondrial DNA control region. Nat Genet. 1997, 15:363-368

Peterson AT. Application of molecular clocks in ornithology revisited. J Avian Biol. 2006, 37: 541-544

Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D. Genetic evidence for complex speciation of humans and chimpanzees. Nature. 2006, 441: 1103–1108

Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia a molecular disease. Science. 1949, 110:543-548

Paxinos EE, James HF, Olson SL, Sorenson MD, Jackson J, Fleischer RC. mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (Branta canadensis). Proc. Natl. Acad. Sci. U.S. A. 2002, 99:1399-1404

Penny D. Relativity for molecular clocks. Nature. 2005, 436:183-184

Pereira SL and Baker AJ. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol Biol Evol. 2006 a, 23:1731–1740

Pereira SL and Baker AJ. A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. Mol Phylogenet Evol. 2006 b, 38:499-509

Pikó L and Matsumoto L. Number of mitochondria and some properties of mitochondrial DNA in the mouse egg. Dev Biol. 1976, 49:1-10

Polly PD. On morphological clocks and paleophylogeography: towards a timescale for Sorex hybrid zones. Genetica. 2001, 112–113:339–357

Polly PD. Paleophylogeography: the tempo of geographic differentiation in marmots (Marmota). Journal of Mammalogy. 2003, 84:369–384

Posada D and Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics. 1998, 14: 817-818

Poulton J and Marchington DR. Segregation of mitochondrial DNA (mtDNA) in human

oocytes and in animal models of mtDNA disease: clinical implications. Reproduction 2002, 123:751–755

Prager EM and Wilson AC. Slow evolutionary loss of the potential for intersspecific hybridization in birds. Proc. Natl. Acad. Sci. U.S. A. 1975 72: 200–204

Pulquerio MJ and Nicholas RA. Dates from the molecular clock: how wrong can we be? Trends Ecol Evol. 2007, 22:180-184

Pyburn WF, Kennedy JP. Hybridisation in US treefrogs of the genus Hyla. Proceedings of the Biological Society of Wash. 1961, 74: 157–160

Quinn TW, Shields GF, Wilson AC. Affinities of the Hawaiian Goose Based on Two Types of Mitochondrial DNA Data. The Auk. 1991. 108: 585-593

Quinn TW. The genetic legacy of Mother Goose--phylogeographic patterns of lesser snow goose Chen caerulescens caerulescens maternal lineages. Mol Ecol. 1992, 1:105-117

Rambaut A and Bromham L. Estimating divergence dates from molecular sequences. Mol Biol Evol. 1998, 15:442–448

Rambaut A and Drummond AJ. Tracer. University of Oxford, 2004

Randi E. A mitochondrial cytochrome B phylogeny of the Alectoris partridges. Mol Phylogenet Evol. 1996, 6:214 -227

Redd AJ and Stoneking M. Peopling of Sahul: mtDNA variation in aboriginal Australian and Papua New Guinean populations. Am J Hum Genet.1999, 65:808–828

Rensch B. Evolution above the species level. Methuen, London. 1959

Rest JS, Ast JC, Austin CC, Waddell PJ, Tibbetts EA, Hay JM, Mindell DP. Molecular systematics of primary reptilian lineages and the tuatara mitochondrial genome. Molecular Phylogenetics and Evolution. 2003, 29: 289-297

Robin ED and Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol. 1988, 136:507–513

Rohde K. Latitudinal gradients in species diversity and their causes. I. A review of the hypotheses explaining the gradients. Biologisches Zentralblatt. 1978, 97: 393–403

Rohde K. Latitudinal gradients in species-diversity–the search for the primary cause. Oikos. 1992, 65: 514–527

Rohland N and Hofreiter M. Comparison and optimization of ancient DNA extraction. Biotechniques. 2007, 42: 343-352

Römpler H, Dear PH, Krause J, Meyer M, Rohland N, Schöneberg T, Spriggs H, Stiller M, Hofreiter M. Multiplex amplification of ancient DNA. Nat Protoc. 2006, 1: 720-728

Rottenberg H. Longevity and the evolution of the mitochondrial DNA-coded proteins in mammals. Mech Ageing Dev. 2006, 127:748–760

Rottenberg H. Exceptional longevity in songbirds is associated with high rates of evolution of cytochrome b, suggesting selection for reduced generation of free radicals. J Exp Biol. 2007 a, 210:2170–2180

Rottenberg H. Coevolution of exceptional longevity, exceptionally high metabolic rates, and mitochondrial DNA-coded proteins in mammals. Exp. Gerontol. 2007 b, 42:364–373

Saarma U, Ho SY, Pybus OG, Kaljuste M, Tumanov IL, Kojola I, Vorobiev AA, Markov NI, Saveljev AP, Valdmann H, Lyapunova EA, Abramov AV, et al. Mitogenetic structure of brown bears (Ursus arctos L.) in northeastern Europe and a new time frame for the formation of European brown bear lineages. Mol Ecol. 2007.16:401-13

Saitou N and Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987, 4:406–425

Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: A laboratory manual (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989

Sanderson MJ. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol. 1997, 14:1218–1231

Sanderson MJ. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 2002, 19:101–109

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain - terminating inhibitors. Proc. Natl. Acad. Sci. U.S. A. 1977, 74: 5463–5467

Santos C, Montiel R, Sierra B, Bettencourt C, Fernandez E, Alvarez L, Lima M, Abade A, Aluja MP. Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). Mol Biol Evol. 2005, 22: 1490–1505

Sasa MM, Chippendale PT, Johnson NA. Patterns of postzygotic isolation in frogs. Evolution, 1998, 52: 1811–1820

Savard LM, Michaud, Bousquet J. Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S, rRNA, and rbcL gene sequences. Mol Phylogenet Evol. 1993, 2:112-118

Schrijver I, Pique LM, Traynis I, Scharfe C, Sehnert AJ. Mitochondrial DNA analysis by multiplex denaturing high-performance liquid chromatography and selective sequencing in pediatric patients with cardiomyopathy. Genet Med. 2009, 11: 118–126

Schwartz M and Vissing J. Paternal inheritance of mitochondrial DNA. N Engl J Med. 2002, 347:576–800

Schwartz M and Vissing J. New patterns of inheritance in mitochondrial disease. Biochem. Biophys Res Commun. 2003, 310:247–251

Schwartz M and Vissing J. No evidence for paternal inheritance of mtDNA in patients with sporadic mtDNA mutations. J Neurol Sci. 2004, 218:99–101

Shackelton LA, Parrish CR, Truyen U, Holmes EC. High rate of viral evolution associated with the emergence of carnivore parvovirus. Proc. Natl. Acad. Sci. USA. 2005, 102:379-384

Shepherd LD, Millar CD, Ballard G, Ainley DG, Wilson PR, Haynes GD, Baroni C, Lambert DM. Microevolution and mega-icebergs in the Antarctic. Proc. Natl.

Acad. Sci. U.S. A. 2005, 102:16717-16722.

Seutin G, White BN, Boag PT. Preservation of avian blood and tissue samples for DNA analyses. Can J Zool .1991, 69:82–90

Shadel GS and Clayton DA. Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem. 1997, 66:409-435

Shapiro B, Drummond AJ, Rambaut A, Wilson MC. et al. Rise and fall of the Beringian steppe bison. 2004, Science 306:1561–1565

Shields GF and Wilson AC. Calibration of mitochondrial DNA evolution in geese. J Mol Evol. 1987, 24: 212–217

Shuster RC, Rubenstein AJ, Wallace DC. Mitochondrial DNA in anucleate human blood cells. Biochem Biophys Res Commun.1988, 155:1360-1365

Sigurğardóttir S, Helgason A, Gulcher JR, Stefansson K, Donnelly P. The mutation rate in the human mtDNA control region. Am J Hum Genet. 2000, 66:1599-609

Simpson GG. In: Washburn SL, Editor, classification and Human Evolution. Adline Press, New York. 1970, pp: 175

Simpson GG. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.

Smith, AB, Lafay B, Christen R. Comparative variation of morphological and molecular evolution through geologic time: 28S ribosomal RNA versus morphology in echinoderms. Phil Trans R Soc Lond B Biol Sci. 1992, 338:365-382

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. Correcting for purifying selection: an improved human mitochondrial molecular clock. Am J Hum Genet. 2009, 84:740-459

Stevens GC. The Latitudinal Gradient in Geographical Range: How so Many Species Coexist in the Tropics. The American Naturalist. 1989, 133:240-256 doi:10.2307/2462300

Stoneking M, Sherry ST, Redd AJ, Vigilant L. New approaches to dating suggest a recent age for the human mtDNA ancestor. Philos Trans R Soc London B Biol Sci. 1992, 337:167–175

Stoneking M. DNA and recent human evolution. Evol Anthropol.1993, 2:60–73

Stoneking M and Soodyall H. Human evolution and the mitochondrial genome. Curr Opin Genet Dev.1996, 6:731–736

Subramanian S, Denver DR, Millar CD, Heupink T, Aschrafi A, Emslie SD, Baroni C, Lambert DM. High mitogenomic evolutionary rates and time dependency. Trends Genet. 2009a, 25:482-486

Subramanian S, Hay JM, Mohandesan E, Millar CD, Lambert DM. Molecular and morphological evolution in tuatara are decoupled. Trends Genet. 2009b, 25:16-18.

Sucena E, Delon I, Jones I, Payre F, Stern DL. Regulatory evolution of shavenbaby/ovo underlies multiple cases of morphological parallelism. Nature. 2003, 424: 935–938

Swami M. Turning up the heat. Nat Rev Genet. 2009, 10:512-513

Swofford DL. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, MA, 2003

Takezaki N, Rzhetsky A, Nei M. Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol. 1995, 12:823–833

Tambets K, Rootsi S, Kivisild T, Help H, Serk P, Loogväli EL, Tolk HV, Reidla M, Metspalu E, Pliss L, Balanovsky O, Pshenichnov A, Balanovska E, Gubina M, et al. The western and eastern roots of the Saami–the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. Am J Hum Genet. 2004, 74:661–682

Tamura K and Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993, 10: 512-526.

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007, 24:1596-1599

Tang S and Huang T. Characterization of mitochondrial DNA heteroplasmy using a parallel sequencing system. Biotechniques. 2010,48:287-96

Taylor RW, McDonnell MT, Blakely EL, Chinnery PF, Taylor GA, Howell N, Zeviani M, Briem E, Carrara F, Turnbull DM. Genotypes from patients indicate no paternal mitochondrial DNA contribution. Ann Neurol. 2003, 54:521–524

Taylor RW and Turnbull DM. Mitochondrial DNA mutations in human disease. Nat Rev Genet. 2005, 6: 389-402

Thomas RK, Nickerson E, Simons JF, Jänne PA, Tengs T, Yuza Y, Garraway LA, LaFramboise T, Lee JC, Shah K, O'Neill K, Sasaki H, Lindeman N, Wong KK et al. Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. Nat Med. 2006, 12: 852–855

Thompson MB and Daugherty CH. Metabolism of tuatara, Sphenodon punctatus. Comp Bichem Phys. 1998, 119: 519-522

Thorne JL, Kishino H, Painter IS. Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol. 1998, 15:1647–1657

Thorne JL and Kishino H. Divergence time and evolutionary rate estimation with multilocus data. Syst Biol. 2002, 51:689–702

Tilley SG, Verrell PA, Arnold SJ. Correspondencebetween sexual isolation and allozyme differentiation: a test in the salamander Desmognathus ochriphaeus. Proceedings of the National Academy of Sciences of the USA. 1990, 87: 2715–2719

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet. 1993, 53:563–590

Trifunovic A, Wredenberg A, Falkenberg M, Spelbrink JN, Rovio AT, Bruder CE, Bohlooly-Y M, Gidlöf S, Oldfors A. Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature. 2004, 429:417-423

Ursprung H and Schabtach E. Fertilization in tunicates: loss of the paternal mitochondria prior to sperm entry. J Exp Zool. 1965, 159: 379-381

Vasta V, Ng SB, Turner EH, Shendure J, Hahn SH. Next generation sequence analysis for mitochondrial disorders. Genome Med. 2009, 1: 100

Venditti C and Pagel M. Speciation as an active force in promoting genetic evolution. Trends Ecol Evol. 2009, 25:14–20

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. Science. 1991, 253:1503-1507

Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin Chem. 2009, 55: 641–658

Wall JD. Estimating ancestral population sizes and divergence times. Genetics. 2003, 163:395–404

Wallace DG, Maxson LR, Wilson AC. Albumin evolution in frogs: a test of the evolutionary clock hypothesis. Proc. Natl. Acad. Sci. U.S. A. 1971, 68:3127-129

Wallace DG, King MC, Wilson AC. Albumin differences among ranid frogs: taxonomic and phylogenetic implications. Syst Zool. 1973, 22:1–14

Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. Gene. 1999, 238:211–230

Ward RH, Frazier BL, Dew-Jager K, Pääbo S. Extensive mitochondrial diversity within a single Amerindian tribe. Proc. Natl. Acad. Sci. U.S. A. 1991, 88:8720-8724

Waters JM, Dijkstra LH, Wallis GP. Biogeography of a southern hemisphere freshwater fish: how important is marine dispersal? Mol Ecol. 2000, 9:1815–1821

Wayne RK, Van Valkenburgh B, O'Brien SJ. Molecular distance and divergence time in carnivores and primates. Mol Biol Evol. 1991, 8:297-319

Webster AJ, Payne RJ, Pagel M. Molecular phylogenies link rates of evolution and speciation. Science. 2003, 301:478

Wenink PW, Allan J. Baker AJ, Rosner H-U, Tilanus MGJ. Global Mitochondrial DNA Phylogeography of Holarctic Breeding Dunlins (Calidris alpina). Evolution. 1996, 50: 318-330

Wettstein OV. Rhynchocephalia. Kukenthal and Krumbach's Handbuch der Zoologie. 1931, 7:1-28, 215-235

Wettstein O. von. Zoologisches Anzeiger 143. 1943, pp: 45–47

Whitaker AH and Daugherty CH. Research on the Tuatara (Sphenodon punctatus) of Little Barrier Island. Threatened Species Unit, Department of Conservation, Wellington, 1991

Wiens JJ, Graham CH, Moen DS, Smith SA, Reeder TW. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: treefrog trees unearth the roots of high tropical diversity. Am Nat. 2006, 168:579-596

Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, de Bruijn E, Horvitz HR, Kauppinen S, Plasterk RH. MicroRNA expression in zebrafish embryonic development. Science. 2005, 309:310-311

Wiesner RJ, Rüegg JC, Morano I. Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. Biochem Biophys Res Commun. 1992, 183:553-559

Williams GR and Given DR. The Red Data Book of New Zealand. Nature Conservation Council, Wellington, 1981

Wilson AC, Sarich VM, Maxson LR. The importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein, and anatomical evolution. Proc. Natl. Acad. Sci. U.S. A. 1974 a, 71:3028-3030

Wilson AC, Maxson LR, Sarich VM. Two types of molecular evolution: Evidence from studies of interspecific hybridization. Proc. Natl. Acad. Sci. U.S. A. 1974 b, 71:2843-2847

Wilson AC, Carlson SS, White TJ. Biochemical evolution. Annu Rev Biochem. 1977, 46:573-639

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M. Mitochondrial DNA and two perspectives on evolutionary genetics. Biological Journal of the Linnean Society. 1985, 26:375-400

Wittkopp PJ, Vaccaro K, Carroll SB. Evolution of yellow gene regulation and pigmentation in Drosophila. Curr Biol. 2002, 12: 1547– 1556

Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. Int Rev Cytol. 1992,141:173-216, Review

Wright, Kevin DVM. Tuataras. Reptiles magazine. Fancy Publications. Irvine, California. 1994, 2(1)

Wright S, Keeling J, Gillman L. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. Proc. Natl. Acad. Sci. U.S. A. 2006, 103:7718-7722

Wright SD, Gillman LN, Ross HA, Keeling DJ. Slower tempo of microevolution in island birds: implications for conservation biology. Evolution. 2009, 63:2275-2287

Wright SD, Ross HA, Keeling DJ, McBride P, Gillman LN. Thermal Energy and the Rate of Genetic Evolution in Marine Fishes. Evolutionary Ecology. 2011, 25:525-530

Wu CI and Li WH. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. U.S. A. 1985, 82:1741-1745

Wu C-I and Hollocher H. Subtle is nature: the genetics of species differentiation and speciation. In: Howard DJ, Berlocher SH, Editors. Endless forms: species and speciation. New York: Oxford University Press, 1998, 339–351

Yang Z. Estimating the pattern of nucleotide substitution. J Mol Evol. 1994, 39:105-111

Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci.1997, 13: 555-556

Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007, 24:1586-1591

Yang Z and Nielsen R. Mutation–selection models of codon substitution and their use to estimate selective strengths on codon usage. Mol Biol Evol. 2008, 25: 568–579

Yokoyama S. Molecular evolution of color vision in vertebrates. Gene. 2002, 300: 69–78

Zaragoza MV, Fass J, Diegoli M, Lin D, Arbustini E. Mitochondrial DNA variant discovery and evaluation in human Cardiomyopathies through next-generation sequencing. PLoS One. 2010, 5: e12295 doi:10.1371/journal.pone.0012295

Zhang J, Zhang YP, Rosenberg HF. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet. 2002, 30: 411–415

Zhao X, Li N, Guo W, Hu X, Liu Z, Gong G, Wang A, Feng J, Wu C. Further evidence for paternal inheritance of mitochondrial DNA in the sheep (Ovis aries). Heredity. 2004, 93:399-403

Zharkikh A. Estimation of evolutionary distances between nucleotide sequences. J Mol Evol. 1994, 39:315-329

Zink RM. A new perspective on the evolutionary history of Darwin's finches. Auk. 2002, 119: 864⊠871

Zouros E. Genetic differentiation associated with the early stages of speciation in the mulleri subgroup of Drosophila. Evolution. 1973, 27: 601–621

Zouros E. The chromosomal basis of sexual isolation in two sibling species of Drosophila: D. arizonensis and D. mojavensis. Genetics. 1981, 97: 703–718

Zouros E, Freeman KR, Ball AO, Pogson GH. Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel Mytilus. Nature. 1992, 359:412–414

Zuckerkandl E and Pauling L. Molecular disease, evolution, and genetic heterogeneity. In Kasha M and Pullman B, Editors, Horizons in biochemistry. Academic Press, New York, 1962, pp:189-225

Zuckerkandl E, Pauling L. Molecules as documents of evolutionary history. J Theor Biol. 1965, 8:357–368

Appendix A

Comparison of Phenol-Chloroform and Silica-Based DNA Extraction Methods

A.1 Comparison of Various DNA Extraction Methods

It has been suggested that phenol/chloroform DNA extraction typically result in higher DNA yields, in comparison to other methods (Cattaneo *et al.* 1997; Yang *et al.* 1998). In this study, two different methods of DNA extraction were tested and the DNA mass yields were compared, using gel electrophoresis and a nanodrop spectrophotometer. First, DNA was extracted by proteinase K (PK) digestion followed by phenol/chloroform approach (Sambrook *et al.* 1989) and second, a Silica-based method (Boom *et al.* 1990; Höss & Pääbo 1993) were investigated. Using these two methods, eight different tuatara blood samples were extracted in duplicate. Although no significant difference in DNA yield were recorded for these two methods, (Table A.1) (Figure A.1), the phenol/chloroform method was chosen for DNA extraction from blood samples since it was quicker and less cost effective compared to the silica method. This was important for DNA extraction of a large number of samples.

Table A.1: Comparison of silica and phenol/chloroform-based DNA extraction methods

Sample ID	Locations	DNA Yield (ng/µl)	
		Phenol /chloroform	**Silica**
FT200	North Brother	767.9	750.2
FT2409	Hen	194.2	200.6
FT2391	Green Mercury	376.4	370
FT2365	Middle Mercury	105	120.7
FT2704	Stanley	300.3	320.5
CD1333	Stephens	482.6	490
FT4926		17.6	17.2
FT2522	Tawhiti Rahi	93	104.5
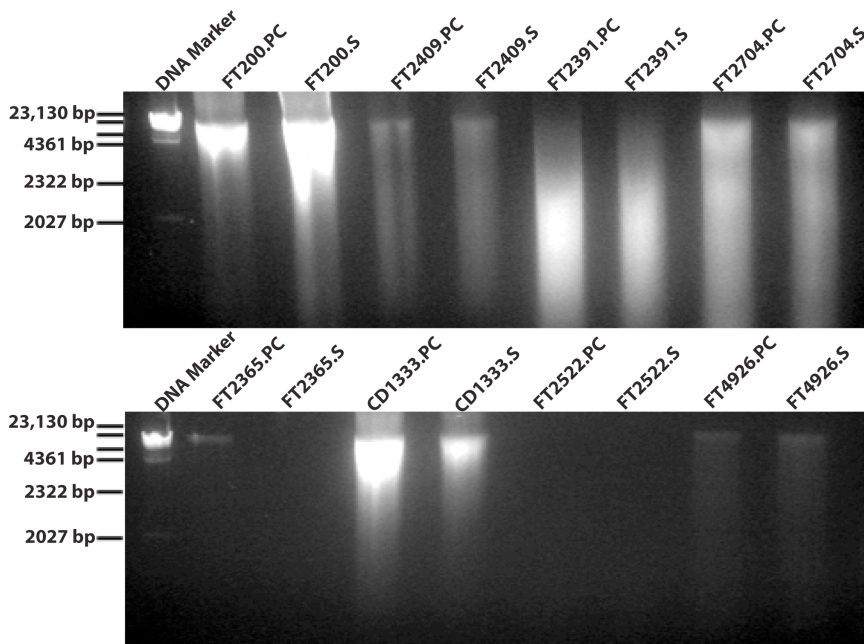
Figure A.1: Comparison of phenol/chloroform (PC) and silica (S) based DNA extraction methods. Using phenol/chloroform and silica methods, DNA were extracted from eight tuatara blood samples in duplicate. DNA was subjected to electrophoresis in 1.5% agarose and visualized by ethidium bromide staining over ultra violet (UV) radiation. The results show that there is no significant different between these two methods in relation to DNA yield.

Appendix B

**Supplementary Material for Chapter Seven:
Potential Heteroplamies in Tuatara mtDNA**

Table B.1: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4241, mapped to NCBI Reference (NC_004815). The variants are single nucleotide substitutions or indels (insertions or deletions). The total depth is the average number of mapped sequences (non-redundant and redundant sequences) at each base position. The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4241

Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	102	102	A	G	100	47	100	100	25	22	25	22
NC_004815	1571	1571	G	T	25.8	31	18.2	30	2	6	11	20
NC_004815	2229	2229	A	C	100	10	100	100	4	6	4	6
NC_004815	4092	4092		-	15.2	46	0	15.6	0	7	1	45
NC_004815	5021	5021	T	C	23.4	47	20.8	26.1	5	6	24	23
NC_004815	5671	5671	T	CC	100	41	100	100	3	38	3	38
NC_004815	7605	7605	A	G	98.4	64	92.9	100	13	50	14	50
NC_004815	11478	11478	T	C	100	49	100	100	10	39	10	39
NC_004815	13034	13034	-	AACGCACACCTATAAAAA	100	20	-	100	0	20	0	20
NC_004815	13064	13064	-	A	100	28	100	100	3	25	3	25
NC_004815	13092	13092	T	A	100	24	100	100	3	21	3	21
NC_004815	13108	13108	-	A	100	25	100	100	4	21	4	21
NC_004815	14351	14351	-	A	100	45	100	100	8	37	8	37
NC_004815	14356	14356	T	-	100	47	100	100	10	37	10	37
NC_004815	14442	14442	-	C	100	34	100	100	10	24	10	24
NC_004815	14515	14515	A	G	100	39	100	100	11	28	11	28
NC_004815	14565	14565	-	A	100	46	100	100	19	27	19	27
NC_004815	14592	14592	-	A	100	56	100	100	27	29	27	29

Table B.2: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4244, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4244												
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	102	102	A	G	100	73	100	100	38	35	38	35
NC_004815	2229	2229	A	C	100	30	100	100	10	20	10	20
NC_004815	4127	4128	CA	TG	12.9	62	0	16.3	0	8	13	49
NC_004815	4757	4757	T	CTA	13.3	30	25	11.5	1	3	4	26
NC_004815	5671	5671	T	C	100	45	100	100	5	40	5	40
NC_004815	7605	7605	A	G	100	86	100	100	20	66	20	66
NC_004815	11478	11478	T	C	100	47	100	100	12	35	12	35
NC_004815	13034	13034	-	AACGCACACCTATAAAAA	100	25	-	100	0	25	0	25
NC_004815	13064	13064	-	A	100	38	100	100	5	33	5	33
NC_004815	13092	13092	T	A	100	35	100	100	5	30	5	30
NC_004815	13108	13108	-	A	97.6	41	100	96.9	9	31	9	32
NC_004815	14351	14351	-	A	100	44	100	100	17	27	17	27
NC_004815	14356	14356	T	-	100	47	100	100	19	28	19	28
NC_004815	14442	14442	-	C	100	42	100	100	17	25	17	25
NC_004815	14515	14515	A	G	97.1	35	88.9	100	8	26	9	26
NC_004815	14565	14565	-	A	100	45	100	100	18	27	18	27
NC_004815	14592	14592	-	A	100	46	100	100	22	24	22	24

Table B.3: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4252, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4252												
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	102	102	A	G	100	64	100	100	38	26	38	26
NC_004815	2229	2229	A	C	100	16	100	100	9	7	9	7
NC_004815	5671	5671	T	C	100	19	100	100	4	15	4	15
NC_004815	7605	7605	A	G	98.3	60	100	98	11	48	11	49
NC_004815	8179	8179	A	G	100	37	100	100	14	23	14	23
NC_004815	13034	13034	-	AACGCACACCTATAAAAA	100	24	-	100	0	24	0	24
NC_004815	13064	13064	-	A	97.1	35	100	96.6	6	28	6	29
NC_004815	13092	13092	T	A	93.1	29	85.7	95.5	6	21	7	22
NC_004815	13108	13108	-	A	100	30	100	100	8	22	8	22
NC_004815	14351	14351	-	A	100	35	100	100	13	22	13	22
NC_004815	14356	14356	T	-	100	38	100	100	14	24	14	24
NC_004815	14442	14442	-	C	100	36	100	100	12	24	12	24
NC_004815	14515	14515	A	G	100	31	100	100	6	25	6	25
NC_004815	14565	14565	-	A	100	40	100	100	12	28	12	28
NC_004815	14592	14592	-	A	100	44	100	100	16	28	16	28

Table B.4: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4251, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4251													
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	102	102	A	G	100	67	100	100	37	30	37	30
NC_004815	2229	2229	A	C	100	12	100	100	5	7	5	7
NC_004815	5671	5671	T	C	100	46	100	100	9	37	9	37
NC_004815	7605	7605	A	G	98.6	72	95.2	100	20	51	21	51
NC_004815	8179	8179	A	G	100	48	100	100	15	33	15	33
NC_004815	9197	9197	T	A	13.5	37	6.7	18.2	1	4	15	22
NC_004815	13064	13064	-	A	100	28	100	100	2	26	2	26
NC_004815	13092	13092	T	A	100	29	100	100	4	25	4	25
NC_004815	13108	13108	-	A	100	32	100	100	5	27	5	27
NC_004815	14351	14351	-	A	100	34	100	100	9	25	9	25
NC_004815	14356	14356	T	-	100	34	100	100	9	25	9	25
NC_004815	14442	14442	-	C	100	43	100	100	13	30	13	30
NC_004815	14515	14515	A	G	100	35	100	100	9	26	9	26
NC_004815	14565	14565	-	A	100	49	100	100	18	31	18	31
NC_004815	14592	14592	-	A	100	52	100	100	24	28	24	28

Table B.5: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4250, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4250												
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	87	87	C	T	98.6	73	100	97.2	37	35	37	36
NC_004815	102	102	A	G	98.7	75	100	97	42	32	42	33
NC_004815	275	275	C	T	100	43	100	100	26	17	26	17
NC_004815	701	701	G	A	100	24	100	100	13	11	13	11
NC_004815	944	944	A	G	100	39	100	100	10	29	10	29
NC_004815	1334	1334	T	C	100	38	100	100	11	27	11	27
NC_004815	1514	1514	C	A	100	25	100	100	17	8	17	8
NC_004815	1706	1706	A	G	100	22	100	100	6	16	6	16
NC_004815	1899	1899	A	G	100	24	100	100	12	12	12	12
NC_004815	2063	2063	C	T	100	31	100	100	7	24	7	24
NC_004815	2229	2229	A	C	100	12	100	100	4	8	4	8
NC_004815	2340	2340	T	C	100	26	100	100	3	23	3	23
NC_004815	2356	2356	-	G	100	25	100	100	5	20	5	20
NC_004815	2420	2420	C	T	96.8	31	100	96	6	24	6	25
NC_004815	2548	2548	A	G	100	30	100	100	17	13	17	13
NC_004815	2607	2607	G	A	100	28	100	100	19	9	19	9
NC_004815	2731	2731	A	C	100	23	100	100	12	11	12	11
NC_004815	2965	2965	G	A	100	29	100	100	10	19	10	19
NC_004815	3607	3607	A	G	100	17	100	100	10	7	10	7
NC_004815	3916	3916	C	T	91.9	37	93.3	90.9	14	20	15	22
NC_004815	4114	4114	T	C	76.9	39	100	74.3	4	26	4	35
NC_004815	4213	4213	G	A	100	50	100	100	21	29	21	29
NC_004815	4485	4485	G	A	100	35	100	100	12	23	12	23
NC_004815	4713	4713	C	G	100	13	100	100	2	11	2	11
NC_004815	5018	5018	G	A	100	29	100	100	11	18	11	18

NC_004815	5021	5021	T	CC	100	29	100	100	11	18	11	18
NC_004815	5671	5671	T	C	100	15	100	100	2	13	2	13
NC_004815	6141	6141	A	G	100	28	100	100	12	16	12	16
NC_004815	6402	6402	C	T	100	33	100	100	21	12	21	12
NC_004815	7325	7325	A	G	91.1	45	100	78.9	26	15	26	19
NC_004815	7413	7413	C	T	100	35	100	100	14	21	14	21
NC_004815	7605	7605	A	G	96.1	102	84.6	100	22	76	26	76
NC_004815	7690	7690	A	T	96.9	98	96.6	97.1	28	67	29	69
NC_004815	7903	7903		-	44	125	32	52	16	39	50	75
NC_004815	7909	7909	-	C	100	130	100	100	50	80	50	80
NC_004815	8105	8105	G	A	71	62	71.8	69.6	28	16	39	23
NC_004815	8269	8269	A	C	74.1	58	100	55.9	24	19	24	34
NC_004815	8596	8596	T	C	100	71	100	100	21	50	21	50
NC_004815	9205	9205	G	A	98.4	64	93.3	100	14	49	15	49
NC_004815	9250	9250	C	A	100	57	100	100	23	34	23	34
NC_004815	9298	9298	A	G	80.6	31	100	78.6	3	22	3	28
NC_004815	9536	9536	C	T	100	13	100	100	1	12	1	12
NC_004815	10373	10373	A	T	100	8	100	100	5	3	5	3
NC_004815	10402	10405	ATTC	CAAT	100	3	100	100	2	1	2	1
NC_004815	10408	10410	AAA	GT	100	3	100	100	2	1	2	1
NC_004815	11259	11264	AGGGGT	GGGGTA	84.6	26	100	81	5	17	5	21
NC_004815	11265	11265	G	-	73.1	26	0	90.5	0	19	5	21
NC_004815	11732	11732	C	T	100	34	100	100	17	17	17	17
NC_004815	12414	12414	C	T	100	10	100	100	7	3	7	3
NC_004815	12528	12528	C	-	100	24	100	100	6	18	6	18
NC_004815	12587	12587	G	A	96.2	26	100	95.5	4	21	4	22
NC_004815	12709	12709	G	A	100	61	100	100	13	48	13	48
NC_004815	12860	12860	T	C	100	46	100	100	15	31	15	31
NC_004815	12890	12890	T	C	100	48	100	100	19	29	19	29
NC_004815	13030	13030	-	AAAAAACGCACACCTATA	84.2	19	0	88.9	0	16	1	18
NC_004815	13059	13059	C	A	100	32	100	100	2	30	2	30
NC_004815	13064	13064	-	A	100	32	100	100	2	30	2	30
NC_004815	13092	13092	T	A	100	30	100	100	3	27	3	27
NC_004815	13108	13108	-	A	100	28	100	100	5	23	5	23

NC_004815	13320	13320	C	T	100	48	100	100	23	25	23	25
NC_004815	14351	14351	-	A	100	35	100	100	10	25	10	25
NC_004815	14356	14356	T	-	100	35	100	100	10	25	10	25
NC_004815	14442	14442	-	C	100	44	100	100	20	24	20	24
NC_004815	14515	14515	A	G	100	34	100	100	12	22	12	22
NC_004815	14565	14565	-	A	100	48	100	100	19	29	19	29
NC_004815	14592	14592	-	A	100	49	100	100	25	24	25	24

Table B.6: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4249, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4249												
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	102	102	A	G	100	61	100	100	26	35	26	35
NC_004815	2229	2229	A	C	100	11	100	100	4	7	4	7
NC_004815	4127	4127	C	T	15.2	46	0	17.5	0	7	6	40
NC_004815	4128	4128	A	G	15.2	46	0	17.5	0	7	6	40
NC_004815	5671	5671	T	CC	100	26	100	100	1	25	1	25
NC_004815	7605	7605	A	G	100	68	100	100	11	57	11	57
NC_004815	7841	7841	G	A	20	45	21.1	19.2	4	5	19	26
NC_004815	8030	8030	A	G	11.7	60	14.8	9.1	4	3	27	33
NC_004815	13034	13034	-	AACGCACACCTATAAAAA	100	17	-	100	0	17	0	17
NC_004815	13064	13064	-	A	100	30	100	100	6	24	6	24
NC_004815	13092	13092	T	A	74.2	31	50	77.8	2	21	4	27
NC_004815	13108	13108	-	A	100	31	100	100	3	28	3	28
NC_004815	14351	14351	-	A	100	33	100	100	7	26	7	26
NC_004815	14356	14356	T	-	100	35	100	100	8	27	8	27
NC_004815	14442	14442	-	C	100	38	100	100	9	29	9	29
NC_004815	14515	14515	A	G	100	34	100	100	13	21	13	21
NC_004815	14565	14565	-	A	100	51	100	100	19	32	19	32
NC_004815	14592	14592	-	A	98.1	54	100	96.6	25	28	25	29

Table B.7: Summary of potential heteroplasmic mtDNA variants recorded in complete mitochondrial genomes of FT4246, mapped to NCBI Reference (NC_004815). The potential heteroplasmic substitution variants are colored in yellow.

Sample ID: FT4246												
Reference Accession Number	Start Position in Ref	End Position in Ref	Reference Bases	Variation Bases	Total Variation Percent	Total Depth	Percent Forward With Variation	Percent Reverse With Variation	Num Forward With Variation	Num Reverse With Variation	Total Num Forward Reads	Total Num Reverse Reads
NC_004815	87	87	C	T	100	85	100	100	36	49	36	49
NC_004815	101	101	-	G	98.9	89	97.7	100	42	46	43	46
NC_004815	275	275	C	T	100	47	100	100	24	23	24	23
NC_004815	701	701	G	A	100	22	100	100	15	7	15	7
NC_004815	944	944	A	G	100	34	100	100	8	26	8	26
NC_004815	1334	1334	T	C	100	40	100	100	10	30	10	30
NC_004815	1514	1514	C	A	100	24	100	100	7	17	7	17
NC_004815	1706	1706	A	G	100	19	100	100	4	15	4	15
NC_004815	1899	1899	A	G	100	14	100	100	6	8	6	8
NC_004815	2063	2063	C	T	95.2	21	100	94.4	3	17	3	18
NC_004815	2229	2229	A	C	100	13	100	100	5	8	5	8
NC_004815	2340	2340	T	C	100	12	100	100	2	10	2	10
NC_004815	2356	2356	-	G	100	13	100	100	2	11	2	11
NC_004815	2420	2420	C	T	100	37	100	100	6	31	6	31
NC_004815	2548	2548	A	G	100	39	100	100	13	26	13	26
NC_004815	2607	2607	G	A	97	33	100	95.7	10	22	10	23
NC_004815	2731	2731	A	C	100	25	100	100	12	13	12	13
NC_004815	2965	2965	G	A	100	40	100	100	11	29	11	29
NC_004815	3607	3607	A	G	100	16	100	100	8	8	8	8
NC_004815	3916	3916	C	T	90	40	100	85.2	13	23	13	27
NC_004815	4114	4114	T	C	80.3	61	100	77.4	8	41	8	53
NC_004815	4125	4128	CACA	ATG	16.1	56	0	19.1	0	9	9	47
NC_004815	4213	4213	G	A	100	51	100	100	17	34	17	34
NC_004815	4485	4485	G	A	100	47	100	100	11	36	11	36
NC_004815	4713	4713	C	G	91.3	23	66.7	100	4	17	6	17

NC_004815	5018	5018	G	A	95.7	46	93.8	96.7	15	29	16	30
NC_004815	5021	5021	T	CC	100	45	100	100	17	28	17	28
NC_004815	5671	5671	T	CC	100	43	100	100	5	38	5	38
NC_004815	6141	6141	A	G	100	22	100	100	13	9	13	9
NC_004815	6402	6402	C	T	97.4	39	95.8	100	23	15	24	15
NC_004815	7325	7325	A	-	100	54	100	100	24	30	24	30
NC_004815	7413	7413	C	T	100	35	100	100	13	22	13	22
NC_004815	7605	7605	A	G	98.9	95	95	100	19	75	20	75
NC_004815	7690	7690	A	T	100	82	100	100	16	66	16	66
NC_004815	7909	7909	-	C	100	90	100	100	32	58	32	58
NC_004815	8105	8105	G	A	81.7	60	84.4	78.6	27	22	32	28
NC_004815	8269	8269	A	C	97.7	43	100	95.7	20	22	20	23
NC_004815	8596	8596	T	C	98.7	79	100	98	28	50	28	51
NC_004815	8689	8689	A	C	68.5	73	66.7	69.2	14	36	21	52
NC_004815	9205	9205	G	A	100	53	100	100	17	36	17	36
NC_004815	9250	9250	C	A	93.5	46	82.4	100	14	29	17	29
NC_004815	9298	9298	A	G	100	21	100	100	1	20	1	20
NC_004815	10373	10373	A	T	80	5	100	66.7	2	2	2	3
NC_004815	11265	11265	G	A	88.9	27	0	100	0	24	3	24
NC_004815	11732	11732	C	T	100	46	100	100	12	34	12	34
NC_004815	12414	12414	C	T	100	13	100	100	8	5	8	5
NC_004815	12526	12526	C	G	100	34	100	100	5	29	5	29
NC_004815	12587	12587	G	A	93.2	44	80	94.9	4	37	5	39
NC_004815	12709	12709	G	A	98.9	89	93.8	100	15	73	16	73
NC_004815	12860	12860	T	C	98.2	57	93.8	100	15	41	16	41
NC_004815	12890	12890	T	C	100	57	100	100	19	38	19	38
NC_004815	13030	13030	-	AAAAAACGCACACCTATA	95.7	23	0	100	0	22	1	22
NC_004815	13059	13059	C	A	100	41	100	100	3	38	3	38
NC_004815	13064	13064	-	A	100	41	100	100	3	38	3	38
NC_004815	13092	13092	T	A	100	40	100	100	7	33	7	33
NC_004815	13108	13108	-	A	97.7	43	100	97	10	32	10	33
NC_004815	13320	13320	C	T	100	65	100	100	26	39	26	39
NC_004815	14351	14351	-	A	100	46	100	100	11	35	11	35
NC_004815	14356	14356	T	-	100	49	100	100	11	38	11	38

NC_004815	14442	14442	-	C	100	52	100	100	11	41	11	41
NC_004815	14515	14515	A	G	100	46	100	100	13	33	13	33
NC_004815	14565	14565	-	A	98.6	69	100	97.7	26	42	26	43
NC_004815	14592	14592	-	A	98.7	75	100	97.4	36	38	36	39

Appendix C

Radiocarbon Dating of Sub-fossil Bones

RAFTER RADIOCARBON LABORATORY

R29386/15

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
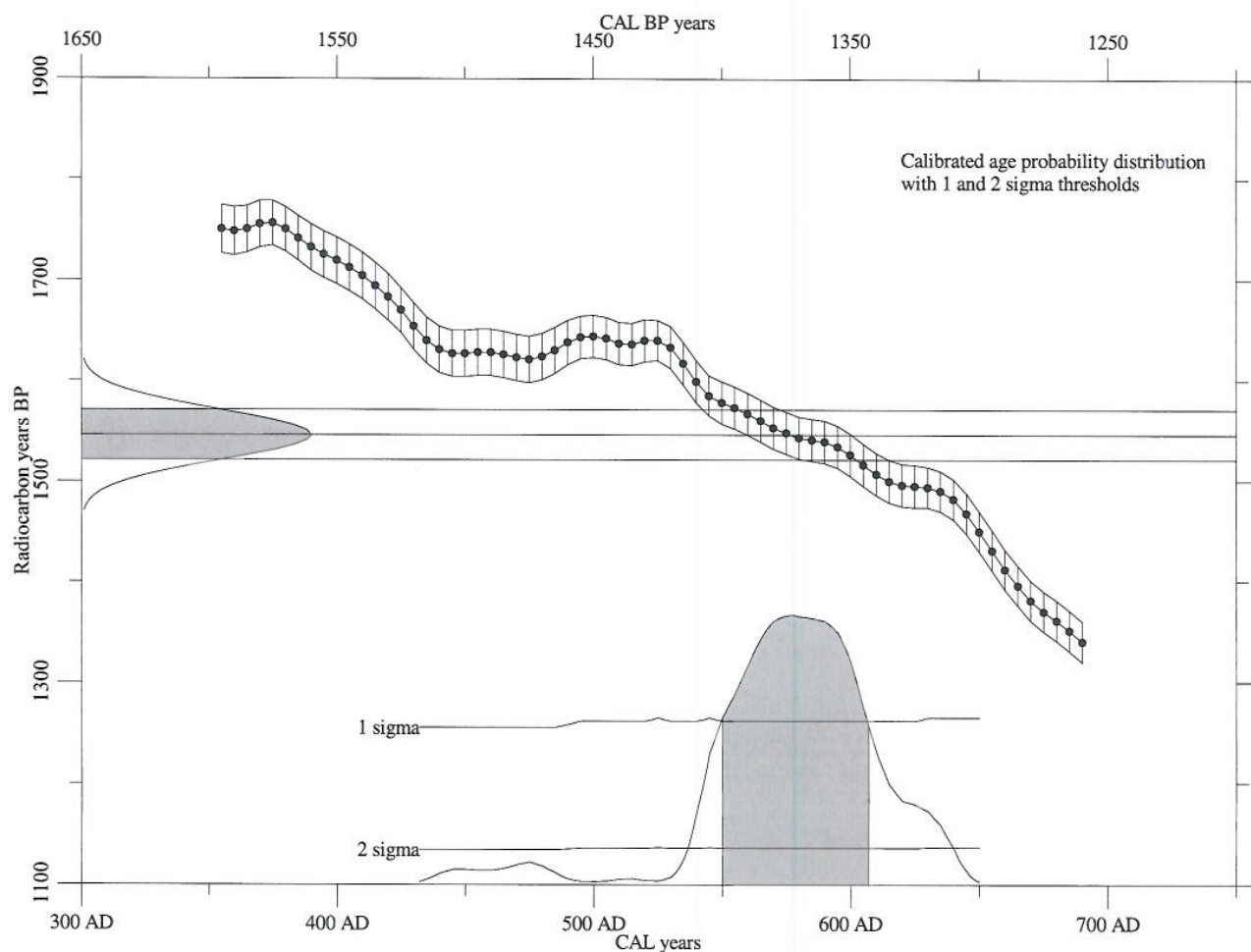Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26697 CONVENTIONAL RADIOCARBON AGE   1546 ± 25 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 537 AD to 641 AD	1413 BP to 1309 BP (93.3% of area)
1 sigma interval is 550 AD to 607 AD	1400 BP to 1343 BP (70.9% of area)

RAFTER
GNS SCIENCE

NZA	**26697**
R	29386/15
Job No	37701
Measured	15-Nov-06
Issued	21-Nov-06
	Page 3 of 6

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T102 MNZ S34098 max
Description	Bone
Fraction Dated	bone gel
Submitter	Jennifer Hay Massey University

* **Radiocarbon Age**   **1546 ± 25 BP** $\delta^{13}C =$ -20.2 ‰

** **Per cent modern =**   81.93 ± 0.28 $\delta^{14}C =$ -172.6 ± 2.8 ‰ $\Delta^{14}C =$ -180.7 ± 2.8 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI)
   corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample: tuatara bone , Knobby Range, Alexandra - hard, well preserved. Microscopic exam
revealed very clean, hard, very well preserved bone. Scraped overall with scalpel. Ground to 1cm
and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

RAFTER RADIOCARBON LABORATORY

R29386/24

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.

PO Box 31312, Lower Hutt, New Zealand
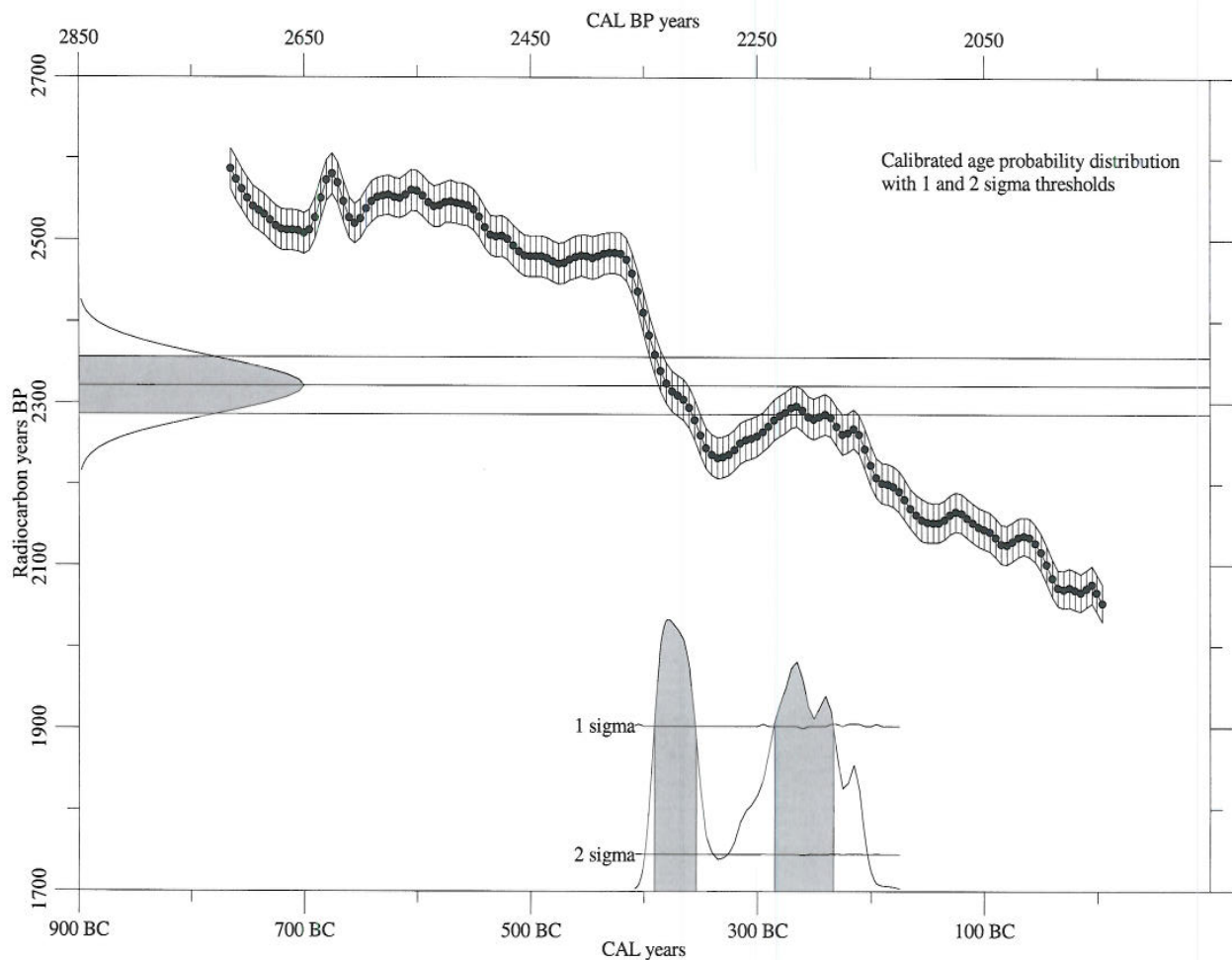Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26797 CONVENTIONAL RADIOCARBON AGE 2321 ± 35 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 399 BC to 202 BC	2348 BP to 2151 BP (97.1% of area)
1 sigma interval is 391 BC to 354 BC	2340 BP to 2303 BP (30.8% of area)
plus 285 BC to 233 BC	2234 BP to 2182 BP (35.3% of area)

Calibrated age probability distribution with 1 and 2 sigma thresholds

RAFTER
GNS SCIENCE

NZA	**26797**
R	29386/24
Job No	37710
Measured	13-Dec-06
TW No	2149
Issued	18-Dec-06
	Page 6 of 6

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T151 SM6 Rd
Description	Bone
Fraction Dated	bone gel
Submitter	Jennifer Hay Massey University

---

* **Radiocarbon Age** 2321 ± 35 **BP** $\delta^{13}C = $ -20.3 ‰

** **Per cent modern** = 74.4 ± 0.33 $\delta^{14}C = $ -248.8 ± 3.3 ‰ $\Delta^{14}C = $ -256 ± 3.3 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI)
    corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample: tuatara bone , Wakapaki Coastal area - hard, well preserved. Ink marking from museum archiving. Microscopic exam revealed very clean, hard, very well preserved bone. Scraped overall with scalpel. Ground to 1cm and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

The reported errors comprise statistical errors in sample and standard determinations, combined in quadrature with a system error component based on the analysis of an ongoing series of measurements on an oxalic acid standard.
For the present result the system error component is conservatively estimated as 0.18% (= ± 14 radiocarbon years).

National Isotope Centre, Institute of Geological and Nuclear Sciences Ltd (GNS Science)
PO Box 31-312 Lower Hutt, New Zealand   Fax +64 4 570 4657 Phone +64 4 570 4644
www.RafterRadiocarbon.co.nz

RAFTER RADIOCARBON LABORATORY

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
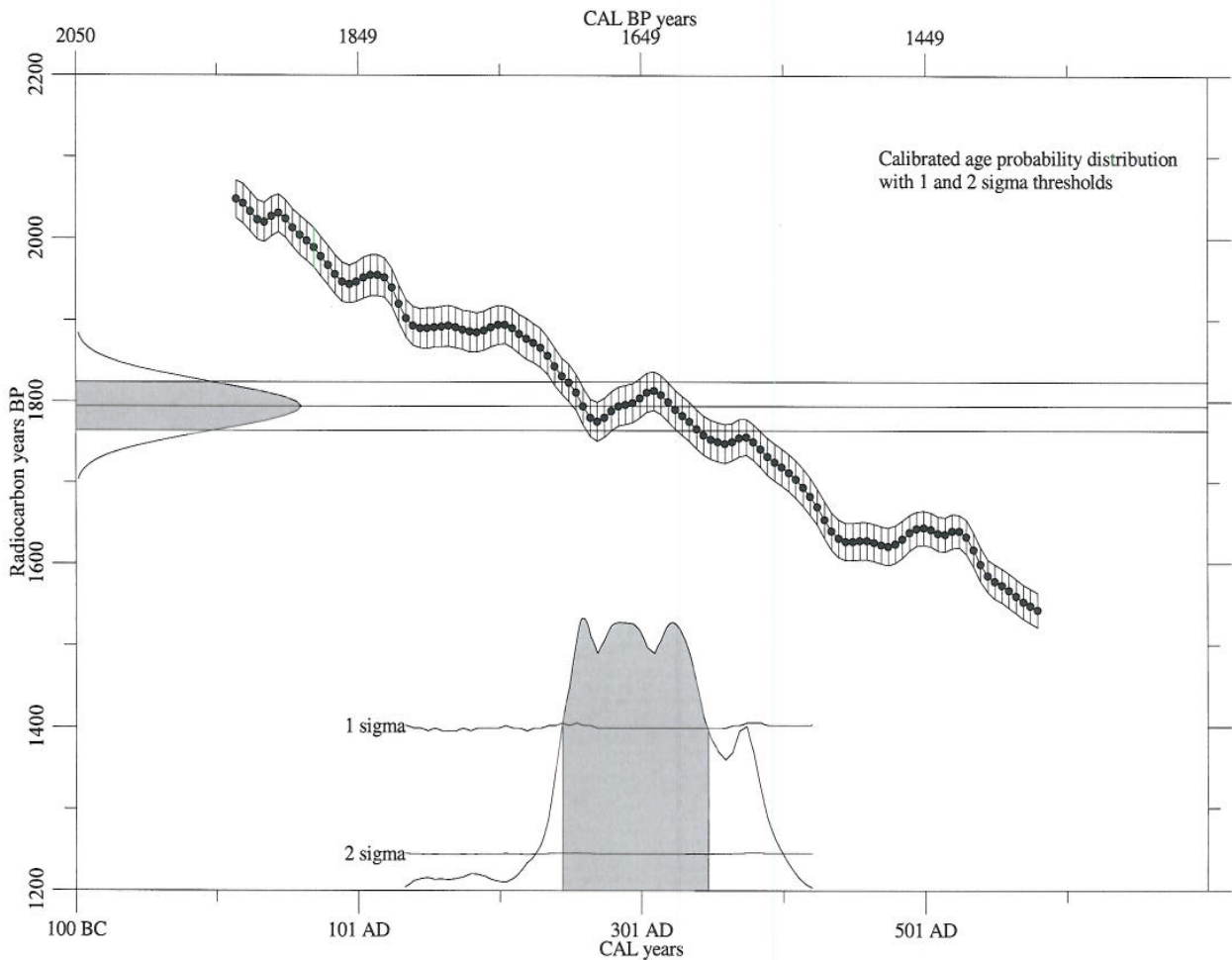Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26685 CONVENTIONAL RADIOCARBON AGE   1794 ± 30 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 226 AD to 401 AD	1724 BP to 1549 BP (95.4% of area)
1 sigma interval is 245 AD to 348 AD	1705 BP to 1602 BP (72.6% of area)

NZA   **26685**

R   29386/17
Job No   37703
Measured   15-Nov-06
Issued   21-Nov-06

Page 5 of 6

RAFTER
GNS SCIENCE

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T110 MNZ S33366 Ltib
Description	Bone
Fraction Dated	bone gel
Submitter	Jennifer Hay Massey University

---

* **Radiocarbon Age**   **1794 ± 30 BP** $\delta^{13}C =$ -19.4 ‰

** **Per cent modern =**   79.44 ± 0.29 $\delta^{14}C =$ -196.4 ± 2.9 ‰ $\Delta^{14}C =$ -205.6 ± 2.9 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI)
   corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample: tuatara bone , Holocene Cave, Mt Cook, Canterbury- hard, well preserved. Microscopic
exam revealed very clean, hard, very well preserved bone. Scraped overall with scalpel. Ground to
1cm and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

RAFTER RADIOCARBON LABORATORY

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
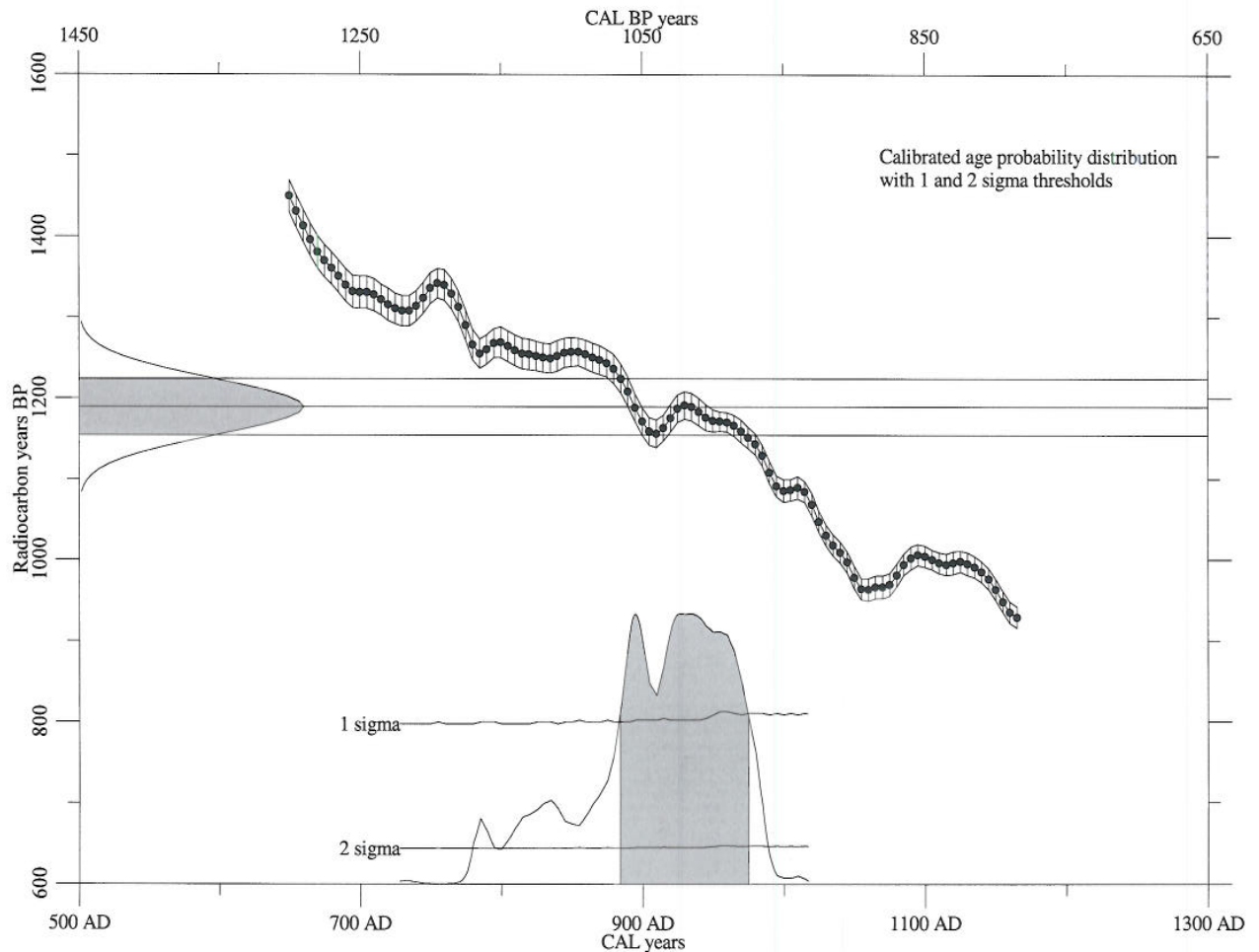Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26793   CONVENTIONAL RADIOCARBON AGE   1189 ± 35 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 780 AD to 989 AD	1170 BP to 961 BP (97.7% of area)
1 sigma interval is 884 AD to 975 AD	1066 BP to 975 BP (70.0% of area)

Calibrated age probability distribution with 1 and 2 sigma thresholds

NZA   **26793**

R	29386/22
Job No	37708
Measured	13-Dec-06
TW No	2149
Issued	18-Dec-06
	Page 4 of 6

RAFTER
GNS SCIENCE

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T142 CM REP 335 vert 2+3
Description	Bone
Fraction Dated	bone gel
Submitter	Jennifer Hay Massey University

---

* Radiocarbon Age   **1189 ± 35 BP** $\delta^{13}C =$ -20 ‰

** Per cent modern =   85.66 ± 0.36 $\delta^{14}C =$ -134.6 ± 3.6 ‰ $\Delta^{14}C =$ -143.4 ± 3.6 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI) corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample: tuatara bone , Marfell's Beach, Marlbourgh- hard, well preserved. Microscopic exam revealed very clean, hard, very well preserved bone. Scraped overall with scalpel. Ground to 1cm and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

RAFTER RADIOCARBON LABORATORY

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
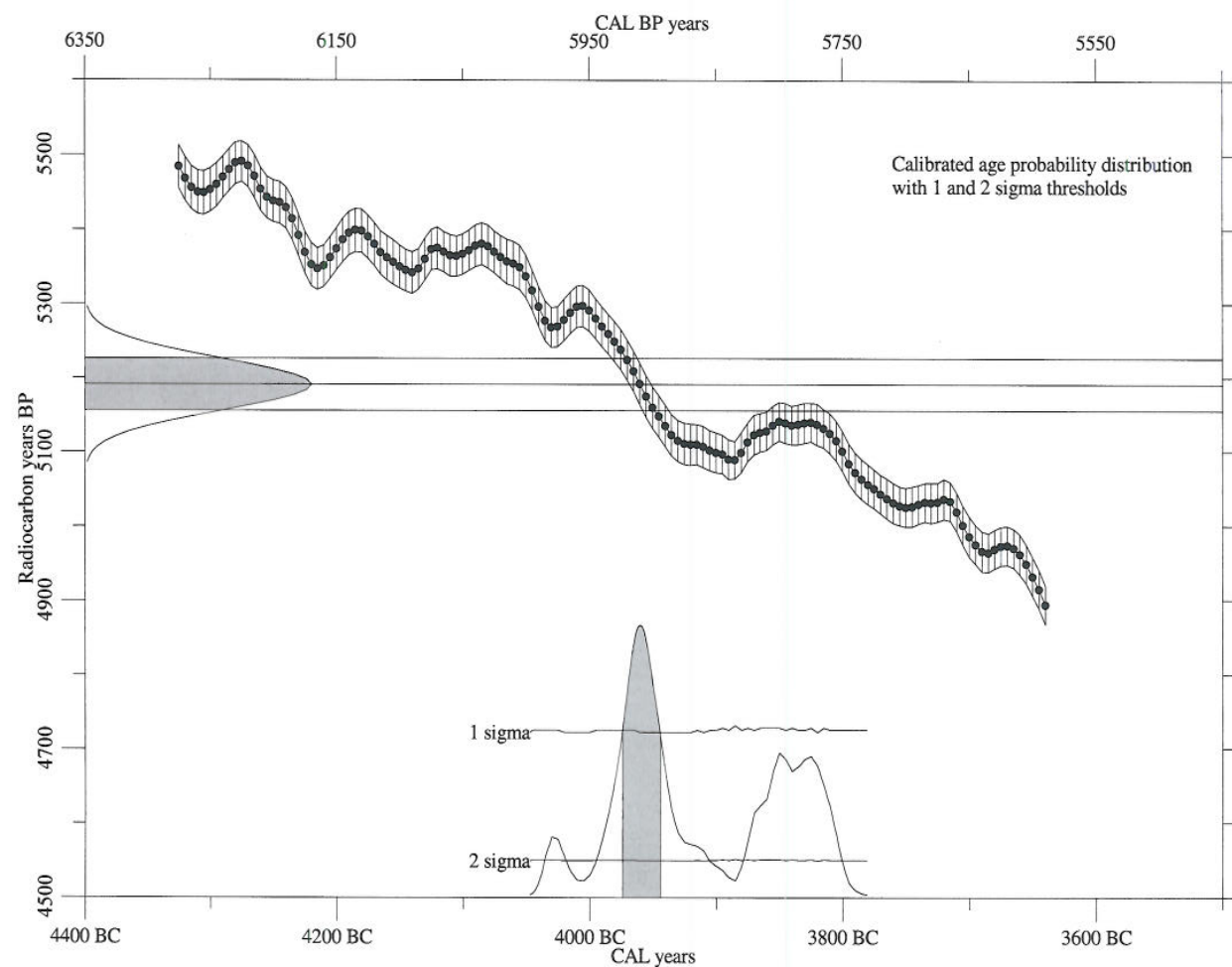Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26678 CONVENTIONAL RADIOCARBON AGE   5191 ± 35 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 4035 BC to 4018 BC	5984 BP to 5967 BP (3.9% of area)
plus 3994 BC to 3905 BC	5943 BP to 5854 BP (51.0% of area)
plus 3878 BC to 3800 BC	5827 BP to 5749 BP (37.3% of area)
1 sigma interval is 3974 BC to 3944 BC	5923 BP to 5893 BP (30.5% of area)

NZA	26678
R	29386/16
Job No	37702
Measured	15-Nov-06
Issued	21-Nov-06
	Page 4 of 6

RAFTER
GNS SCIENCE

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T106 MNZ S43850 Rd partial
Description	Bone
Fraction Dated	bone gel
Submitter	Jennifer Hay Massey University

---

* Radiocarbon Age 5191 ± 35 **BP** $\delta\,^{13}C =$ $-19.5\,‰$

** Per cent modern = 52.05 ± 0.21 $\delta\,^{14}C =$ $-473.6 \pm 2.2\,‰$ $\Delta^{14}C =$ $-479.5 \pm 2.1\,‰$

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI)
  corrected for decay since 1950.

Age, $\Delta\,^{14}C$, $\delta\,^{14}C$ and absolute per cent modern are as defined by Stuiver ,Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample: tuatara bone , Tarakohe, Takaka- hard, well preserved. Microscopic exam revealed very clean, hard, very well preserved bone. Scraped overall with scalpel. Ground to 1cm and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

The reported errors comprise statistical errors in sample and standard determinations, combined in quadrature with a system
error component based on the analysis of an ongoing series of measurements on an oxalic acid standard.
For the present result the system error component is conservatively estimated as 0.18% (= ± 14 radiocarbon years).

National Isotope Centre, Institute of Geological and Nuclear Sciences Ltd (GNS Science)
PO Box 31-312 Lower Hutt, New Zealand   Fax +64 4 570 4657 Phone +64 4 570 4644
www.RafterRadiocarbon.co.nz

RAFTER RADIOCARBON LABORATORY

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
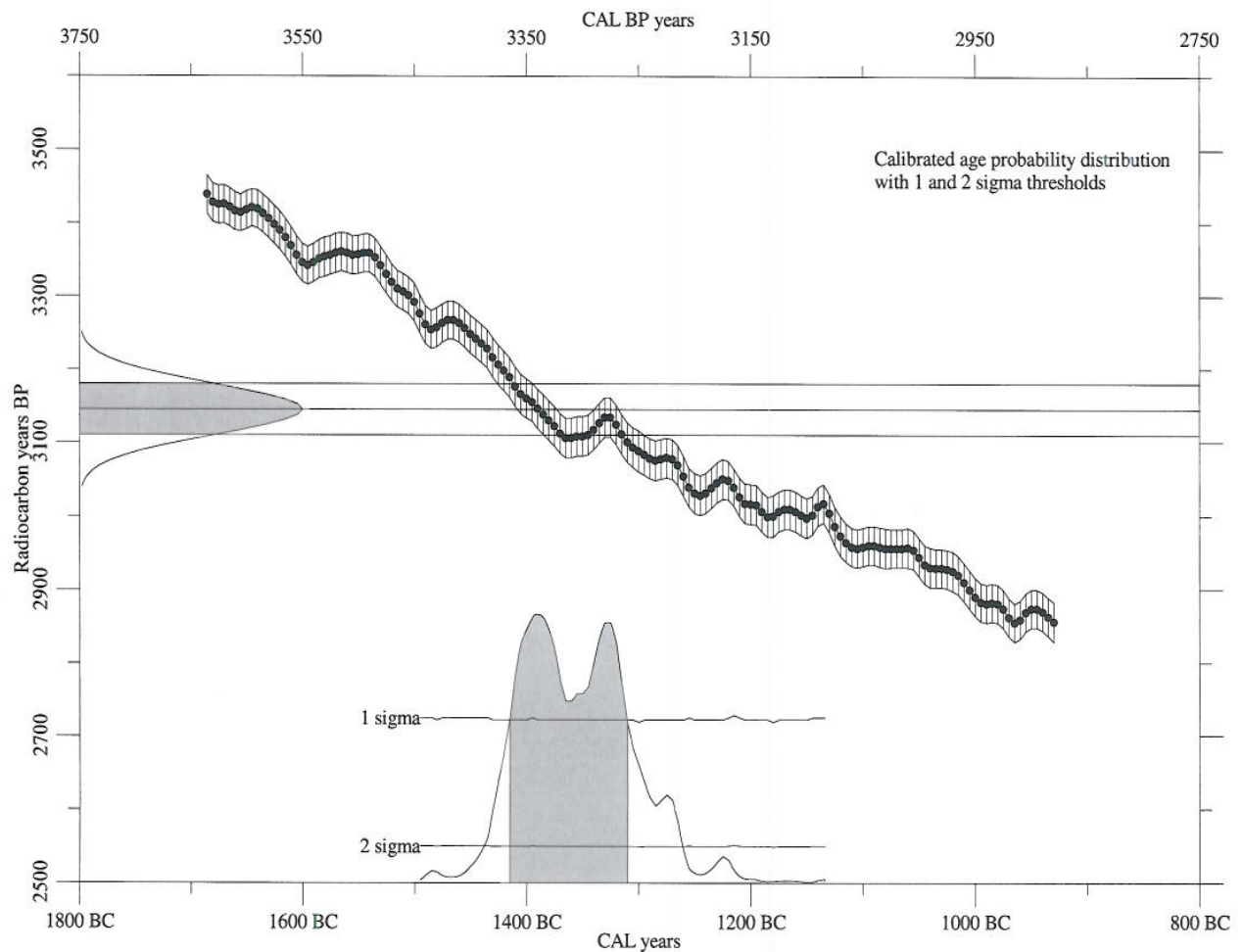Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26648 CONVENTIONAL RADIOCARBON AGE   3145 ± 35 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 1438 BC to 1260 BC	3387 BP to 3209 BP (94.8% of area)
1 sigma interval is 1415 BC to 1310 BC	3364 BP to 3259 BP (73.6% of area)

NZA   26648

R	29386/12
Job No	37698
Measured	13-Nov-06
Issued	17-Nov-06
	Page 3 of 7

RAFTER
GNS SCIENCE

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T87 MNZ S44277.54 Lmax
Description	Bone
Fraction Dated	gelatin
Submitter	Jennifer Hay Massey University

---

* Radiocarbon Age   **3145 ± 35 BP** $\delta^{13}C =$ -19.8 ‰

** Per cent modern =   67.15 ± 0.29 $\delta^{14}C =$ -321.3 ± 3 ‰ $\Delta^{14}C =$ -328.5 ± 2.9 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard   (HOxI)
    corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver ,Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample consisted of clean, very well preserved bone.Microscopic exam revealed very clean.
Treated with overnight acid demineralisation and gelatinised. Freeze dried.

Stored   none

Comments

The reported errors comprise statistical errors in sample and standard determinations, combined in quadrature with a system
error component based on the analysis of an ongoing series of measurements on an oxalic acid standard.
For the present result the system error component is conservatively estimated as 0.18% (= ± 14 radiocarbon years).

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.
PO Box 31312, Lower Hutt, New Zealand
Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26007 CONVENTIONAL RADIOCARBON AGE 696 ± 30 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 1283 AD to 1394 AD	667 BP to 556 BP (98.8% of area)
1 sigma interval is 1293 AD to 1322 AD	657 BP to 628 BP (33.4% of area)
plus 1349 AD to 1387 AD	601 BP to 563 BP (44.9% of area)
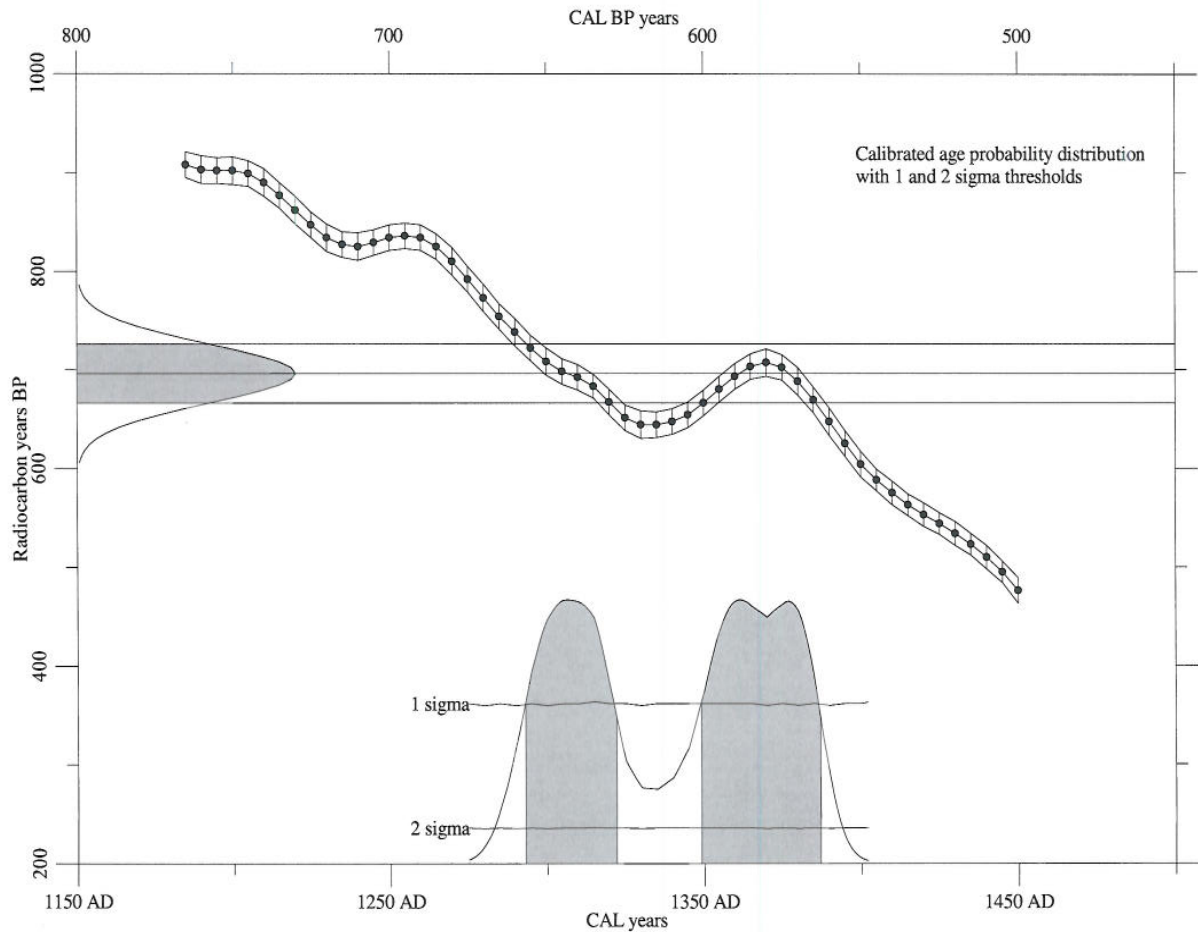
NZA   **26007**

R   29328/3

Job No   36999

Measured   24-Aug-06

Issued   31-Aug-06

Page 1 of 4

RAFTER
GNS SCIENCE

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	MNZ S44276-Ld 1
Description	Extraction T91
Fraction Dated	Gelatin
Submitter	Jennifer Hay Massey University

---

* **Radiocarbon Age**   **696 ± 30 BP** $\delta\ ^{13}C =$ -19 ‰

** **Per cent modern =** 91.08 ± 0.34 $\delta\ ^{14}C =$ -77.9 ± 3.4 ‰ $\Delta^{14}C =$ -89.2 ± 3.4 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard (HOxI)
corrected for decay since 1950.

Age, $\Delta\ ^{14}C$, $\delta\ ^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample consisted of left lower jaw from tuatara. Microscopic exam revealed very hard, clean, very well preserved bone. Scraped overall with scalpel. Ground to 1cm and treated with acid demineralisation and gelatinisation. Freeze dried.

Stored   none

Comments

INSTITUTE OF GEOLOGICAL AND NUCLEAR SCIENCES LTD.

PO Box 31312, Lower Hutt, New Zealand
Phone (+64 4) 570 4671, Fax (+64 4) 570 4657

RADIOCARBON CALIBRATION REPORT

NZA 26573 CONVENTIONAL RADIOCARBON AGE   1684 ± 35 years BP

Southern Hemisphere Atmospheric data from McCormac et al (2004);
FG McCormac, AG Hogg, PG Blackwell, CE Buck, TFG Higham, and PJ Reimer (2004)
Radiocarbon 46, 1087-1092

CALIBRATED AGE in terms of confidence intervals (Smoothing parameter: 1, Offset: 0)

2 sigma interval is 339 AD to 540 AD	1611 BP to 1410 BP (95.3% of area)
1 sigma interval is 395 AD to 435 AD	1555 BP to 1515 BP (32.4% of area)
plus 496 AD to 503 AD	1454 BP to 1447 BP (4.3% of area)
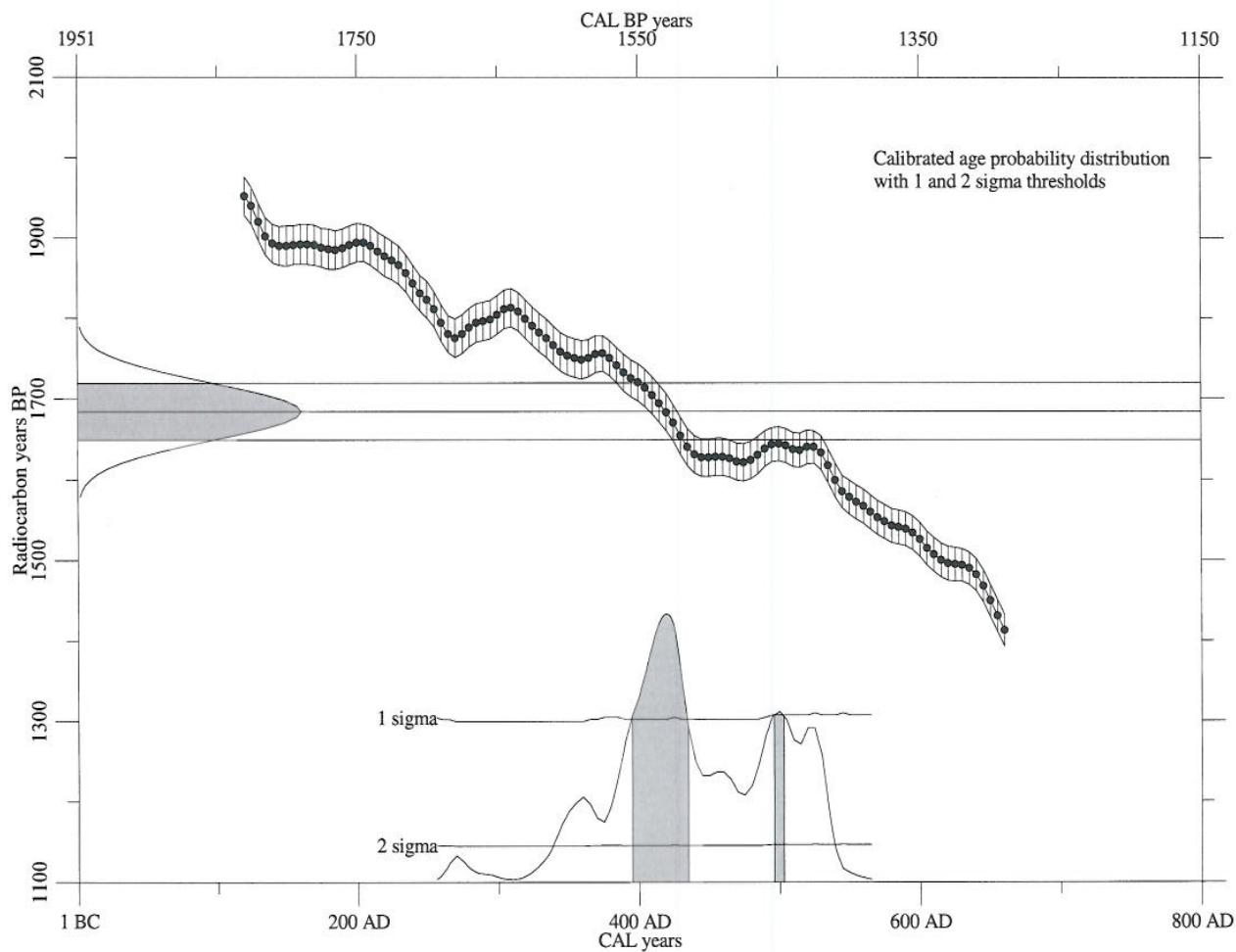
NZA   **26573**

R   29386/2
Job No   37688
Measured   26-Oct-06
Issued   27-Oct-06

Page 4 of 5

Accelerator Mass Spectrometry Result

This result for the sample submitted is for the exclusive use of the submitter.
All liability whatsoever to any third party is excluded.

Sample ID	T31 AU4030 Rd5
Description	Bone
Fraction Dated	gelatin
Submitter	Jennifer Hay Massey University

* **Radiocarbon Age**   **1684 ± 35 BP** $\delta^{13}C =$ -19 ‰

** **Per cent modern =**   80.54 ± 0.37 $\delta^{14}C = $ -184.7 ± 3.7 ‰ $\Delta^{14}C =$ -194.6 ± 3.7 ‰

* Reported age is the conventional radiocarbon age before present (BP)

** Per cent modern means absolute per cent modern relative to the NBS oxalic acid standard (HOxI)
corrected for decay since 1950.

Age, $\Delta^{14}C$, $\delta^{14}C$ and absolute per cent modern are as defined by Stuiver Polach, Radiocarbon 19:355-363 (1977)

Sample Treatment Details

Sample consisted of clean, hard, well preserved bone. Microscopic exam revealed no root hairs,
burial dirt. Scraped with scalpel overall and broke to 1cm. Treated with overnight acid
demineralisation and gelatinised. Freeze dried.

Stored   none

Comments