Molecular Systematics of Australian and New Guinea Pythons

Lesley H. Rawlings

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Genetics
The University of Adelaide

March 2001

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains not material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Lesley H. Rawlings

March 2001

Abstract

Mitochondrial *control region* and *cytochrome b* gene sequence comparisons were used to investigate the molecular systematics of pythons. Duplicate, non-tandem copies of *control region* sequence from single mitochondrial genomes were compared for divergent genera and within populations of pythons. There was almost total sequence identity between the two regions within individuals but there were differences in copy number of an ~88 bp repeat between the two regions within individuals. Structures such as conserved sequence blocks and termination sequences previously documented in vertebrate *control regions* were present and there was a $tRNA^{Ile}$ paralogue and a large 5' hairpin in the Australasian pythons that was not present in the Afro-Asiatic pythons, boids, colubrids or viperids. A range of squamate reptiles was tested for the presence of the duplication, which was found to be absent in all the squamate reptiles tested and present in all the snake lineages tested except Scolecophidians.

Phylogenetic analyses of python species suggested that the genus *Python* is paraphyletic and the Afro-Asian *Python* species formed the sister lineage to all other pythons. *Aspidites* was nested amongst the Australo-Papuan pythons, albeit without strong statistical support for this arrangement. There was low genetic divergence among all taxa, and there was monophyly of the genera *Antaresia*, *Aspidites* and *Liasis* only.

Two studies focused on phylogenetic relationships at the population level. Analyses of *Morelia viridis* nucleotide sequences and allozyme loci revealed that in New Guinea, populations north of the central cordillera were genetically different from those of the south. The Australian population was closely related to the southern form.

Analysis of *control region* sequences of *Liasis* and *Apodora* concluded that *Apodora papuana* was the sister lineage to all *Liasis* and that *Liasis fuscus*/*L. mackloti* and *L. olivaceus* are separate lineages. There was strong support for the recognition of two lineages ("eastern" and "western") of water pythons in Australia.

To conclude, the genetic split found in green pythons across the central cordillera of New Guinea was investigated further. To test if the most recent episodes of the Plio-Pleistocene climatic oscillations have had an effect on genetic subdivision of the lowland rainforest fauna, molecular phylogenies were estimated for six vertebrate taxa. The distribution of *cytochrome b* lineages and allozyme fixed differences demonstrated a pattern of historical genetic subdivision that correlates with populations distributed either side of the central cordillera. Under a molecular clock assumption, the magnitude of sequence divergence within the six taxa suggests that uplift of the central cordillera acting in concert with the earlier climatic oscillations of the Plio-Pleistocene was the most likely factor involved in genetic subdivision.

Acknowledgments

If I have seen further, it is by standing of the shoulders of giants.
NEWTON, Sir Isaac (1642-1727)

If anyone had asked me in 1990 what I thought I might be doing in the year 2001, I am sure that finishing a PhD about pythons would not have been anywhere near the possibilities of my imaginings. Living in Asia and caring for three small children was a world away from tertiary educational pursuits. When the door of opportunity opened, I stepped through into undergraduate studies and then into the challenge of the python project. Ultimately, the challenge was to maintain a balance between family life and the demands of the project/thesis and for better or worse, the challenge has been met.

This thesis and the work that it reports is not representative of my own efforts alone. Like Isaac Newton, I need to acknowledge a multitude of people who have far more experience and expertise than I, in the fields of herpetology, systematics and molecular evolution. As always with science, I have just added a few more pieces to a giant puzzle. As well as those who are experts in their fields, I have also become aware of the tremendous support from an army of encouragers. The unflagging support and genuine interest of my supervisor, Steve Donnellan has been invaluable to keep the components of this project moving forward. His professionalism, in-depth knowledge of the field of molecular evolution and vast range of contacts in fauna research and government areas gave me a great start into this project. There are so many other "bits and pieces" that didn't make it into the thesis that I hope will find their usefulness someday, that are legacies of Steve's contagious enthusiasm for the python project. I was fortunate among PhD students in that Steve was always willing to spend time explaining the nuances of terminology and statistical methodology that are part of molecular evolution. Sometimes, my brain was very slow in grasping hold of these ideas. I am also grateful for the encouragement and opportunities provided to attend conferences and present my work. These stretched and challenged my reserved nature in ways that were

painful, but profitable. I especially appreciated the opportunity to pursue a tangential dabble into the forensic science of fauna enforcement using the sequencing methods and sample database I had established for green pythons. Seeing a direct application resulting from this work was very rewarding and involvement in court proceedings was enlightening. In the long term, I have appreciated the encouragement to try new methods when they appeared in the literature and the encouragement to persist with the difficult museum specimens that have been important to the study.

I especially want to thank my beautiful family, Malcolm, Nathan, Joel and Lauren and my mother-in-law, Rita. They released me from a stereo-type and family background that said that women of my age, family status, education and work history did not start a university education and certainly not in science. They accompanied me with interest and eager participation through my undergraduate days of making sundials, counting lichens and catching slaters. They have been my biggest fans and encouragers and without their support, I would not have wanted to continue. They believed that I could go on, and they have constantly encouraged me to make it to the end, when sometimes I doubted myself. They have demonstrated pride in my achievement and in these last few months, have been willing to bear the cost of my monopoly of the computer without complaining and my single-mindedness, by picking up the domestic slack. I am looking forward to enjoying our family fun again!

I must also thank our many friends, who think I'm crazy for setting out on such a task, but who have always offered interest and encouragement and cheered me on to the finish. Even Jennie, who can hardly mention the word "snake" without becoming a neurotic wreck has been brave enough to show polite interest in Monty, our pet carpet python and offered words of encouragement when I needed them. Special thanks go to Roxanne, Dave and Steve H. who have patiently listened through four years of triumphs and self-doubts with love, acceptance, encouragement and prayer. I'm looking forward to the party!

Throughout my time in the Evolutionary Biology Unit, there has been a very happy

blend of hard work and fun with fellow EBUers. Agatha, Brenda, Brian, Caroline, Cat, Chris, David, Debbie, Jan, Kathy, Kylie, Kym, Leanne, Malcolm, Mark, Mike, Penny, Ralph, Remko, Sarah, Scott, Steve C. and Terry have all added to a great time. Agatha, Jan, Steve C., Terry and Jenny Thurmer have all given wonderful moral support during times when I wasn't sure that I could keep going, particularly following the death my mother.

From the purely practical requirement to obtain tissue samples, the list of those who have assisted in the field and amongst the museum pickles is very long. The SA museum's python sampling across Australia is now very extensive and even if I had been given the funding to make wide-ranging field trips, I could not have covered such a vast geographic area myself. I would like to thank A. Allison, K. Aplin, D. Armstrong, D.G. Barker, G. Bedford, S. Bielke, F. Bonnacorso, M. Bonnett, J. Bredl, B. Bush, A. Carey, M.Chambers, N. Charles, S. Cooke, P. Couper, M. Cunningham, F. Delpiva, S. Donnellan, A. Emmett, D. Farlam, T. Flannery, R. Foster, G. Fyfe, G. Gillespie, R. Hancock, T. Herbert, J-M. Hero, G. Hollis, R. How, S. Hudson, G. Husband, M. Hutchinson, S. Irwin, R & J. Jarman, C. Kishinami, M. Kearney, P. Kendrick, M. King, P. Krauss, D. Leyden, D. McCrae, J. McKenzie, W. Mannion, R. Martin, G. & S. Maynes, J. McKenzie, M. McCoy, T. Morley, L. Naylor, D. Neijelske, M. O'Brien, D. O'Connor, H. Owens, D. Pearson, P. Quill, T. Reardon, J. Reid, S. Richards, D. Robinson, P. Robinson, M. Ryan, R. Sadlier, C. Schneider, T. Schwaner M. I. Sidik, L. Smith, S. Stone, D. Storch, S. Sutherland, G. Torr, R. Turnbull, L. Valentine and F. Yuwono for providing materials. These are people who have an appreciation for pythons from many varied perspectives, some are researchers, some are fauna rescuers, some are keepers and some just have an affection for their friend who lives in the garden shed.

I would like to thank the Trustees of the D.R. Strank's Travel Fellowship for awarding me the opportunity to travel to Indonesia. I thank S.N. Prijono, director of the Bogor museum and I. Sidik, curator of reptiles, for their assistance and a very special thank you to Hari S. Basuki for giving up his time to be my driver and translator. I also thank C. Kishinami and L.

Smith for their hospitality at the Bernice P. Bishop and Western Australian Museums respectively.

I thank G. Baldock, J. Birrell, R. Foster, C. Horne and S. Simms for laboratory assistance, D. Broderick for performing some of the caesium-chloride mtDNA purifications, P. Ellinger and B. James for writing computer programmes and S. Chenoweth for assistance with Raggedness analysis. Thanks to S. Cooper and Y. Kumazawa for providing unpublished primer sequences and to M. Hutchinson for helping me to understand snake morphology. A special thank you must go to Jan Birrell for her willingness to share her vast knowledge of computer programs and laboratory experience. I would also like to thank K. Aplin, M. Adams, S. Cooper, R. Foster, G. Hope, M. Hutchinson, R. Leijs, B. Pillans, and S. Smith, for critical comments on various chapters of this thesis. The study was funded in part by an Australian Research Council grant - A19602789 to S. C. Donnellan and M. N. Hutchinson.

I dedicate this thesis to my encouragers:

God,

the author of the mysteries of the universe;

Malcolm,

my soul-mate, who always believed that I could;

Joel, Lauren & Nathan,

who've shared their mum with this thesis for four years;

Roxanne, Dave & Steve H.,

who think I'm crazy, but have listened anyway.

Preface

Pythons are animals that have been feared and revered by thousands of people for thousands of years. They have captured the imaginations' of storytellers and artists and the fascination of lovers of natural history by their beauty and variety. Certainly, having never been exposed to pythons in my own natural environment, I started these projects with a very limited understanding of the their habits and biology. As such, my appreciation for these creatures increased considerably as I came into contact with those who live with wild pythons in their sheds, breed them for a hobby, who study them from a fascination to learn more or who are passionate about preserving their environment.

Attempts to describe, categorise and classify pythons cladistically on the basis of morphology have caused many debates as authors agree and disagree with each others terms of reference. As molecular techniques have entered the systematics arena, there have been similar discussions and debates. Techniques and methods of analysis are themselves continually evolving and often seemingly clear-cut results are later superseded. The results presented in the following chapters add to the information available in the python puzzle and hopefully will provide building blocks for further development of our understanding of the evolutionary relationships of pythons.

The investigations of the present study have used molecular techniques to address a varied range of issues involving pythons, from intra-individual similarity to population variation to subfamily systematics. Chapters 3-7 are each complete projects within themselves. Therefore, each chapter has been written as a separate unit, almost ready for publication, complete with samples tested, results and discussion. There is some overlap between projects in samples used and therefore, in nucleotide sequences produced. In order to maintain the integrity of each data set used for analysis, the sequences have been duplicated in the sequence alignments. However, to reduce the amount of duplication, all materials other than tissue samples tested, and all methods have been collated together in Chapter 2 with reference made in each project chapter as necessary.

Table of contents

Chapter 1.

Introduction

1.1 Overview

The molecular systematics of Australian and New Guinean pythons is an incredibly broad topic for study and it has been very difficult to limit the questions pursued in the present study to a number that could be reasonably completed within an appropriate time frame. In fact, the temptation to pursue tangents has resulted in a number of pilot studies that will not be reported here. The scope of the present study ranges from the nucleotide and secondary structure evolution at intra-organelle and gene levels to the evolutionary relationships determined from genetic divergences at the inter-generic and population levels.

There are several snake groups, including pythons, that have a duplication in the mitochondrial genome of the region responsible for the control of genome replication, the *control region* (Kumazawa *et al.*, 1996). I chose to commence my study of python systematics with an investigation of the molecular evolution of the *control region* (*CR*) and *control region-like* (*CRL*) genes in pythons. The *CR* and *CRL* genes are evolving in a concerted manner and I investigated whether the sequence evolution between the two copies could be detected at a population level. I compared intra-organism (*Morelia spilota* and *M. viridis*) sequences from individuals that had inter-organism *CR* divergences of between 0 and 6% and also incorporated some comparisons within species of other genera and only detected two base changes within a 700 bp region for all the taxa examined. I also discuss some of the secondary structure features that are present in pythons that are not found in other snakes examined to date.

At the outset of the project, a comprehensive molecular systematics study of pythons had not been reported and there was a general disagreement in the literature as to the evolutionary relationships amongst pythons. It was appropriate to start the present study from the broad perspective of python phylogeny, by comparing the genetic relatedness of

most python species, as determined from two mitochondrial genes, to the proposed phylogenetic hypotheses of three comprehensive morphological studies each of which suggested quite different outcomes.

Large geographic ranges and diverse habitat use of many pythons species, close morphological similarities between species and descriptions of species complex(es), invite the pursuit of many interesting intra-specific python studies (*e.g.* population studies of the Woma python, *Aspidites ramsayi*, the carpet python, *Morelia spilota* and the children's python (*Antaresia* species complex). However, in the present project I limited my investigation to two species, the green python *Morelia viridis* (Chapter 5) and the water python *Liasis mackloti* (Chapter 6) to determine the phylogeographic structure of populations. I chose these species because their distributions span the Australian-Melanesian region, an important region in the biogeographic history of Australia.

Because of the green python's significance to the Australian fauna constituency, (*i.e.* it has a very small distribution in Australia relative to offshore) and its popularity and high market price amongst reptile keepers, the green python study has implications for wildlife management, fauna enforcement and captive breeding programs. Chapter 5 deals with three populations, northern New Guinea, southern New Guinea and Australia and reports a distinct phylogenetic split between populations separated by the central mountain range of New Guinea.

For water pythons, there has been debate as to whether there are two or more species forming a species complex, with the Australian and New Guinea populations being separated morphologically from the Indonesian populations and the Indonesian island populations being considered by some authors as sufficiently distinct to warrant their recognition as species (Barker and Barker, 1994; Brongersma, 1956; Stuhl, 1932). In Chapter 6, I demonstrate the monophyly of most of the Indonesian island populations and report a phylogenetic split that is not divided by the Arafura Sea, but is found within Australia. I discuss the implications of the Australian Plio-Pleistocene landscape in the formation of this genetic differentiation.

To conclude, the theme of genetic diversification being correlated with the central New Guinea mountain range is further considered in Chapter 7 by expanding the findings of the green python study to include a range of organisms whose rainforest distributions occur either side of the mountain range to an elevation of approximately 2000 metres. The present study included two pythons, two boas, a colubrid snake and a rodent. I used partial *cytochrome b* (*cytb*) sequence to establish a phylogeny that showed a genetic subdivision between populations either side of the mountain range. I used a molecular clock on genetic distances established from almost complete *cytb* sequences to estimate a time of divergence for the northern and southern populations.

1.3 Phylogeography

In phylogeographic analysis clades established by allelic phylogenies of populations are compared with the geographic distributions of those populations (Avise *et al.*, 1987). The mitochondrial gene phylogenies of many species show geographical localisation of clades (Lamb and Avise, 1992; Quinn, 1992). An example of phylogeographic structuring has been shown in a compilation of data summarised from studies conducted by Avise (1992) on the mtDNA variation of North American east coast marine and coastal species such as seaside sparrows, American oysters, black sea bass and the horseshoe crab. Samples taken from the Gulf of Mexico were compared with samples from the Atlantic coastline. Out of 10 species studied, at least 5 and possibly 8 showed subdivision in the mtDNA haplotypes corresponding to either Atlantic or Gulf populations. Thus, Avise concluded that species should not be viewed as monotypic, but as different geographic populations with hierarchical genetic and historical structure. Conservation of these different structures is an important consideration in maintaining genetic diversity. I use the concept of phylogeography to make biogeographic assessments of evolutionary history for water pythons and six New Guinea taxa.

1.4 Choosing the marker

The choice of genetic marker, and thus the analytical techniques applied, varies according to the question being asked. Mitochondrial DNA analysis is widely used as a tool in the study of evolutionary biology. Some advantages of using mt DNA are: i) it has a rapid rate of base substitution compared with nuclear DNA (0.5% - 1% per million years for primates, 5-10 times that of nDNA (Moritz *et al.*, 1987), ii) it is effectively haploid, iii) it is assumed there is no recombination, although evidence of mitochondrial recombination is emerging (Lunt and Hyman, 1997; Thyagarajan *et al.*, 1996), iv) mtDNA is transmitted through the maternal lineage, reducing the effective population size, and increasing the sensitivity to genetic drift (Birky Jr *et al.*, 1989), v) it is compartmentalised in the cell cytoplasm away from the nuclear genome and can be easily isolated (Dowling *et al.*, 1996) yet it is widely distributed so that comparisons can be made between evolutionally divergent organisms, and vii) it lacks introns, pseudogenes (Avise *et al.*, 1987; although see chapter 3 and Macey *et al.*, 1998; Moritz, 1994) etc. making its structure relatively simple. The mitochondrial genome is a closed circular molecule about 16 kilobases (kb) long for mammals (Cantatore and Saccone, 1987) and about 16-22 kb long for snakes (Densmore III *et al.*, 1992; Kumazawa *et al.*, 1998). There are 37 encoding genes, comprising 13 polypeptides, 2 ribosomal RNAs (12S and 16S) and 22 transfer RNAs. The *CR* is the major non-coding sequence of about 1.2 kb that contains the origin of replication for the heavy strand (Saccone *et al.*, 1991). The *CR* is prone to insertions, deletions and accumulations of short tandem repeats accounting for the differences in the size of the genome (Moritz and Brown, 1987). In the placental mammals, fish and frogs studied to date, the gene order is conserved, though it is different from that found in birds (Desjardin and Morais, 1990) and marsupials (Pääbo *et al.*, 1991). The *CR* is often used for intra-specific phylogenetic studies because it is the most variable part of the mitochondrial genome (Walberg and Clayton, 1981) with an estimated sequence divergence of about 2% per million years in mammals. For taxa separated by longer evolutionary distance a more slowly evolving gene such as *cytochrome b* (*cytb*) or *12S rRNA*

can be used to give information about the evolutionary relationships (Crozier, 1990). Liver is a suitable tissue from which to amplify mitochondrial genes as mitochondria may be present in 5,000-10,000 copies per cell.

Cytochrome b has been the gene of choice for many phylogenetic studies because it has an intermediate rate of mutation, however, its use has been confounded by the presence of nuclear paralogues that can be non-discriminately amplified by the Polymerase Chain Reaction (PCR) (Arctander, 1995; Thorpe *et al.*, 1994). Because it has a faster rate of nucleotide substitution than coding genes, *CR* has often been used to determine intra-species phylogenies.

Nuclear markers have been used in mammalian phylogenetics for many years, but, with the exception of rRNA loci, have not been accessible for non-mammalian vertebrates until recently (e.g. Saint *et al.*, 1998). Microsatellites provide information for familial studies, but nuclear markers that incorporate introns with sufficient variability to be useful at a population level are rare for non-mammalian species. In pythons, there is insufficient variability (1-3%) between divergent genera for protein encoding genes such as *c-mos* and *RAG1* (data not shown). Where the tissue had been collected and stored appropriately (Richardson *et al.*, 1986), allozyme electrophoresis was also done by Steve Donnellan to screen for a large number of nuclear loci, as allozyme electrophoresis is still the most cost effective method of screening large numbers of nuclear loci (Murphy *et al.*, 1996) for studies of population structure. Allozyme fixed differences could be used to differentiate between northern and southern New Guinea populations of six taxa.

For the studies within the scope of the present project, *cytb* and *CR* were piloted as possible markers to differentiate between and within species. *Cytochrome b* has been useful to establish the north-south New Guinea split in two taxa, but was not variable enough to provide a marker to distinguish the Australian green pythons from the southern New Guinea populations. The *CR* was sufficiently variable to differentiate between these two populations. The timing of the python radiation(s) has not been determined, so for the python phylogeny

mid/fast-evolving *cytb* gene and fast-evolving *CR* were used.

1.5 Gene Tree vs. Species Tree

Chapter 6 discusses the phylogeny and phylogeography of water pythons. The tentative conclusions of the present study are based of the examination of a single gene, the mitochondrial *control region*. On the basis that a gene tree may not be representative of the species tree argument, I conclude that further studies should be undertaken to test my hypotheses. Even though any family of bi-parental, sexually reproducing organisms has a single pedigree relating them through a chain of parent-offspring genetic transmission, individual genes within this pedigree may not have filtered through in an identical way (Avise, 1994). Due to stochastic lineage sorting, gene duplication or horizontal transfer, phylogenetic analysis of one gene locus may give a different tree topology to a tree derived from the analysis of another unlinked locus (Fitch, 1970). As some alleles may become extinct through time, allelic associations that will determine the gene tree of a single gene, may not be representative of the species tree. To give a more robust estimate of the species tree, it is necessary to make phylogenetic analyses of more than one gene. Because the mitochondrial genome is inherited as a single unit, analysis of a second mitochondrial gene should yield the same, albeit more robust, phylogenetic information. Thus, if possible, it is advisable to also include a number of unlinked nuclear genes in the analysis (Brower *et al.*, 1996). The occurrence of incongruence between gene trees causes difficulties in interpretation of the species tree. There are many possible causes of incongruence (Brower *et al.*, 1996) including incompatibility of analytical methods and character incompatibility (homoplasy). The development of a models of evolution and statistical tests to estimate when the data are not conforming to the model (de Queiroz *et al.*, 1995) attempt to reduce the latter cause of incongruence (see below). Brower *et al.* (1996) conclude that "explaining incongruence is not a task of systematists", I, however, in the present study, have attempted to offer possible explanations for such inconsistencies.

Chapter 2.

Methods

In this chapter, I detail all the methods used in the present study and therefore, when outlining methods in each subsequent chapter, I will only mention the procedure name and any modifications appropriate to that chapter. All specimens used will be listed in the appropriate chapter with institution codes following Leviton *et al.* (1985), except for ABTC = Australian Biological Tissue Collection, South Australian Museum, Adelaide.

2.1 Tissue collection and storage.

Liver, heart, muscle or scale tissues were used for DNA extraction. Tissues collected from freshly killed animals were snap-frozen in liquid nitrogen and stored at -80°C. Tissues (preferably liver) collected opportunistically from road-kills and scales were preserved in either 70% ethanol or ethanol:saline (1:1) solution. Scales were clipped from gently-restrained, live snakes using clean scissors and forceps. The snake, if wild caught, was then released at the point of capture. Shed skins were frozen at -20°C as a quarantine measure to kill any parasites and kept frozen at -20°C for long term storage.

2.2 Genomic DNA Extraction

DNA was extracted using a standard phenol-chloroform (Sambrook *et al.*, 1989) or salt (Miller *et al.*, 1988) extraction method. Preliminary steps in the extraction process varied according to the tissue type and method of preservation.

2.2.1 Tissues in alcohol

To remove the alcohol, the tissue was centrifuged at 10,000 rpm for 1 min and the alcohol was removed. The tissue was washed in 1ml Tris buffer, pH8.0. spun at 10,000 rpm for 5 minutes and the supernatant discarded. The standard phenol or salt extraction protocol was then followed.

2.2.2 Phenol Chloroform Extraction

0.2-0.5 gms of frozen or alcohol preserved liver, heart, muscle tissue was homogenised in 900µl Tris buffer, pH8.0. When using blood, 200µl aliquots were mixed with 500µl dH_2O and centrifuged at 10,000 rpm for 5 minutes in a Universal 30RF refrigerated centrifuge to pellet cells and the pellet was resuspended in 900µl Tris buffer, pH8.0. The tissue homogenate, resuspended blood pellet or allozyme electrophoresis lysate was mixed with sodium dodecyl sulfate (SDS) (1% final concentration) and Proteinase K (0.2mg/ml final concentration) and incubated at 55°C for 2-3 hrs to lyse the cells. To extract the DNA, an equal volume of phenol: chloroform: isoamylalcohol (25:24:1) was added, mixed well and centrifuged at 10,000 rpm for 5 mins. The phenol/chloroform layer was removed and the process was repeated until the lysate was clear, with the final extraction using an equal volume of chloroform only. The DNA was precipitated with 0.1X vol 3M ammonium acetate (NH_4Ac) (pH5.2), and 2X vol. 100% ethanol at -20°C for at least 3 hours (or overnight) or at -80°C for at least 20 mins. The DNA was pelleted by centrifuging at 10,000 rpm for 30 mins at 4°C, then washed in 70% ethanol, dried at 50°C for 10 mins and resuspended in 100µl dH2O.

2.2.3 Salt extraction

Tissue was digested overnight at 37°C in a solution of 50 mM Tris-HCl, pH 8.0, 20 mM Ethylenediaminetetraacetic Acid (EDTA), pH 8.0, 2% SDS and 0.2mg/ml ProteinaseK. Proteins were precipitated by adding 0.5 vol saturated sodium chloride (NaCl) solution, chilled for 5 mins then centrifuged at 8000 rpm for 15 mins. The supernatant was collected and the DNA was precipitated by adding 2X vol. 95-100% ethanol at -20°C overnight. After centrifuging for 30 mins at room temperature, the DNA pellet was washed with 70% ethanol dried at 55°C for 10 mins and resuspended in 100µl dH2O.

2.3 Mitochondrial enrichment methods

2.3.1 Isolation of mitochondrial DNA using caesium chloride PI gradient

This method follows Dowling *et al.* (1996). Approximately 0.5gm of frozen liver tissue was homogenised thoroughly in cold Sodium Chloride-Tris-EDTA-Sucrose (STES) buffer, then centrifuged for 5 mins at 1200*g* (2590 rpm in a 4.5 rotor Sorvall RC-5B centrifuge) at 4°C to pellet the nuclear debris. The supernatant was transferred to a 50ml Oakridge centrifuge tube and centrifuged at 23,000*g* (11,339 rpm) for 20 mins at 4°C to pellet the mitochondria. The supernatant was discarded and the pellet resuspended in 1.0ml Tris-EDTA (ThE) buffer and 0.125ml (1/8 resuspended volume) 20% SDS (w/v in H_2O), for 10 mins to lyse membranes. To precipitate nuclear DNA-SDS, 0.188ml (1/6 volume) Caesium Chloride (CsCl)-saturated water was added and left at 4°C overnight.

The sample was centrifuged for 10 mins at 17,000g (12,000 rpm in 4.25" SS 24 rotor) at 4°C to pellet nuclear DNA. The supernatant was transferred to a 5ml ultracentrifuge tube and layered in a CsCl gradient. Caesium-chloride was added to bring the sample density to 1.40gm/ml. For 1.28gm of sample, 0.81gm of CsCl and 0.23mg/ml propidium iodide (PI) stock (in ThE) was added (Dowling *et al.*, 1996). To form a step gradient the sample was underlaid with 1.33ml of 1.70gm/ml CsCl per ml of sample (1.70ml) and overlaid with mineral oil. Samples were centrifuged for 24 hours at 140,000g (36,000 rpm in a SW60Ti rotor, Beckman L8-70 Ultracentrifuge) in a temperature range of 21°C-35°C.

Mitochondrial DNA was collected from the tube by taking fractions with a sterile syringe and 23 gauge needle under UV light. The nuclear fraction was also collected separately. The PI was extracted with an equal volume of CsCl-saturated isopropyl alcohol. The DNA samples were placed into prepared dialysis tubing (see below) and dialyzed for 24 hours in two changes of 500mls 0.5 X ThE buffer. The mtDNA was then removed from the tubing to a sterile 1.5 ml eppendorf and stored at 4°C until required.

2.3.1.1 Preparation of Dialysis Tubing

Tubing was cut into 20cm lengths and boiled vigorously in 2% sodium bicarbonate, 10mM EDTA pH 8.0 for 10 mins. After rinsing thoroughly in distilled water it was boiled a further 10 mins in 10mM EDTA pH 8.0 and allowed to cool. The tubing was stored in sterile dH_20 until needed.

2.3.2 Mitochondrial DNA enrichment

This method of mitochondrial DNA enrichment was modified from the plasmid isolation mtDNA enrichment method of Welter *et al.* (1989). 0.1-0.2gm of liver tissue was homogenised in 1 ml ice-cold homogenisation buffer (50mM glucose, 10mM EDTA, 20mM Tris-HCl pH8.0) and centrifuged at 1600rpm for 10 mins. The supernatant was transferred to a fresh tube and centrifuged a second time at 1000rpm for 10 mins. The supernatant was collected and centrifuged at 12,000rpm for 20 mins to pellet the mitochondria. The pellet was gently resuspended in 1ml 1X TrisHCl-EDTA (TE) and placed on ice for 10 mins. To concentrate the mitochondria, the samples were centrifuged at 12,000rpm for 20 mins and the pellet was resuspended in 100µl homogenisation buffer. 200µl of lysis solution (200mM sodium hydroxide [NaOH], 1% SDS) was added, mixed gently and placed on ice for 5 mins. 150µl of acetate neutralising solution (11.5% glacial acetic acid, 3M potassium acetate) was added, vortexed for 10 secs and placed on ice for 60 mins. The mixture was then centrifuged at 10,000 rpm for 20 mins to pellet any remaining cellular debris. The supernatant was extracted with phenol/chloroform and the mitochondrial DNA was precipitated with 0.1 vol 3M sodium acetate (NaAc) pH5.2 and 2X vol 100% ethanol at -20°C overnight. The DNA was pelleted by spinning at 10,000 rpm for 30 mins and the pellet washed in 70% ethanol and air dried. The DNA was resuspended in 50µl TE/RNAase (10mg/ml) and incubated at 37°C for 30 mins.

2.4 Titration test to reduce the amplification of nuclear paralogues

The presence of paralogous sequences in the genomes of organisms confounds phylogenetic analysis, which requires the comparison of orthologous sequences to correctly estimate relatedness of alleles by descent. Paralogous sequences occur when a gene or part of a gene is duplicated in the genome and then each copy evolves independently. The time since the duplication event and the rate of nucleotide substitution will determine the amount of similarity between the paralogues. Often, paralogous sequences are detected by the presence of stop codons or frame shifts in coding genes (Lopez *et al.*, 1994; Quinn, 1992).

There is now an accumulation of evidence supporting the translocation of paralogous sequences to other parts of the genome. Nuclear genes translocate to different chromosomes (deGrouchy and Turleau, 1984) and organellar components (*i.e.* mitochondrial and chloroplast genes) translocate into the nuclear genome (Ayliffe and Timmis, 1992; Collura and Stewart, 1995; Lopez *et al.*, 1994; Quinn, 1992; Smith *et al.*, 1992; Sorenson and Fleischer, 1996; Zhang and Hewitt, 1996). PCR amplification of paralogous sequences and their subsequent inclusion into any phylogenetic analyses will lead to inaccurate estimates of phylogeny (Arctander, 1995; Collura and Stewart, 1995). Donnellan *et al.* (1999) developed a procedure to test the potential for mitochondrial primers to amplify mitochondrial genes rather than nuclear paralogues. Serial dilutions from neat to 10^{-5} of caesium-chloride purified or enriched mtDNA were amplified with three sets of primers: the "universal" *12S rRNA* (L1091/H1478) or *16S rRNA* (16Sar/16Sbr) and the nuclear *18S rRNA* primers (18e and G59) and the mitochondrial primers being tested. (All primer sequences and authors are listed in Table 2.1). Usually the mitochondrial primer pairs amplified products at a dilution of 10^{-5} and the nuclear primers amplified products to a dilution of 10^{-2}. If the endpoint for the test primers was similar to the *12S rRNA* primer pair and both amplified a product from an aliquot at least 100 fold more dilute than the nuclear primers, it was concluded that the product was most likely derived from mtDNA. If the endpoint of the test primers was similar to that of the nuclear primers, it was concluded that the test primers had most likely amplified a nuclear

Table 2.1 PRC primers used in this study
Sources: 1 – Kumazawa *et al.* (1996); 2 – Y. Kumazawa, Nagoya University, Japan; 3 – Kumazawa and Nishida (1993); 4 – Kocher *et al.* (1989); 5 – Cunningham *et al.* (1992); 6 – Hillis and Dixon (1991); 7 – S. Cooper, South Australian Museum, Adelaide, Australia.

Primer	Sequence	Source
CR		
L14973	5'-CACATCACYCGAGATGTCCCCTACGG-3'	1
H690	5'-GTTGAGGCTTGCATGTATA-3'	1
L15926	5'-AAAGCATTGGTCTTGTAARC-3'	2
Snake17 H	5'-TATGTCTAACAAGCATTAAG-3'	2
Snake1 L	5'-CCTATGTATAATAATACATTAA-3'	2
Snake6 L	5'-ACCCTTCCCGTGAAATCC-3'	2
Snake7 H	5'-TGAAAGGATAGAGGATTTCACG-3'	2
L4160m	5'-CGATTCCGATATGACCARCT-3'	3
H4433	5'-AACCAACATTTTCGGGGTAT-3'	3
Cytb		
H15916	5'-GCCCAGCTTTGGTTTACAAGA-3'	2
Snake12 L	5'-CAGCCAAYATCAAYCTAGCATTTTCATC-3'	2
L14861	5'-AAAAGCTTCCATCCAACATCTCAGATAAA-3'	4
H15149	5'-AAACTGGCAGCCCCTCAGAATGATATTGCCTCA-3'	4
12S RrNA		
L1091	5'-AAAAAGCTTCAAACTGGGATTAGATACCCCACTAT-3'	4
H1478	5'- TGACTGCAGAGGGTGACGGGCGGTGTGT-3'	4
16S rRNA		
16Sar	5'-CCGGTCTGAACTCAGATCACGT-3'	5
16Sbr	5'-CGCCTGTTTATCAAAAACAT-3'	5
18S rRNA		
18e (F)	5'-GGTTGATCCTGCCAGTAG-3'	6
G59 (R)	5'-GCTGGCACCAGACTTGCCCTCC-3'	7

paralogue and an alternative set of test primers was considered. Also, the sequence of the product derived from a high dilution of the enriched mtDNA was compared with the sequence of a product derived from total cellular DNA using the sequencing procedures given above to test the origin of sequence derived from total cellular DNA. If these sequences were indistinguishable, it was concluded that the primers only amplified mtDNA.

2.5 Gel Electrophoresis

2.5.1 Agarose

To visualise the extracted DNA and PCR products, 5µl aliquots with 2µl of loading buffer were loaded onto a 1.5% agarose gel. The gel was covered with 1X Tris-boric acid-EDTA (TBE) running buffer and electrophoresed under standard conditions at 105V (50mA), stained with 0.5µg/ml ethidium bromide and viewed under UV light. The size of PCR products less than 1kb were gauged against a 100 bp ladder marker (GeneWorks) and products greater than 1kb were gauged against *Hind*III/*Eco*RI digested DNA standard size marker (GeneWorks).

For gel purifying single bands, either 1.2% low-melt NuSieve® GTG® agarose (Bioproducts, USA) or 6% polyacrylamide (PAGE) gels were used.

2.5.2 Polyacrylamide gels

PAGE gels were made with 3.5ml 10X TBE buffer, 26.25 ml H_2O and 5.25ml acrylamide, to which 245µl 10% ammonium persulfate (APS) and 40 µl NNNN-Tetramethylethylenediamine (TEMED) was added. The acrylamide mixture was syringed between two prepared glass plates, a comb was inserted and the gel was polymerised for 30 mins. The plate was attached to an upright gel rig, immersed in 1X TBE buffer and pre-electrophoresed for 15 mins at 250V. 5µl of PCR product with 1µl loading buffer was loaded into each well and electrophoresed for 1.5 hrs.

2.6 Oligonucleotide PCR and sequencing primers

Oligonucleotide primers for PCR and sequencing were ordered from Gene Works (formerly Bresatec Pty Ltd.) or Applied Biosystems, USA. Primer sequences used are listed in Table 2.1. Each primer was diluted in dH2O to give concentrations of 5μM for PCR and sequencing.

2.7 PCR mixtures and thermocycling conditions

2.7.1 PCR1

A 50μl PCR mixture was used containing 0.75U of *Taq* (Promega) or Ampli*Taq*Gold (Perkin Elmer) polymerase, 5μl 10X Reaction Buffer (Promega), 4mM Magnesium Chloride ($MgCl_2$) (Promega), 0.2mM of each nucleotide (dNTP), 0.2μM each primer and 20-100ng of template DNA.

2.7.2 PCR2 - (Long and Accurate PCR, LA-PCR)

50μl PCR reactions were prepared in two separate mixes. Mix 1 contained 0.2mM of each dNTP, 0.2μM of each primer, 20-100ng of template DNA and water to make the volume to 20μl; Mix 2 contained 1.6mM $MgCl_2$ (Buffer A + Buffer B), 1μl of ELONGASE Enzyme Mix Gibco/BRL) and water to make the volume to 30μl. This reaction was hot-started by mixing Mix 1 and 2 at 94°C.

2.7.3 PCR programs

2.7.3.1 Program 1
(annealing temp 48°C, 52°C or 55°C) on a FTS-320 Thermal Sequencer (Corbett Research).

94°C (2 mins), 48°, (1 min), 72° C (20 secs) X 1;

94°C (45 secs), 48°, (45 secs), 72°C (1 min) X 34 cycles ;

72° C (6 mins), 26°C (10 secs) X 1

2.7.3.2 Program 2
for LA-PCR on an Omn-E Thermal Sequencer (Hybaid)

94°C (30 secs);

94°C (30 secs), 52°C (1 min), 72°C (15 mins) X 30;

30°C (10 secs)

2.7.2.3 Program 3
(sequencing) on a FTS-1 Thermal Sequencer (Corbett Research)

94°C (30 secs), 50°C (15 secs), 60°C (4 mins) X 25 cycles;

60°C (4 mins), 25°C (10secs)

2.8 Nested PCR

2.8.1 Nested PCR of the control region

To avoid the amplification of a *control region-like* gene that is present in snakes between the *ND1* and *ND2* genes (Kumazawa *et al.*, 1996), nested PCR was used. An initial amplification used the primers L15926 and H690 situated in the $tRNA^{Thr}$ and *12S rRNA* genes respectively with the ELONGASE PCR protocol (chapter 2, sections 7.2 and 7.2.2). This product was then used as a template for a hemi-nested amplification with primers L15926 and snake 17 situated in the Conserved Sequence Block I of the *control region* sequence to amplify a ~850 bp product that was used for sequence analysis.

2.8.2 Nested PCR of the control region-like *gene*

To selectively PCR amplify the control region-like gene, nested PCR was used. An initial amplification used the primers L4160m and H4433 situated in the ND1 and ND2 genes respectively with the ELONGASE PCR protocol (chapter 2, sections 7.2 and 7.2.2). This product was then used as a template for a hemi-nested amplification with primers L4160 and snake 17 situated in the Conserved Sequence Block I of the *control region-like* sequence to amplify a ~850 bp product that was used for sequence analysis.

2.9 Purification of PCR products

2.9.1 Bresaclean

PCR products were purified using a "BRESA-CLEAN™" DNA purification Kit (Gene Works, formerly Bresatec Pty. Ltd.). 45 µl PCR product was aliquoted into a 1.5ml eppendorf tube to which 3X volume of "BRESA SALT" (6M NaI stock solution) and 5µl "BRESA-BIND™" (silica glass milk) were added. This was incubated with mixing at room temperature for 5-15 mins and centrifuged for 1 min to pellet the BRESA-BIND™/DNA complex. The supernatant was removed, the pellet was resuspended in 500µl of BRESA-WASH™ (Tris, EDTA, NaCl, 50% ethanol) and spun for 1 min. The supernatant was removed and the pellet was dried on a heating block at 55°C for 5 mins. DNA was eluted into solution by resuspending the pellet in 35 µl of dH_2O, incubating for 5 mins at 55°C, centrifuging for 1 min and immediately transferring the supernatant to a new tube. The cleaned PCR product was stored at 4°C. Recovery of DNA was checked by running 4 µl of product with 2 µl loading buffer on a 1.5% agarose gel.

2.9.2 Column purification

Some PCR products were also purified using spin columns (BRESAspin PCR purification kit - GeneWorks) as per manufacturer's instructions.

2.10 DNA Sequencing

PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit containing four ABI dye-labelled dideoxy nucleotides (G, C, A and T Dye Deoxy terminators) was supplied by Applied Biosystems. Terminator premix includes dITP and the thermally stable enzyme, Ampli*Taq* DNA polymerase.

Both strands of the PCR amplified gene fragment were sequenced using a PRISM™ Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit on a FTS-1 thermal sequencer

(Corbett Research) using Program 3 (see above). 20 µl reactions were prepared using 2 µl of PCR product, 8 µl of PRISM™ and 1µl of 5 µM primer and transferred to 20µl Corbett Research capillary tubes. The reaction products were electrophoresed on an Applied Biosystems 373A DNA sequencer and Macintosh Powermac computer at the Department of Molecular Pathology of the Institute of Medical and Veterinary Science (IMVS), Adelaide and the sequences were made available as chromatograms.

2.11 Generating Sequence Data & Data Analysis

DNA sequences were edited on the Sequence Editor (SeqEd) program from Applied Biosystems, Inc. The reversed complement sequence of the heavy (or reverse) strand primer was comparatively aligned with the light (or forward) strand and a consensus sequence was constructed. Where base anomalies occurred a redundancy code was used as follows: R = A or G (puRine); Y= Tor C (pYrimidine); M = A or C (aMino); K = G or T (Keto); W = A or T (Weak - 2 H bonds); S = C or G (Strong - 3 H bonds).

Sequences were aligned either by eye for coding genes or using ClustalW (Thompson *et al.*, 1994) and adjusted manually to improve alignment where necessary when sequences had multiple variable length insertions or deletions (indels). Aligned sequences were analysed using PAUP*4.0b2a (written by Swofford, 1999).

2.12 Analysis

2.12.1 Data Partitions

To deal with the gene tree *vs.* species tree issue, multiple data sets are usually considered in phylogeny building. Often, too, there are other non-molecular data sets such as morphological, physiological or biochemical characters that can be incorporated into the analysis. When different qualitative data partitions are incorporated into a single phylogeny estimate, there are three possible approaches to their combination, i) always combine, ii) never combine and iii) combine in some circumstances. The first approach, referred to as the

total evidence approach (Kluge, 1989) always combines all data sets into a single analysis. Kluge's (1989) justification for this approach was that incorporating all the available data maximises the "informativeness" and "explanatory power" of the information. The second approach, referred to as the consensus approach (de Queiroz, 1993) estimates the phylogeny of each data partition separately and then determines a consensus tree from the separate analyses. The argument for this approach suggests that information can be gained from separate analyses that might be lost in a combined analysis (Miyamoto and Fitch, 1995). Because combining partitions removes the independence of characters provided by analysis of each partition separately, the sampling variance is increased, thus increasing the probability of errors in the phylogeny estimation.

Support for the third approach of conditional combination arises from evidence of heterogeneity between data partitions. For molecular data, different genes, even linked genes such as those of the mitochondrial genome, have differing rates of nucleotide substitution and nucleotide frequencies between lineages (Wu and Li, 1985), between genes (Nei, 1987; Wolfe, 1988) and within genes (Li *et al.*, 1985). With the advancement of computing power, different analytical algorithms have been developed to incorporate new and more complex models of molecular evolution (Jukes and Cantor, 1969; Kimura, 1980; Lanave *et al.*, 1984; Rodriguez *et al.*, 1990; Yang, 1996). These models of evolution can also incorporate variable parameters such as the frequency of each nucleotide, the proportion of invariable sites and the gamma shape parameter (Yang, 1996). The gamma shape parameter (Γ) (Yang, 1996) measures the variation in the rate of evolution between sites within a sequence. If within-sequence rate variation is not accounted for, it can bias the estimation of phylogeny and the estimation of evolutionary rate. Ideally, a continuous gamma distribution best describes the evolutionary model that would fit real data, implying an independent rate for each nucleotide site. However, a continuous distribution is computationally intensive and only feasible for very small data sets (ie. < 6 taxa). The best available alternative is to use a discrete gamma model, which sets discrete classes of rates of equal probability. The gamma shape parameter

can be estimated using ML in PAUP*4.0b2a. With a move into even more complex analyses such as Bayesian probabilities (Larget and Simon, 1999; Shoemaker *et al.*, 1999; Yang and Rannala, 1997), the theory of molecular evolution modelling is still ahead of available computing power that can achieve results in a realistic time frame, particularly for large data sets.

All phylogenetic methods make specific assumptions about modes of evolution. If combining partitions that have different underlying evolutionary processes violates these assumptions, false estimates of evolutionary history may be made (Bull *et al.*, 1993). A first step in dealing with data partitions that may have different underlying evolutionary processes is to identify those partitions that are heterogeneous. There are a number of tests that attempt this. The Incongruence Length Difference Test (ILD, Farris *et al.*, 1995) compares the incongruence length difference (D) for the observed data with D for a series of randomised partitions that are the same size as the original partitions (*i.e.* with sampling replacement). D is calculated as the difference between the length of the most parsimonious tree for a combined analysis and the sum of the lengths of the most parsimonious trees from individual analyses. The null hypothesis for the test assumes combinability of partitions.

A second test of partition heterogeneity is the goodness-of-fit test as determined by the Templeton Test (Bull *et al.*, 1993; Templeton, 1983). This non-parametric test compares two tree topologies using a Wilcoxon signed-ranks test of the number of steps required by each character on each of the respective trees to determine whether the support for a hypothesis is significantly better than would be expected from random variation among the characters. To effect a conditional combinability test, each data partition is mapped on to the most parsimonious trees from a combined analysis and the Templeton test is applied to determine goodness of fit.

The third test discussed here is the Rodrigo test (Rodrigo *et al.*, 1993), named as such by Cunningham (1997). Differences in tree topology may be because the tree reconstruction method is inappropriate, may be due to a lack of significant cladistic information in the data,

may be due to different cladistic processes acting on the data or sampling error. This test uses the symmetric difference index (SDI) (Penny and Hendy, 1985) to measure the degree of similarity between the most parsimonious trees from each partition. The distribution of the SDI statistic is determined for each partition by calculating the mean SDI between the most parsimonious trees estimated from bootstrap pseudoreplicates taken from the same data partition. If the range of possible tree topologies from both partitions do not overlap, the null hypothesis that both trees estimate the same phylogeny can be rejected.

Cunningham (1997) used pairs of data partitions of well-established phylogenies to test the sensitivity of these incongruence tests to detect heterogeneity between data partitions. He found that the ILD test was best able to predict the accuracy of combined data. Also, the ILD test could be applied to multiple data partitions simultaneously and could be reapplied after the model had been adjusted by weighted parsimony to determine whether congruence was increased.

The congruence of data partitions can be tested using maximum likelihood (likelihood heterogeneity test) (Huelsenbeck and Bull, 1996). In this test, the likelihood ratio test statistic Λ compares the log likelihoods of a combined analysis with the total log likelihoods of separate analyses (ie. $\Lambda = 2(\ln L_1 - \ln L_0)$, where L_1 is the sum of log likelihoods for all data partitions). The null hypothesis (L_0) assumes that the same tree underlies all the data partitions (*i.e.* a combined analysis) and the alternate hypothesis suggests that different tree topologies apply to different data partitions.

2.12.2 Choosing the Analysis Method

Once a decision has been determined, whether or not to combine partitions, analysis can proceed. Three methods of analysis were included in this study, maximum parsimony, maximum likelihood and neighbor-joining analysis of genetic distances. Each method of analysis treats information within the data in a different manner. Maximum

parsimony searches for the phylogenetic hypothesis (*i.e.* tree(s)) that can be explained with the least number of changes through direct descent within the data. There are a number of algorithms that calculate the minimum length tree. Because changes between character states may not necessarily occur with equal frequency, parsimony analysis can incorporate weightings or costs for various kinds of transformations, *i.e.* greater importance can be given to transversions over transitions because transitions have a tendency to become saturated *i.e.* mutations accumulate and superimpose over previously mutated sites (Kimura, 1980). Thus, the observed difference in substitutions between two sequences will not necessarily reflect the actual number of mutations that have occurred over time. The transition/transversion ratio can be estimated in the program PAUP*4.0b2a using likelihood functions.

Neighbor-joining (Saitou and Nei, 1987) is a method of cluster analysis that uses pair-wise genetic distances to join taxa into a tree of relationships. Trees are constructed by starting with the taxa that have the smallest distance between them and successively joining the more distant taxa. Unlike some cluster methods (*e.g.* unweighted pair group method using arithmetic averages), neighbor-joining analysis does not assume that all lineages have diverged equally and distances calculated under various models of sequence evolution can be incorporated into the analysis. However, there is only a single tree produced and specific parameters associated with models of evolution cannot be set, making it a less robust test than analysis such as maximum likelihood.

Maximum likelihood (Felsenstein, 1981) searches for the most probable outcome of relationships (tree) that could arise from the observed data according to a specified model of evolution. The model of evolution that is best suited to a particular data set can be estimated using the computer program Modeltest3 (Posada and Crandall, 1998). Once a particular model has been selected, the specification of many parameters can be estimated from the data using PAUP*4.0b2a and a maximum likelihood search done. ML analysis has been shown to be robust to violations of the assumptions of a specified model (Felsenstein, 1981).

2.12.3 Estimating likelihood model parameters

Maximum likelihood model parameters such as base frequencies, transition/transversion (ts/tv) ratio, gamma shape parameter (Γ) and proportion of invariable sites for a particular model were estimated using successive approximations (Swofford *et al.*, 1996) of the parameters calculated using quartet puzzling (Strimmer and von Haeseler, 1996). The most parsimonious unweighted tree was used as a starting point to estimate the parameters by maximum likelihood. These parameters were then used in a puzzling analysis to generate a tree topology for the next round of parameter estimation using maximum likelihood. When the estimated parameters no longer changed, it was concluded that the best estimate of the parameters had been reached. The estimated base frequencies, transition/transversion (ts/tv) ratio and Γ were then used in a full maximum likelihood analysis under the specified model. The estimated ts/tv value was used to weight heuristic searches under the maximum parsimony criterion of optimality analyses and to choose appropriate molecular evolution models to calculate evolutionary distances.

2.13 Molecular clock

The concept of a molecular clock is that genes evolve at a constant rate, with the number of nucleotide substitutions being proportional to time. If the rate of substitution can be calculated, then a time of divergence between sequences can be estimated. The molecular clock has been applied to very deep time scales reaching to the Cretaceous/Tertiary boundary (Bromham *et al.*, 1999; Cooper and Penny, 1997). However, not all genes evolve at the same rate (as opposed to a gene evolving at a constant rate) (Nei, 1987) and sections of mitochondrial DNA (*e.g.* tRNAs) evolve 20-100 times faster than their nuclear counterparts (Brown *et al.*, 1982; Pesole *et al.*, 1999). There is also evidence to suggest that the rate of molecular evolution varies between homeotherms and poikliotherms (Martin *et al.*, 1992) and that apparent observed rates of evolution are correlated with body size (Martin and Palumbi,

1993). Therefore, to use a molecular clock it is essential to first establish that a constant rate of nucleotide substitution exists for the lineage of sequences being considered.

Rates of nucleotide substitutions can be compared between two closely related taxa using the relative rates test (Li *et al.*, 1985; Sarich and Wilson, 1967; Wu and Li, 1985). This test does not require any knowledge of divergence times between the taxa being considered. The sequences of two closely related taxa are each compared to the sequence of a more distant outgroup taxon. If the molecular distance (*i.e.* the difference in the number of nucleotide substitutions between the ingroup taxon and the outgroup taxon) is the same for each taxon being considered, the sequences have evolved according to a molecular clock. The difference between the distances can be tested statistically to determine whether any deviation from zero is significant. Wu and Li's test assumes the substitution model of Kimura (1980) which has two parameters, one for the rate of transitions and one for transversions, base frequencies are assumed to be equal. Muse and Weir (1992) used the same three-taxon principle but devised a likelihood method to compare the likelihood of sequences under the situations of no constraint or the constraint of equal rates of change that incorporated more complex models of sequence evolution. For both of these methods, a knowledge of the phylogenetic relationships among the taxa being tested is required.

An alternative to the relative rates test to determine the heterogeneity of evolutionary rates is the Tajima test (Tajima, 1993). This test is similar to the relative rates test in that it compares three taxa and it accounts for differences in rates of transition and transversion substitutions. Also, unlike the other tests, models of nucleotide substitution do not need to be incorporated into the analysis and the test can be applied even if the substitution rate varies among different sites and if phylogenetic relationships are not known.

Once a constancy of evolutionary rate has been established, a molecular clock can be invoked to estimate sequence divergence times. To calculate sequence divergence times it is necessary to have an independently derived date of origin for some of the taxa included in the sample, usually from fossil evidence (Martin *et al.*, 1992) and sometimes from the dating of

vicariant events (Caccone *et al.*, 1997). To calculate deeper divergences, Rambaut & Bromham (1998) have developed a test that incorporates a test of rate constancy and models of sequence evolution and also provides confidence intervals for the estimated times of divergence. However, this method requires fossil dates for pairs of taxa that diverged at a more recent time than the divergence under consideration.

If fossil evidence suggests a date for a common ancestor of a cluster of taxa (clade), a more recent divergence (t) within the clade (between taxon A and taxon B, Fig. 2.1) can be

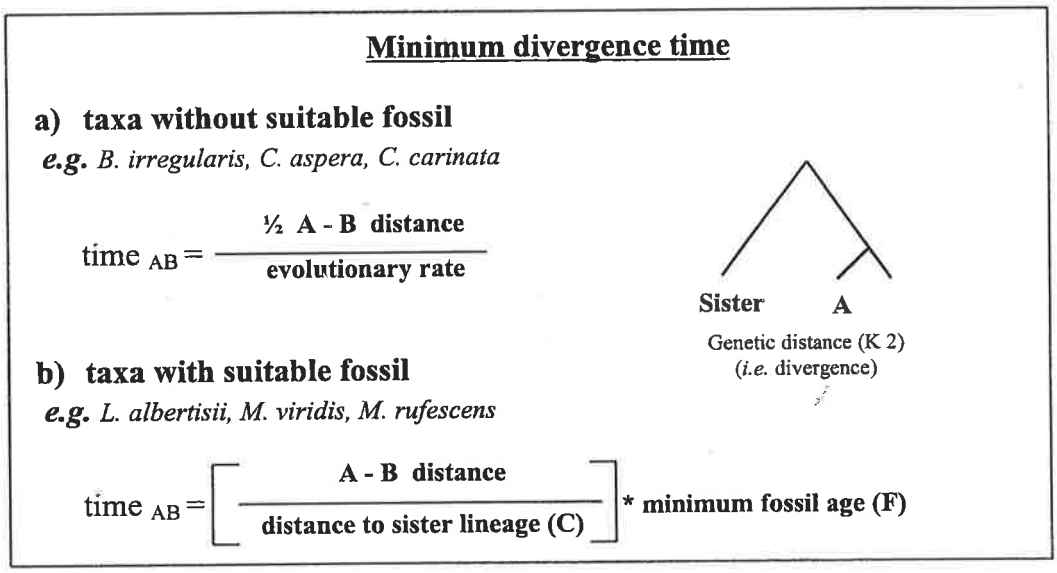


Fig. 2.1 Diagram showing the relationship between evolutionary distance and genetic distance and the formulae used to calculate the minimum divergence times between taxa.

estimated by dividing the sequence divergence between the clades (*divergenceAB*), by the average sequence divergence between the test species and the nearest sister lineage (*divergenceC*) and multiplying this value by the age of the calibration fossil (F).

$$t_{AB} = \frac{divergence\ AB}{divergence\ C} * F \qquad [1]$$

The intrinsic rate of nucleotide substitution or evolutionary rate, ER (% substitutions/million years) can be estimated for a particular lineage using the sequence divergence from a sister lineage with a reliable fossil date by the formula where *divergenceC* is the sequence

divergence between the test species and the living descendant of the nearest sister lineage

$$ER = \frac{\frac{1}{2} divergence\, C}{F} \qquad [2]$$

with a known fossil taxon and (F) is the estimated age of the fossil. The estimate may be a very rough approximation, as the actual time of separation will predate the appearance of a morphologically differentiated fossil by an unknown amount of time. If the ER for a lineage is known, divergence times (t) for taxa within the lineage that do not have suitable fossils can be calculated using the ER.

$$t_{AB} = \frac{\frac{1}{2} divergenceAB}{ER} \qquad [3]$$

2.14 Confidence

2.14.1 Decay indices

The Bremer support indices (Bremer, 1994) are a measure of the amount of support for each internal node in a parsimony phylogeny. The index is calculated from the difference in tree lengths between the shortest trees that contain a particular node and those trees that do not contain the node. Bremer decay indices can be calculated using Autodecay (written by T. Eriksson, 1997) or for multi-partition data, the decay indices can broken down according to support provided by individual data partitions in the program TreeRot v.2 (written by M.D. Sorenson, 1999).

2.14.2 Bootstrapping

Bootstrapping (Felsenstein, 1985) is a measure of whether a phylogenetic tree estimated from a set of nucleotide sequences is truly representative of the signal in the data or has been assembled to a particular topology by purely by chance. Homoplasy within the data will cause different nucleotide sites to support different tree topologies and therefore any given estimate of phylogeny will depend on the characters sampled, with an associated level of sampling error. Bootstrapping estimates sampling error by taking multiple random

samples from the sample (pseudoreplicates). By sampling with replacement, some sites will be represented in the sample more than once and others not at all. A tree is then built from the pseudoreplicate data set. The frequency of occurrence of each node amongst a large number of (100-2000) pseudoreplicate trees is a measure of the sampling error. A bootstrap value of >70% is considered to offer strong support for any particular node in a tree.

2.14.3 Hypothesis testing

When a particular hypothesis is proposed for a data set (eg. monophyly of a clade), the strength of the hypothesis can be statistically tested using the Kishino-Hasegawa (Kishino and Hasegawa, 1989) test for maximum likelihood and maximum parsimony analyses or the Templeton (1983) test for maximum parsimony analysis. These tests compare trees estimated under a constraint with those estimated under no constraint.

2.15 Mismatch distribution and raggedness index

Rogers & Harpending (1992) showed that there is a correlation between historical population size and the distribution of pair-wise nucleotide site differences between individuals. The distribution of nucleotide site differences between pairs of individuals (mismatch distribution, Hartl and Clark, 1997) changes as populations grow or decline (Avise *et al.*, 1988; Felsenstein, 1992; Rogers and Harpending, 1992). In a rapidly growing population, pair-wise differences have a Poisson distribution arising from most alleles being descended from a few ancestral types (Harpending *et al.*, 1993; Rogers and Harpending, 1992). In a theoretical stable or declining population, the equilibrium distribution shows a geometric decrease in the number of matched sites due to a much more structured allelic phylogeny. However, simulations show that stable or declining populations often have multi-modal distributions giving a ragged distribution that does not readily resemble the predicted outcome (Rogers and Harpending, 1992; Slatkin and Hudson, 1991). Bottlenecks in population size generate waves similar to those produced by sudden expansion but with a

steeper leading edge, and they often have ragged empirical distributions with many peaks at large values. Populations that have been through a rapid reduction in size produce L-shaped distributions due to the initial high probability of identity of nucleotide sites, while the right edge converges to a new equilibrium.

To obtain a measure of confidence as to how well the data fit a particular distribution curve, a null distribution is simulated for known parameters and compared to the raggedness index (Rogers, 1995) calculated from the data. High values of raggedness indicate that a population is in equilibrium.

Chapter 3.

Evolution of mitochondrial *control regions* in boid snakes

3.1 Introduction

The circular genome of the mitochondrion of vertebrates is in general approximately 16 kb and in most organisms is maternally inherited. The gene content of the vertebrate mitochondrial genome has been known since the early 1980's (Anderson *et al.*, 1982) and with a multitude of complete genomes now sequenced, has proved for the most part to be highly conserved across vertebrate lineages *i.e.* in amphibians (Roe *et al.*, 1985), birds (Desjardin and Morais, 1990), marsupials (Janke *et al.*, 1997), reptiles (Janke and Árnason, 1997) and sea lampreys (Lee and Kocher, 1995). However, the arrangement of those genes in the mitochondrial genome has proved to be much more dynamic (e.g. birds, Bensch and Härlid, 2000; Desjardin and Morais, 1990; Mindell *et al.*, 1998; marsupials, Pääbo *et al.*, 1991; reptiles, Quinn and Mindell, 1996; Seutin *et al.*, 1994) and gene order is being used as a phylogenetic character to determine evolutionary relationships between groups of organisms (Kumazawa and Nishida, 1993; Macey *et al.*, 1997). Even within the lineage of squamate reptiles, gene rearrangements are not uncommon. In lizards there are tandem duplications of clusters of genes (Moritz and Brown, 1986), tandem duplications of adjacent tRNAs (Macey *et al.*, 1998), and translocations of tRNAs (Kumazawa and Nishida, 1995; Macey *et al.*, 1997). Many of these gene rearrangements are associated with transfer RNAs (Macey *et al.*, 1998). Models implicating slipped-strand mis-pairing during strand replication (Levinson and Gutman, 1987; Streisinger *et al.*, 1966), and an association with hairpin-like structures (Macey *et al.*, 1997; Stanton *et al.*, 1994) have been proposed to explain these rearrangements.

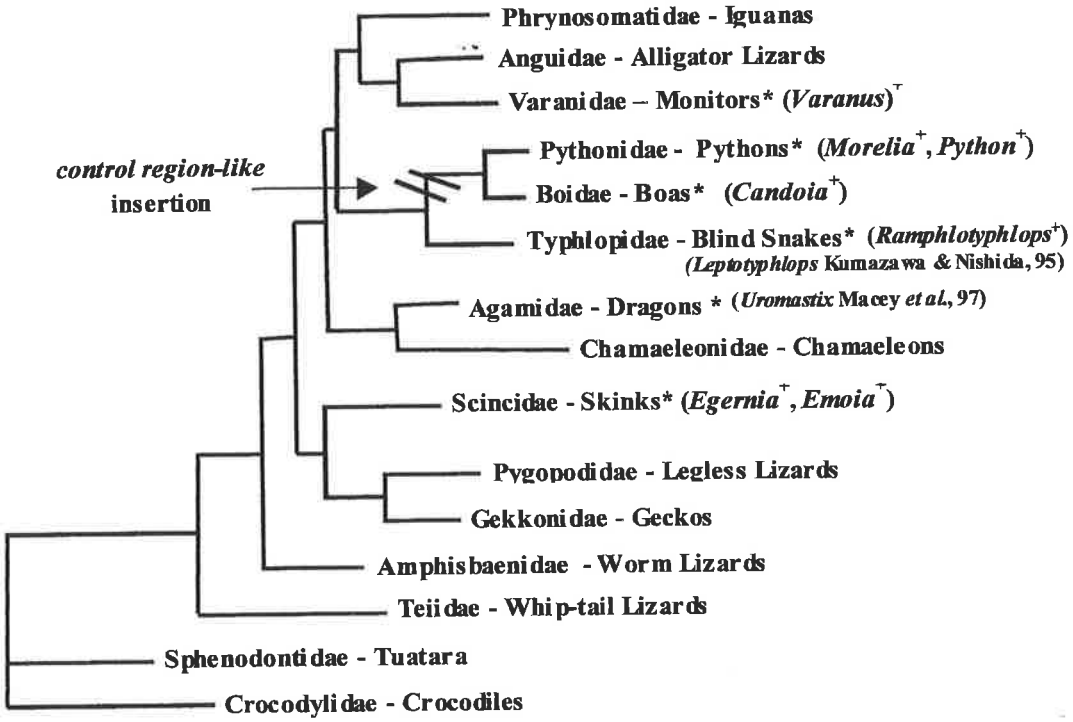
Whilst most non-tRNA duplications occurring in the mitochondrial genome are in tandem array (Moritz and Brown, 1987), in boid, colubrid and viperid snake lineages there is

a non-tandem duplication of the *control region* (*CR*) and associated tRNAs into the *isoleucine*, *glutamine* and *methionine (IQM) tRNA* cluster situated between the *ND1* and *ND2* genes (Kumazawa *et al.*, 1996; Kumazawa *et al.*, 1998). Studies examining the arrangement of gene in the *IQM tRNA* cluster in agamid and iguanid lizards (Macey *et al.*, 1997) and leptotyphlopid snakes (Kumazawa and Nishida, 1995) found no evidence of this inserted *control region-like* (*CRL*) gene.

In three individual snakes, Kumazawa *et al.* (1996) found that the sequences of the *CR* and the *CRL* were identical except for a single base substitution in each individual. Given the rapid rate of evolution in much of the *CR*, it is unlikely that sequence fidelity would be maintained between the *CR* and *CRL* without some kind of recombination mechanism. Until recently, there was no evidence of recombination between vertebrate mitochondrial genomes. Earlier studies showed recombinant mitochondrial sequences in yeast (Dujon, 1981) and more recently, putative recombinases have been found within mammalian mitochondria (Thyagarajan *et al.*, 1996).

The aim of this study was to extend the findings of Kumazawa *et al.* (1996) and investigate the maintenance of sequence fidelity between the *CR* and *CRL* in pythons at varying levels of evolutionary distance, *i.e.* between genera and at a population level within a species. If mitochondrial recombination is a rare event, sequence differences between the *CR* and *CRL* may be detected in a time lag between the sequence mutation in one region and the subsequent homogenization event. I would expect that for shorter divergence times (*i.e.* within or between conspecific populations) there would be more evidence of sequence differences between the *CR* and *CRL* as mutated sequences would have less chance of being involved in a recombination event that would homogenize the two regions.

To determine the evolutionary origin of the *CRL* insertion, I used PCR amplification to detect the presence/absence of a large (>1kb) insertion into the *IQM tRNA* cluster in a range of squamates (Fig. 3.1a and b). To test the idea of concerted evolution between the *CR* and *CRL*, firstly, the *CRs* of two species of *Morelia* pythons were examined for structural features



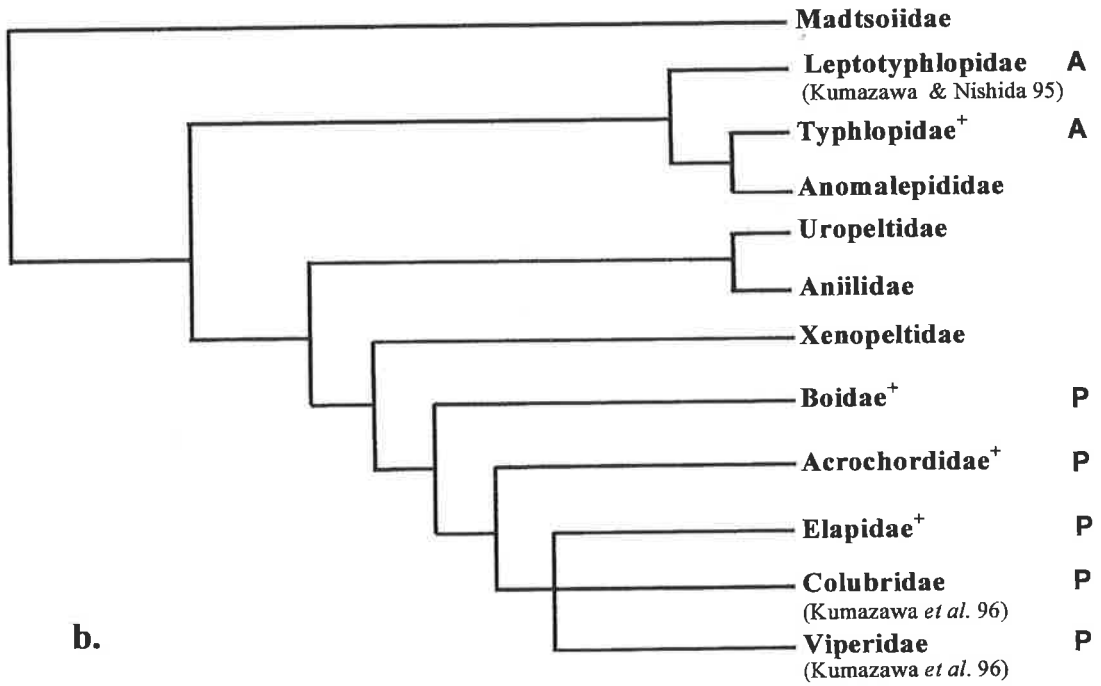


Fig. 3.1a. Evolutionary relationships of major reptile groups from *c-mos* gene sequence (adapted from Saint *et al.* 1998), showing the distribution of a *control region-like* insertion. * = lineages examined for *IQM* region, + = taxa examined in the present study, arrow indicates most likely insertion of *CRL* in squamate phylogeny. **b.** Evolutionary relationships of snakes (after Scanlon & Lee, 2000) showing the derivation of a *control region-like* insertion. + = taxa examined in the present study, P = insertion present, A = insertion absent.

that are known to be present in other vertebrates and colubrid and viperid snakes. Secondly, complete and partial *CR* and *CRL* sequences were compared from a number of individuals of the carpet python *Morelia spilota* of varying geographic and genetic distance (0-6% uncorrected sequence divergence) (Fig. 3.2) and from one representative each of the genus *Acrochordus* and three boid genera, *Antaresia*, *Candoia* and *Python*. Furthermore, to get a more comprehensive picture of the evolution of the *CR* in boid snakes, I also compared the primary and secondary structures of the *CR* and associated tRNAs for a range of python genera with those of other vertebrates.

3.2 Materials & Methods

3.2.1 Specimens examined

Individuals included in this study are given below. * = complete *control region* and *control region-like* sequences, $^{+}$ = extractions enriched for mitochondrial DNA.

Acanthophis antarcticus NTM R17880; *Acrochordus arafurae* SAMA R23999*$^{+}$; *Aipysurus laevis* NTM R17775; *Antaresia childreni* SAMA R21411*$^{+}$; *A. maculosa* ABTC 68227; *A. perthensis* ABTC 68276; *A. stimsoni* SAMA R38794; *Apodora papuana* ABTC 68240; *Aspidites melanocephalus* ABTC 68246; *A. ramsayi*$^{+}$ SAMA R19831; *Bothrochilus boa* AMS R129533; *Candoia aspera*$^{+}$ AMS R115337; *Demansia atra* SAMA R29954; *Egernia stokesii* SAMA R51503; *Emoia longicauda* AMS R122615; *Heloderma horridum* ABTC 64991; *Hydrelaps darwiniensis* NTMR16471; *Leiopython albertisii* AMS R124481^{+}; *Liasis mackloti* SAMA R21422; *Liasis fuscus* ABTC 68263; *Liasis olivaceus* ABTC 6503; *Morelia amethistina* AMS R115347^{+}; *M. boeleni* BPBM 11611; *M. oenpelliensis* ABTC 68277; ten *M. spilota* SAMA R26878*$^{+}$, ABTC 68278/79/80*/81/88, 68308*, ABTC 66175, ABTC 66267, ABTC 66295; six *M. viridis* ABTC 51498, ABTC 65592 (Australia), AMS R115348*$^{+}$, AMS R122363 (Southern PNG), BPBM 11617, BPBM 13798 (Northern PNG); *Python curtus* ABTC 24797; P. *reticulatus* SAMA R28533*; *P. sebae* SAMA R26137; *P. timoriensis* ABTC 68326; *Ramphotyphlops australis* SAMA R38446; *R. bituberculatus*

SAMA R38362; *Sphenodon punctata* ABTC 32244 and *Varanus rosenbergi* AMS R123331.

3.2.2 Mitochondrial genes

Genomic DNA was extracted and mitochondrial DNA (mtDNA) was enriched as per chapter 2.3. Because there is a duplication and translocation of the *CR* in the mitochondrial genome of snakes, as well as rearrangements of flanking transfer RNAs that vary between species (Kumazawa *et al.*, 1996), nested PCR as per chapter 2.8 was used to amplify partial *transfer RNA^{Thr}/CR* and *CRL* products using PCR2 (chapter 2.7.2) and program 2 (chapter 2.7.3.2). Primers L14973 and H690 were used to PCR amplify a large product spanning the *CR* and its flanking transfer RNAs. (For primer sequences see Table 2.1). This product was then used as a template for nested amplification with primer L15926 situated in the *tRNA^{Thr}* gene and H690. A nested PCR, using the first round product as template, and the primers L15926 and snake17 situated in the Conserved Sequence Block I (CSB-1) of the *CR*, generated PCR products of ~850 bp in length. To determine the source of size variation between *CR*s within species, a second nested PCR, using primers Snake16 and H690, was done from the *transfer RNA^{Thr}/CR* templates to amplify the 3' repeat region of some *Morelia* taxa.

To PCR amplify the *CRL* gene that is located between the *ND1* and *ND2* genes (Kumazawa *et al.*, 1996), a PCR product was amplified using primers L4160m and H4433 situated in the *tRNA^{Thr}* and *tRNA^{Gln}* respectively.

3.2.3 Sequencing and alignment of the CR and CRL

Both strands of the PCR products were sequenced as per chapter 2.10. Sequencing reactions in both directions of the *CR* and *CRL* products often stalled at a C-rich region described by Kumazawa *et al.* (1998). Additional nested primers were used to sequence across this region in each direction and to obtain sequence of both strands for other parts of the products. These primers were Snake7, Snake6 and Snake1. Secondary structures in the

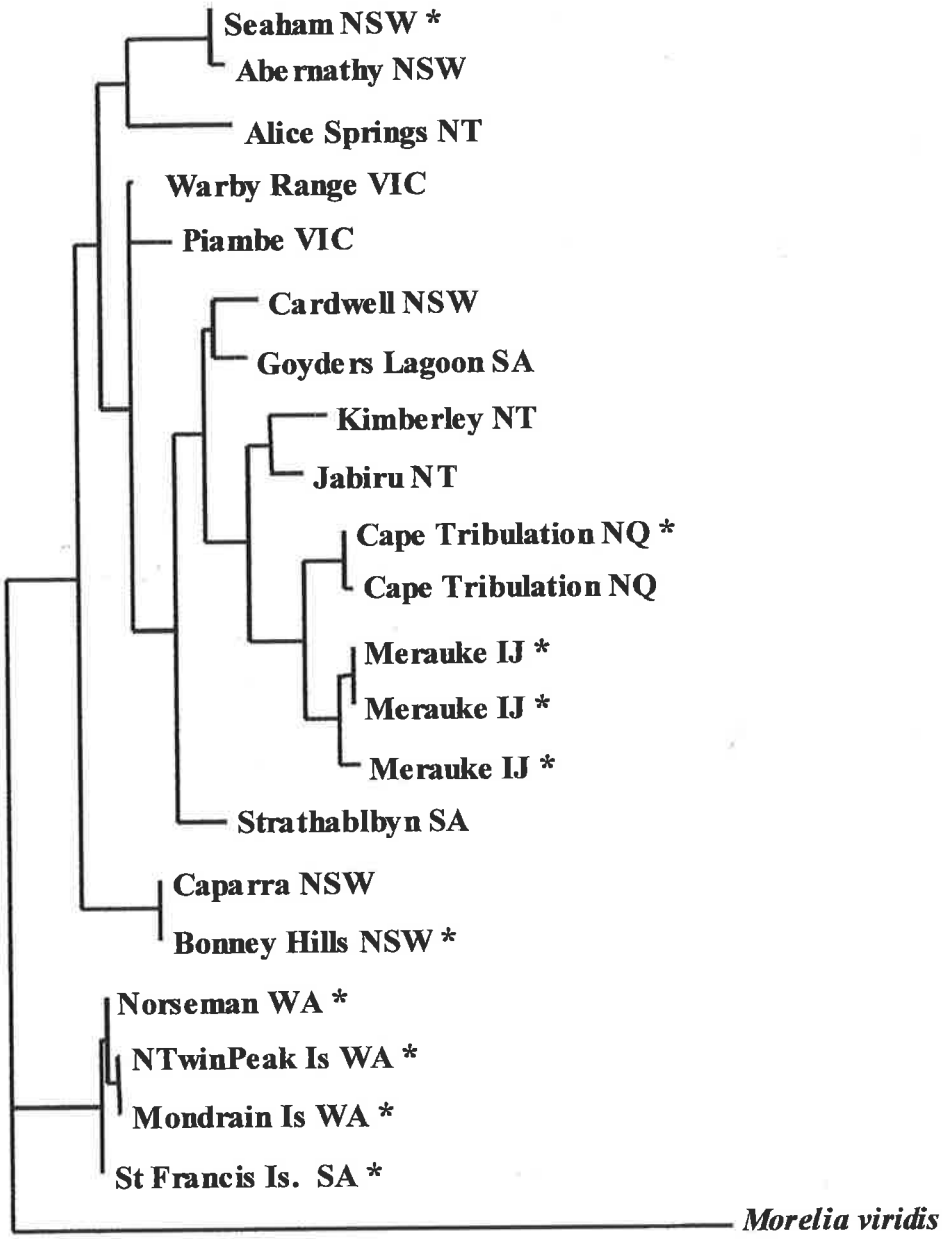


Fig. 3.2. Evolutionary relationships within *Morelia spilota* from 792 bp of mitochondrial *control region* sequence. * = taxa sequenced for *CR* and *CRL*, IJ = Irian Jaya, Indonesia; NSW = New South Wales; NT = Northern Territory; SA = South Australia; NQ = Northern Queensland; Vic = Victoria; WA = Western Australia

CR were identified using the ECGC Wisconsin Package version 8.1.0 (1996) available through WebAngis.

3.2.4 Paralogue test

The potential for the *CR* and *CRL* primers to amplify nuclear paralogues rather than mitochondrial genes (Zhang and Hewitt, 1996) was tested by a serial dilution PCR test as per chapter 2.4. Caesium-chloride gradient purified mtDNA and total cellular DNA of *Morelia spilota*, SAMA R26878 and *M. viridis*, AMS R115348 were used in the titration test. Primer H15916 situated in the $tRNA^{Thr}$ was also used to sequence the H-strand of the 3' end of the *cytb* gene of the large *CR* fragment to compare with published python sequence of this gene and to check for the absence of stop codons that would indicate a paralogue.

3.3 Results

3.3.1 Mitochondrially-derived sequences

Serially-diluted, enriched mtDNA from *Morelia spilota* and *M. viridis* were amplified to dilutions one thousand-fold greater with the *CR* and *CRL* primers than with the nuclear *18S rRNA* primers. These products were used as template DNA for a second round nested PCR. The second round products were then sequenced. These sequences were found to be indistinguishable from those obtained from total cellular DNA. Also, the amino acid sequence of the 5' *cytb* portion of the large *CR* product was the same as the published GenBank sequences for *M. spilota* and *M. viridis*.

3.3.2 Determining the presence of an insertion in the squamate IQM tRNA cluster

Genomic DNA was PCR amplified for the mitochondrial region between the *ND1* and *ND2* genes to determine the presence of a duplication of the *CR* (the *CRL*) in the *IQM tRNA* cluster in a variety of squamate taxa. Taxa surveyed included representatives of blind snakes, *Ramphotyphlops*; monitors, *Varanus*; skinks *Egernia* and *Emoia*; helodermatid lizards,

Heloderma and tuataras, *Sphenodon*. Representatives of the python genera *Antaresia*, *Aspidites*, *Leiopython*, *Morelia*, *Python*, the boid genus *Candoia*, elapid snakes *Acanthophis, Aipysurus, Demansia* and *Hydrelaps* and file snakes *Acrochordus* were also tested. Short ~300 bp products were amplified from *Egernia*, *Emoia*, *Heloderma* and *Varanus* and a shorter product was amplified from *Ramphotyphlops* (Fig. 3.3). In the blind snake, *Leptotyphlops,* the *IQM tRNA* region between *ND1* and *ND2* lacks the $tRNA^{Glu}$ (Kumazawa and Nishida, 1995). It is extrapolated that this is also the explanation for the shorter product amplified for *ND1-ND2* region in *Ramphlotyplops*. These shorter PCR products are consistent with a region that does not have the *CRL* insertion. Large products (1300->2000 bp) were amplified from *Antaresia*, *Leiopython*, *Morelia* and *Python* (Fig. 3.3) and larger products (~3.5 - 4kb) were amplified from *Acanthophis* and *Candoia aspera.*

Two PCR products of 1200 and 800 bp for *CR* and 1500 and 950 bp for *CR* and *CRL* were amplified for *Acrochordus arafurae* (Fig. 3.3 lanes 7 &13), see below. PCR amplification of the *CRL* region from the snakes *Aipysurus, Aspidites*, *Demansia* and *Hydrelaps*, and the lizard *Sphenodon* was not successful, most likely due a lack of sequence similarity of the primers to the target gene in these taxa.

3.3.3 CR *structure in pythons*

Extreme size variation was observed in PCR products of the mitochondrial *CR* from the $tRNA^{Thr}$ gene to the 5' end of *12S rRNA* gene for a range of python genera. Products ranged from ~ 1300-1500 bp for most taxa to >3500 bp in *Candoia aspera*. In order to locate the source of the *CR* size variation, nested primer pairs L15962/Snake 17 and Snake16/H690 were used to amplify the 5' and 3' ends of the *CR* respectively. The majority of the size variation was found in the 3' end of the *CR*, even within a species (Fig. 3.3).

The structural features of the *CR* were compared from sequenced PCR products spanning the $tRNA^{Thr}$ to *12S rRNA* region obtained from six individuals of *Morelia viridis* and three *M. spilota* (Figs. 3.4 and 3.5). For these individuals, the PCR products ranged in size

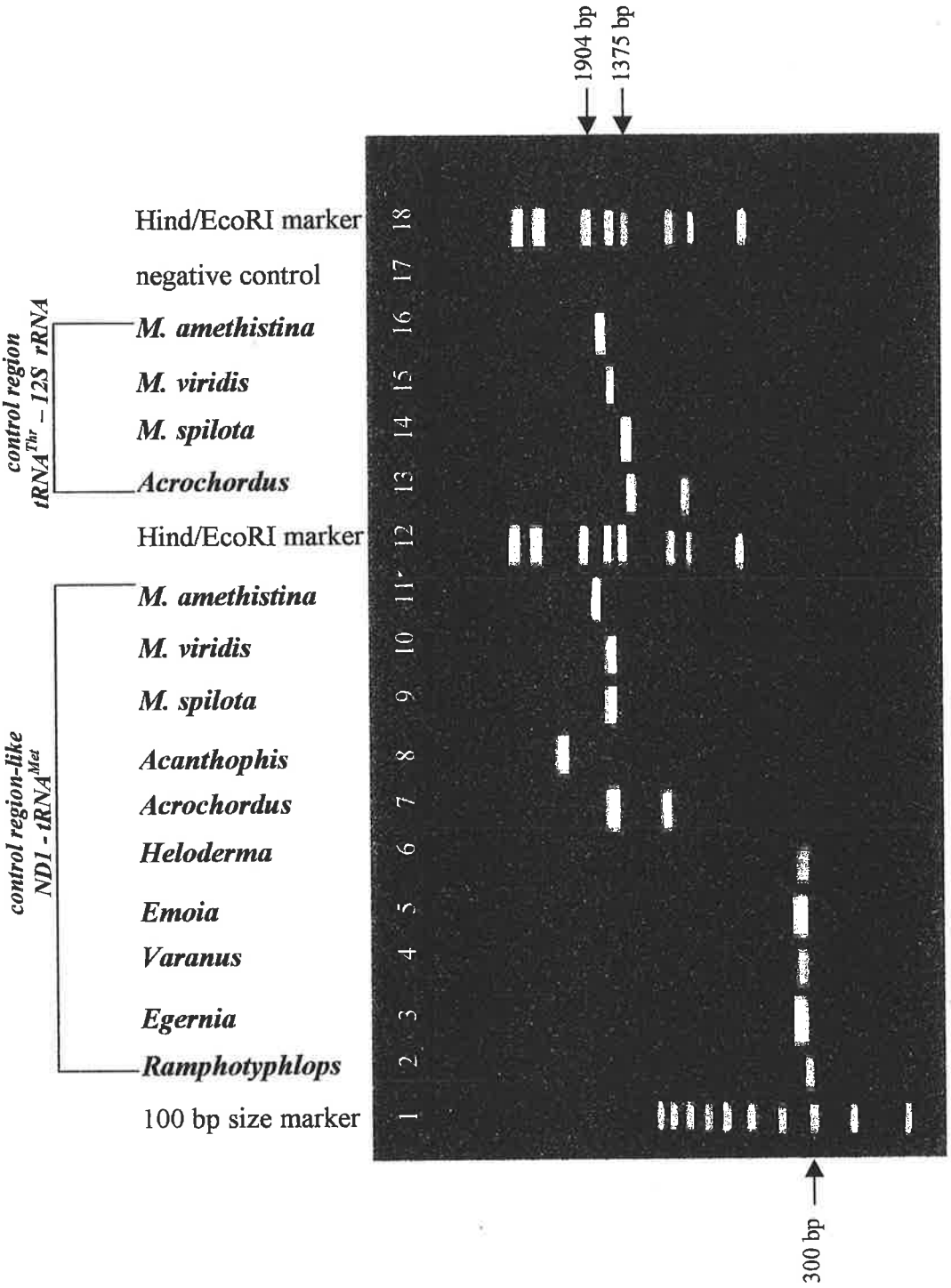


Fig. 3.3 PCR amplified products for the *control region-like* fragment for a range of squamate taxa, showing the absence of a large insertion in the blind snake (*Ramphotyphlops*) and lizards (lanes 2-6) and its presence in other snake lineages (lanes 7-11). Lanes 13-16 show PCR products for *control region* showing size variation. Lanes 7 & 13 show two products amplified for *Acrochordus arafurae*.

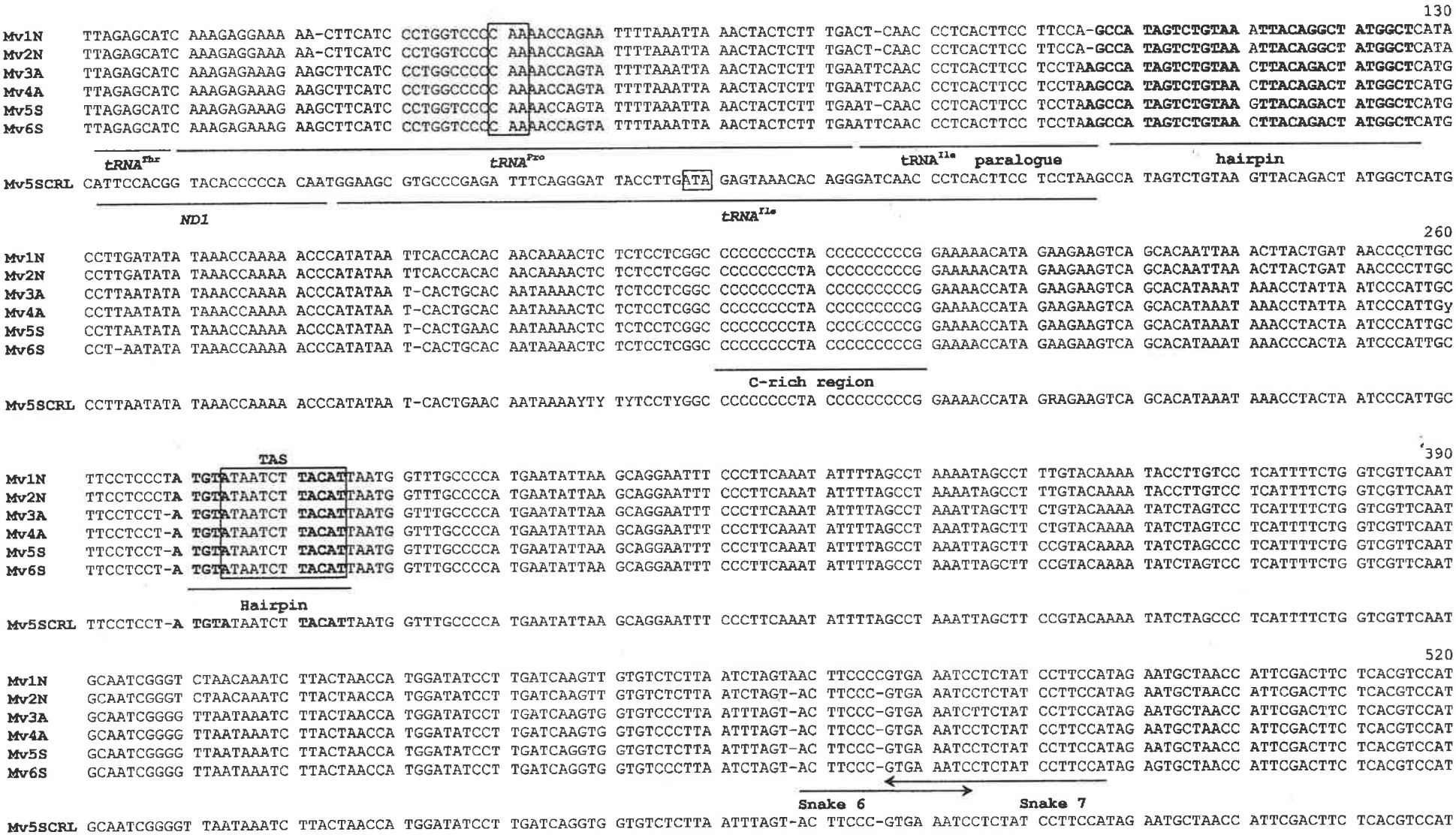
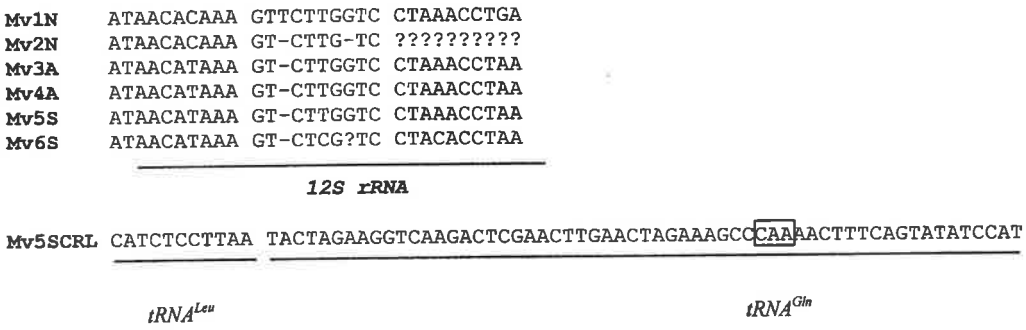


Fig.3.4. Nucleotide sequence alignments of complete *control region* for six *Morelia viridis* individuals from three populations showing secondary structures. **Mv1**=BPBM 11617; **Mv2**=BPBM 13798; **Mv3**=ABTC 65592; **Mv4**=ABTC 51498; **Mv5**=AMS R115348; **Mv6**=AMS R122363; N=northern PNG; S=southern PNG; A=Australia. A complete *control region-like* sequence is also shown for comparison.

```
                                                                                                                         650
Mv1N     ATAATGCCAA TCCTCCCTTT AAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA
Mv2N     ATAATGCCAA TCCTCCCTTT AAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA
Mv3A     ATTATGCCAA TCCTCCCTTC TAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA
Mv4A     ATTATGCCAA TCCTCCCTTC TAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA
Mv5S     ATTATGTCAA TCCTCCCTTC TAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA
Mv6S     ATTATGTCAA TCCTCCCTTC TAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA

Mv5SCRL  ATTATGTCAA TCCTCCCTTC TAGCTTTTTC CAAGGCCGCT GGTTACACTC TCAAGATCAT CTCAATGGTC CGGAACCACC CCTCCATCCT AGCTTTTTCC AAGACCTTTG GTCGCACCCT TTATATGGTA

                                                                                                                         780
Mv1N     CATCTTGTCT CATGTTCTGA TCACCTATGT TAGTCCGCCA CTGGTTTCCC TTTTTTTCTC TGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCT-TAC GGTCCGGGTG GAACATTATG
Mv2N     CATCTTGTCT CATGTTCTGA TCACCTATGT TAGTCCGCCA CTGGTTTCCC TTTTTTTCTC TGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCT-TAC GGTCCGGGTG GAACATTATG
Mv3A     CATCTTGTCT CATGTTCTGA TCACCTATGC TAGTCCACCC CTGGT-TCCC TTTTTTTCTC GGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCTCTTT AGTCCGGGTG GAACATGATG
Mv4A     CATCTTGTCT CATGTTCTGA TCACCTATGC TAGTCCACCC CTGGTT-CCC TTTTTTTCTC GGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCTCTTT AGTCCGGGTG GAACATGATG
Mv5S     CATCTTGTCT CATGTTCTGA TCACCTATGC TAGTCCACCA CTGGTTTCCC TTTTTTTCTC GGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCTCTTT AGTCCGGGTG GAACATGATG
Mv6S     CATCTTGTCT CATGTTCTGA TCACCTATGC TAGTCCACCA CTGGTTTCCC TTTTTTTCTC GGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCTCTTT AGTCCGGGTG GAACATGATG

Mv5SCRL  CATCTTGTCT CATGTTCTGA TCACCTATGC TAGTCCACCA CTGGTTTCCC TTTTTTTCTC GGTACCTTTC ATCTGACTAC CATATATGCA CACACACAGT AAGGCTCTTT AGTCCGGGTG GAACATGATG

                                                                                                                         910
Mv1N     TTTTCTCCCC ATTTACCCCC ATGGATACA- TCTCTTAATG CTTGTTAGAC ATATGAATTC TCCTTACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCA-TT AAATAAACAA ACCACAAC--
Mv2N     TTTTCTCCCC ATTTACCCCC ATGGATACA- TCTCTTAATG CTTGTTAGAC ATATGAATTC TCCTTACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCA-TT AAATAAACAA ACCACAAC--
Mv3A     TTACTTCCTC ATTTACCCCA ATGGATACAG TTTCTTAATG CTTGTTAGAC ATATTAATTC TCCTAACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCAACG ATAAAAACAA ACCCAAACCC
Mv4A     TTACTTCCTC ATTTACCCCA ATGGATACAG TTTCTTAATG CTTGTTAGAC ATATTAATTC TCCTAACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCAACG ATAAAAACAA ACCCAAACCC
Mv5S     TTACTTCCTC ATTTACCCCA ATGGATACAG TTTCTTAATG CTTGTCAGAC ATATTAATTC TCCTAACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCAACG ATAAAAACAA ACCCAAACCC
Mv6S     TTACTTCCTC ATTTACCCCA ATGGATACAG TTTCTTAATG CTTGTCAGAC ATATTAATTC TCCTAACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCAACG ATAAAAACAA ACCCAAACCC
                                             __________________________                                  ________________
                                            <-------CBS-1----------------                                     hairpin
                                   Snake 16 ---------------------------->  Snake 17
Mv5SCRL  TTACTTCCTC ATTTACCCCA ATGGATACAG TTTCTTAATG CTTGTCAGAC ATATTAATTC TCCTAACCGA AATTTCATAT ACCGAAACTT CGGAAAAACC CCAATCAACG ATAAAAACAA ACCCAAACCC

                                                                                                                         1040
Mv1N     ---------TAT TTCTATCACC TTCCAACACG CCCTGAGCGC AATTATACAC AAACTGGCTT TTATTATTAA CAGGCCCATC GTTTCTGTAT AATTTCTTTT AAATCACCTT CCAACACGCC CTGGGCGG
Mv2N     ---------TAT TTCTATCACC TTCCAACACG CCCTGAGCGC AATTATACAC AAACTGGCTT TTATTATTAA CAGGCCCATC GTTTCTGTAT AATTTCTTTT AAATCACCTT CCAACACGCC CTGAGCGC
Mv3A     ACACAAC-ATTT TTGAATCACC TTTTTACACG CTCCTGGTGT AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGT
Mv4A     ACACAAC-ATTT TTGAATCACC TTTTTACACG CTCCTGGTGT AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGT
Mv5S     ACACAACCATTT TTGAATCACC TTTTTACACA CTCCTGGTGT AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGT
Mv6S     ACACAAC-ATTT TTGAATCACC TTTTTACACA CTCCTGGTGT AATTATACAC AAATTGGCTT TTAATATTAG CAGGCCCATC GTTTCTGTAT AATTTTCTT- AGATCACCTT TTTACACGCT CCTGGTGT

Mv5SCRL  ACACAACCATTT TTGAATCACC TTTTTACACA CTCCTGGTGT AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGT
```


1170
Mv1N AATTATACAC AAACTGGCTT TTATTATTAA CAGGCCCATC GTTTCTGTAT AATTTCTTTT AAATCACCTT CCAACACGCC CTGGGCGGAA TTATACACAA ACTGGCTTTT ATTATTAA-C AGGCCCATCG
Mv2N AATTATACAC AAACTGGCTT TTATTATTAA CAGGCCCATC GTTTCTGTAT AATTTCTTTT AAATCACCTT CCAACACGCC CTGAGCGCAA TTATACACAA ACTGGCTTTT ATTATTAA-C AGGCCCATCG
Mv3A AATTATACAC AAATTGGCTT TAAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGTAA TTATACACAA ATTGGCTTTT AATATTAA-C AGGCCCATCG
Mv4A AATTATACAC AAATTGGCTT TAAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGTAA TTATACACAA ATTGGCTTTT AATATTAAAC AGGCCCATCG
Mv5S AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGTAA TTATACACAA ATTGGCTTTT AATATTAA-C AGGCCCATCG
Mv6S AATTATACAC AAATTGGCTT T-AATATTAG CAGGCCCATC GTTTCTGTAT AATTTTCTT- AGATCACCTT TTTACACGCT CCTGGTGTAA TTATACACAA ATTGGCTTTT AATATTAG-C AGGCCCATCG
Mv5SCRL AATTATACAC AAATTGGCTT TTAATATTAA CAGGCCCATC GTTTCTGTAT AATTTTCTT- AAATCACCTT TTTACACGCT CCTGGTGTAA TTATACACAA ATTGGCTTTT AATATTAA-C AGGCCCATCG
1300
Mv1N TTTCTGTATA ATTTCTTTTA AATCACCTTC CAACACGCCC TGAGCGCAAT TATATACAAA CTGGCTTTTA TTATTAACAG GCCCATCGTT TCTGTATAAT TTCTTTTAAA TCACCTTCCA ACACGCCCTG
Mv2N TTTCTGTATA ATTTCTTTTA AATCACCTTC CAACACGCCC TGGGCGGAAT TATACACAAA CTGGCTTTTA TTATTAACAG GCCCATCGTT TCTGTATAAT TTCTTTTAAA TCACCTTCCA ACACGCCCTG
Mv3A TTTCTGTATA ATTTTCTT-A AATCACCTTT TTACACGCTC CTGGTGTAAT TATACACAAA TTGGCTTTTA ATATTAACAG GCCCATCGTT TCTGTATAAT TTTCTT-AAA TCACCTTTTT ACACGCTCCT
Mv4A TTTCTGTATA ATTTTCTT-A AATCACCTTT TTACACGCTC CTGGTGTAAT TATACACAAA TTGGCTTTTA ATATTAACAG GCCCATCGTT TCTGTATAAT TTTCTT-AAA TCACCTTTTT ACACGCTCCT
Mv5S TTTCTGTATA ATTTTCTT-A AATCACCTTT TTACACGCTC CTGGTGTAAT TATACACAAA TTGGCTTTTA ATATTAACAG GCCCATCGTT TCTGTATAAT TTTCTT-AAA TCACCTTTTT ACACGCTCCT
Mv6S TTTCTGTATA ATTTTCTT-A GATCACCTTT TTACACGCTC CTGGTGTAAT TATACACAAA TTGGCTTTTA ATATTAGCAG GCCCATCGTT TCTGTATAAT TTTCTT-AGA TCACCTTTTT ACACGCTCCT
Mv5SCRL TTTCTGTATA ATTTTC---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------- ---------
1430
Mv1N AGCGCAATTA TATACAAACT GGCTTTTATT ATTAACAGGC CCATCGTTTC TGTATAATTT CTTTTAAATC ACCTTCCAAC ACGCCCTGGG CGGAATTAT- ---------- ---------- ----------
Mv2N AGCGCAATTA TACACAAACT GGCTTTTATT ATTAACAGGC CCATCGTTTC TGTATAATTT CTTTTAAATC ACCTTCCAAC ACGCCCTGGG CGGAATTAT- ---------- ---------- ----------
Mv3A AGTGTAATTA TACACAAATT GGCTTTTAAT ATTAACAGGC CCATCGTTTC TGTATAATTT TCTT-AAATC ACCTTTTTAC AC-------- ---------- ---------- ----------
Mv4A AGTGTAATTA TACACAAATT GGCTTTTAAT ATTAACAGGC CCATCGTTTC TGTATAATTT TCTT-AAATC ACCTTTTTAC ACGCTCCTGG TGTAATTATA CACAAATTGG CTTTTAATAT TAACAGGCCC
Mv5S AGTGTAATTA TACACAAAT- GGCTTTTAAT ATTAACAGGC CCGTA----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Mv6S AGTGTAATTA TACACAAATT GGCTTTTAAT ATTAGCAGGC CCATCGTTTC TGTATAATTT TCTT-AGATC ACCTTTTTAC AC-------- ---------- ---------- ---------- ----------
Mv5SCRL ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------- ---------- ----------
1560
Mv1N ---------- ---------- ---------- ---------- ---------- ------TACA GACCCACTGT TATTGTAGCT TACACTA-CA AAGCACAGCA CTGAAAATGC TGAGACGGTA CTACACCCAA
Mv2N ---------- ---------- ---------- ---------- ---------- ------TACA GACCCACTGT TATTGTAGCT TACACTA-CA AAGCACAGCA CTGAAAATGC TGAGACGGTA CTACACCCAA
Mv3A ---------- ---------- ---------- ---------A AATGGCTTTT AATAATAACA AACCCATCGT TTTTGTAGCT TACACTATCA AAGCACAGCA CTGAAAATGC TGAGACGGTA GCATACCCAA
Mv4A ATCGTTTCTG TATAATTTCT TTAAATCACC TTTTTACACA AATGGCTTTT AATAATAACA AACCCATCGT TTTTGTAGCT TACACTATCA AAGCACAGCA CTGAAAATGC TGAGACGGTA GCATACCCAA
Mv5S ---------- ---------- ---------- ---------- ---------- ---------- --------GT TTTTGTAGCT TACACTATCA AAGCACAGCA CTGAAAATGC TGAGACGGTA GCATACCCAA
Mv6S ---------- ---------- ---------- --------AA ATTGGCTTTT AATATTAACA GACCCATCGT TTTTGTAGCT TACACTATCA AAGCACAGCA CTGAAAATGC TGAGACGGTA GCATACCCAA
tRNA^Phe
Mv5SCRL ---------- ---------- ---------- ---------- ---------- ---------- ------ATCC ATTAAGGTAG CAAAGTCAGG CCATGCAAGA GGCTTAAAAC CTTGACACAG ATGTTCAAAT
tRNA^Leu


Mv1N ATAACACAAA GTTCTTGGTC CTAAACCTGA
Mv2N ATAACACAAA GT-CTTG-TC ??????????
Mv3A ATAACATAAA GT-CTTGGTC CTAAACCTAA
Mv4A ATAACATAAA GT-CTTGGTC CTAAACCTAA
Mv5S ATAACATAAA GT-CTTGGTC CTAAACCTAA
Mv6S ATAACATAAA GT-CTCG?TC CTACACCTAA
12S rRNA
Mv5SCRL CATCTCCTTAA TACTAGAAGGTCAAGACTCGAACTTGAACTAGAAAGCCCAAAACTTTCAGTATATCCAT
tRNALeu
tRNAGln

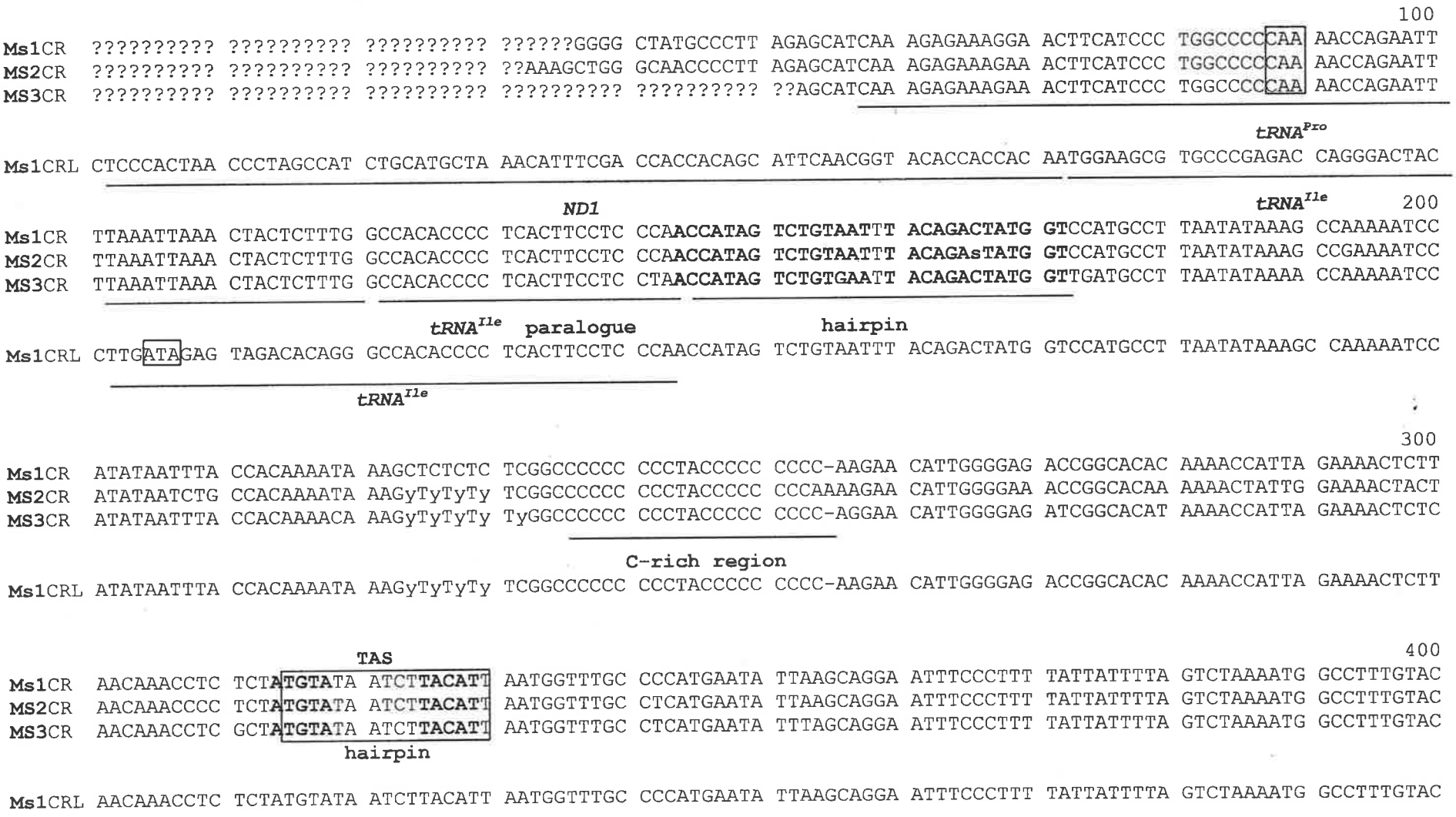
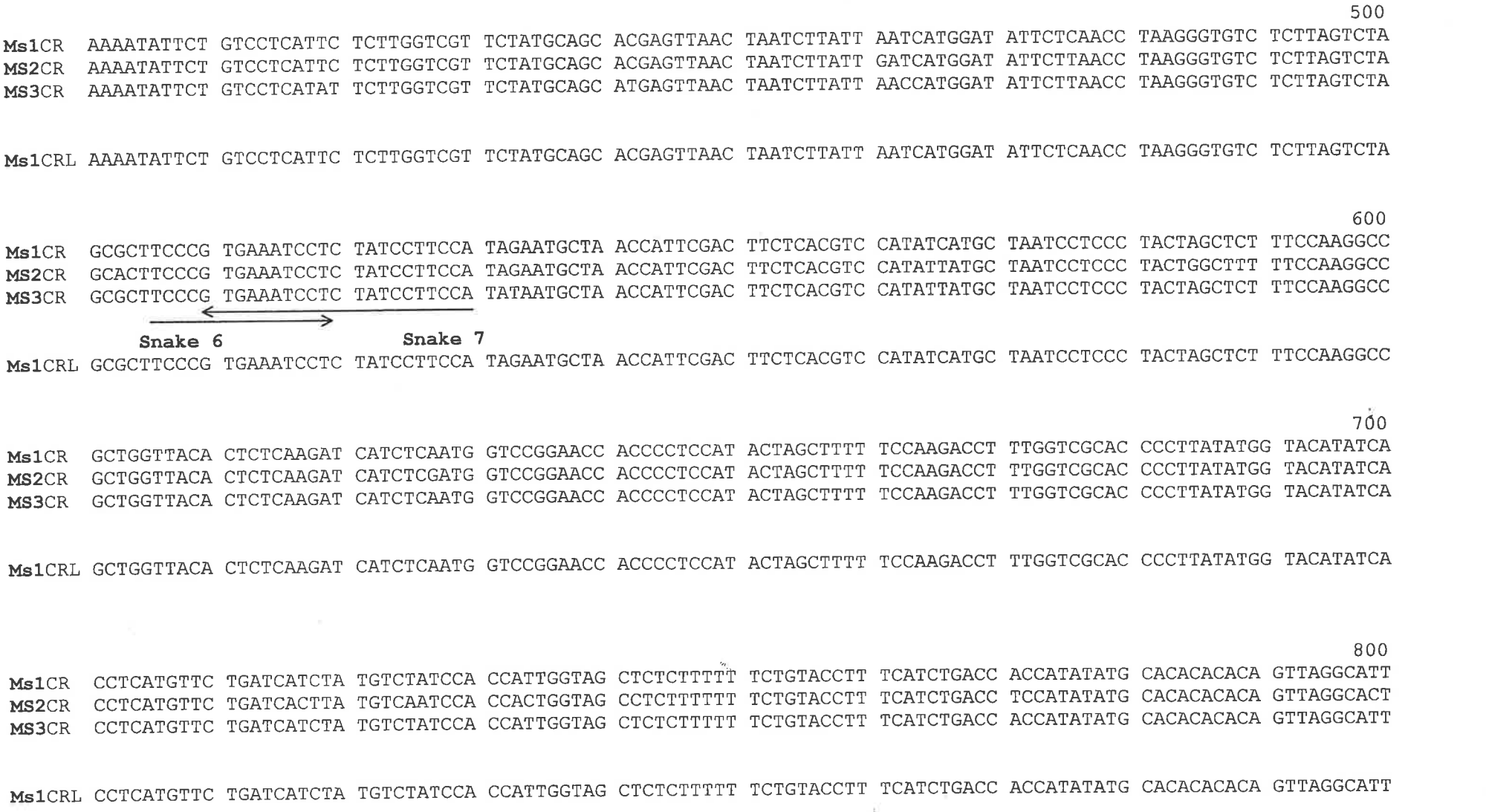
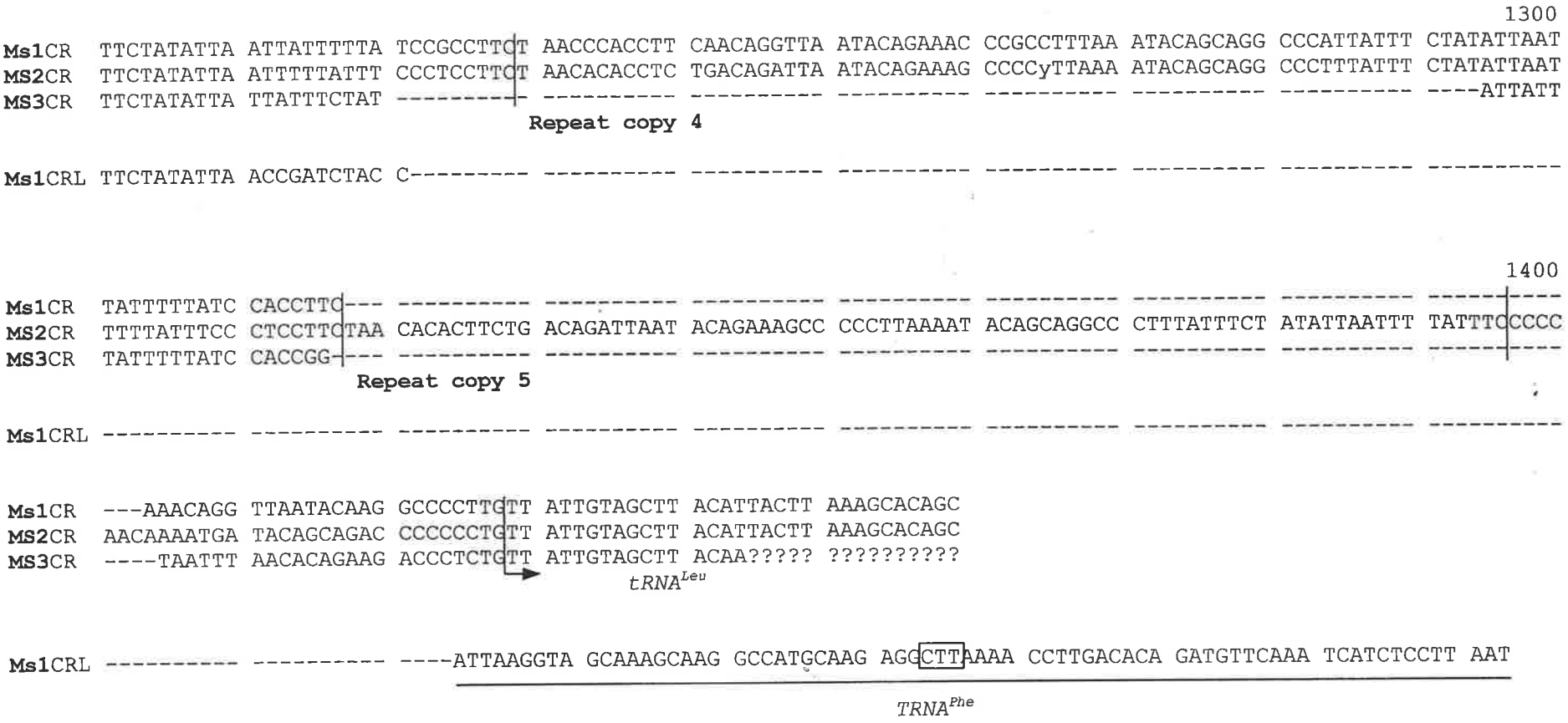


Fig.3.5. Nucleotide sequence alignments of complete *control region* for three *Morelia spilota* individuals from three populations showing secondary structures. A complete *control region-like* sequence is also shown for comparison. Ms1=SAMA R26878; Ms2=ABTC 68308; Ms3=ABTC 68280.

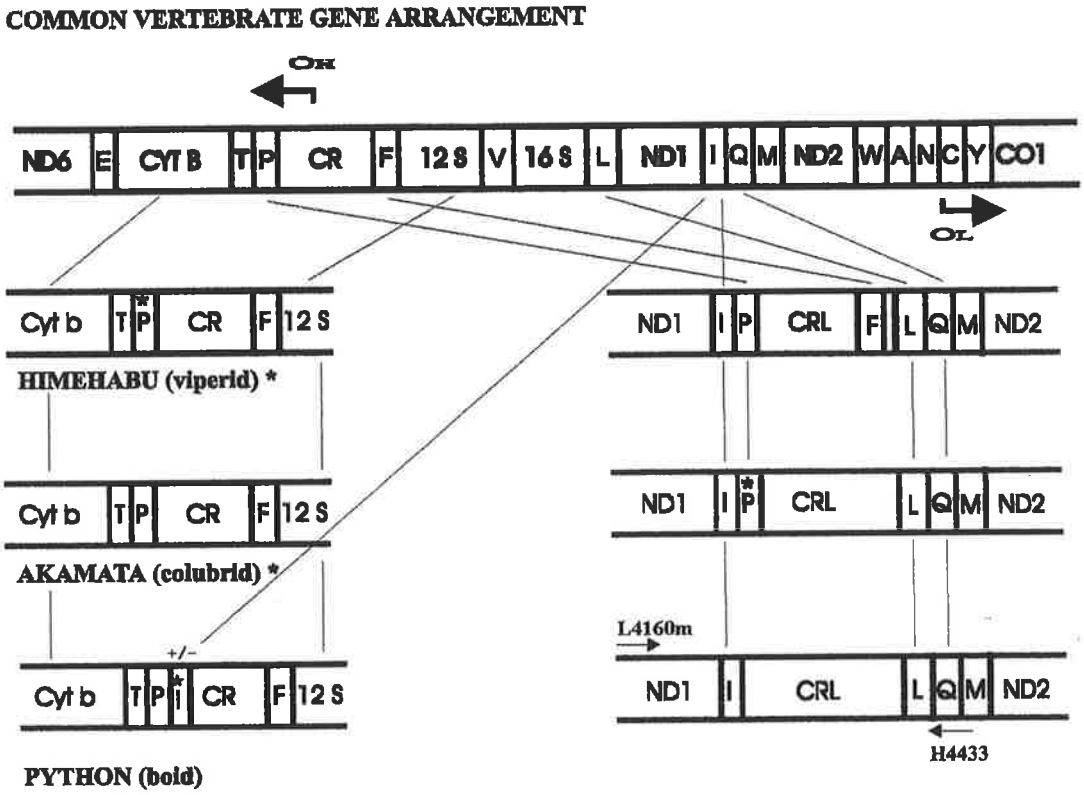

500
Ms1CR AAAATATTCT GTCCTCATTC TCTTGGTCGT TCTATGCAGC ACGAGTTAAC TAATCTTATT AATCATGGAT ATTCTCAACC TAAGGGTGTC TCTTAGTCTA
MS2CR AAAATATTCT GTCCTCATTC TCTTGGTCGT TCTATGCAGC ACGAGTTAAC TAATCTTATT GATCATGGAT ATTCTTAACC TAAGGGTGTC TCTTAGTCTA
MS3CR AAAATATTCT GTCCTCATAT TCTTGGTCGT TCTATGCAGC ATGAGTTAAC TAATCTTATT AACCATGGAT ATTCTTAACC TAAGGGTGTC TCTTAGTCTA
Ms1CRL AAAATATTCT GTCCTCATTC TCTTGGTCGT TCTATGCAGC ACGAGTTAAC TAATCTTATT AATCATGGAT ATTCTCAACC TAAGGGTGTC TCTTAGTCTA
600
Ms1CR GCGCTTCCCG TGAAATCCTC TATCCTTCCA TAGAATGCTA ACCATTCGAC TTCTCACGTC CATATCATGC TAATCCTCCC TACTAGCTCT TTCCAAGGCC
MS2CR GCACTTCCCG TGAAATCCTC TATCCTTCCA TAGAATGCTA ACCATTCGAC TTCTCACGTC CATATTATGC TAATCCTCCC TACTGGCTTT TTCCAAGGCC
MS3CR GCGCTTCCCG TGAAATCCTC TATCCTTCCA TATAATGCTA ACCATTCGAC TTCTCACGTC CATATTATGC TAATCCTCCC TACTAGCTCT TTCCAAGGCC
Snake 6
Snake 7
Ms1CRL GCGCTTCCCG TGAAATCCTC TATCCTTCCA TAGAATGCTA ACCATTCGAC TTCTCACGTC CATATCATGC TAATCCTCCC TACTAGCTCT TTCCAAGGCC
700
Ms1CR GCTGGTTACA CTCTCAAGAT CATCTCAATG GTCCGGAACC ACCCCTCCAT ACTAGCTTTT TCCAAGACCT TTGGTCGCAC CCCTTATATGG TACATATCA
MS2CR GCTGGTTACA CTCTCAAGAT CATCTCGATG GTCCGGAACC ACCCCTCCAT ACTAGCTTTT TCCAAGACCT TTGGTCGCAC CCCTTATATGG TACATATCA
MS3CR GCTGGTTACA CTCTCAAGAT CATCTCAATG GTCCGGAACC ACCCCTCCAT ACTAGCTTTT TCCAAGACCT TTGGTCGCAC CCCTTATATGG TACATATCA
Ms1CRL GCTGGTTACA CTCTCAAGAT CATCTCAATG GTCCGGAACC ACCCCTCCAT ACTAGCTTTT TCCAAGACCT TTGGTCGCAC CCCTTATATGG TACATATCA
800
Ms1CR CCTCATGTTC TGATCATCTA TGTCTATCCA CCATTGGTAG CTCTCTTTTT TCTGTACCTT TCATCTGACC ACCATATATG CACACACACA GTTAGGCATT
MS2CR CCTCATGTTC TGATCACTTA TGTCAATCCA CCACTGGTAG CCTCTTTTTT TCTGTACCTT TCATCTGACC TCCATATATG CACACACACA GTTAGGCACT
MS3CR CCTCATGTTC TGATCATCTA TGTCTATCCA CCATTGGTAG CTCTCTTTTT TCTGTACCTT TCATCTGACC ACCATATATG CACACACACA GTTAGGCATT
Ms1CRL CCTCATGTTC TGATCATCTA TGTCTATCCA CCATTGGTAG CTCTCTTTTT TCTGTACCTT TCATCTGACC ACCATATATG CACACACACA GTTAGGCATT

```
                                                                                                                    900
Ms1CR   CAGTCCGGGT TGAGCACGCT GCTTCCTCCT CATTTACCCC TAGGGATACA TCTCTTAATG CTTGTTAGAC ATATGTATTC TCCTTCCCGA AATTTCATAT
MS2CR   TAGTCCAGGT TGAGCACGCT GTTTCCTCCC CATTTACCCC TATGGATACA TCTCTTAATG CTTGTTAGAC ATATGTATTC TCCTTCCCGA AATTTCATAT
MS3CR   CAGTCCGGGT TGAGCACGCT GCTTCCTCCT CATTTACCCC TAGGGATACA TCTCTTAATG CTTGTTAGAC ATATGTATTC TCCTTCCCGA AATTTCATAT
                                                                                  ----------
                                                                                    CSB-1
Ms1CRL  CAGTCCGGGT TGAGCACGCT GCTTCCTCCT CATTTACCCC TAGGGATACA TCTCTTAATG CTTGTTAGAC ATATGTATTC TCCTTCCCGA AATTTCATAT

                                                                                                                   1000
Ms1CR   ACCAAAACTT CGGAAAAACC CCAATCCTGA CAAAAATCTC ACACAACTTT TT-ACTTTTT CCTCTAACCC AC-CTTCAAC AGGTTAATAC AGAAACCCGC
MS2CR   ACCGAAACTT CGGAAAAACC CCAATCCTTT CAACAACCTC CCCACAGATT TT-CATTTTT CCTCTAACAC AC-CTCTGAC AGATTAATAC AGAAAGCCCC
MS3CR   ACCGAAACTT CGGAAAAAAC CCAATCCTGA CAAAA-CCTC ACACAACTTT TTTCCTTTTT CCTCTAACCC TCTCTCCAAG AGGTTAATAC AAAAACCCGC
        ----------                                                                                  Repeat copy 1
          hairpin
Ms1CRL  ACCAAAACTT CGGAAAAACC CCAATCCTGA CAAAAATCTC ACACAACTTT TT-ACTTTTT CCTCTAACCC AC-CTTCAAC AGGTTAATAC AGAAACCCGC

                                                                                                                   1100
Ms1CR   CTTTAAATAC AGCAGGCCCA TTATTTCTAT ATTAATTATT TTTATCCGCC TTCTAACCCA CCTTCAACAG GTTAATACAG AAACCCGCCT TTAAATACAG
MS2CR   CTTAAAATAC AGCAGGCCCT TTATTTCTAT ATTAATTTTT ATTTCCCTCC TTCTAACACA CCTCTGACAG ATTAATACAG AAAGCCCCCT TAAAATACAG
MS3CR   CTTTAAACAC AGCAGGCCCA TTATTTCTAT ATTATTTATT TTTATCCGCC GTCTAAACCC TCTCCAAGAG GTTAATACAA AAACCCGCCT TTAAACACAG
                                                                                  Repeat copy 2
Ms1CRL  CTTTAAATAC AGCAGGCCCA TTATTTCTAT ATTAATTATT TTTATCCGCC TTCTAACCCA CCTTCAACAG GTTAATACAG AAACCCGCCT TTAAATACAG

                                                                                                                   1200
Ms1CR   CAGGCCCATT ATTTCTATAT TAATTATTTT TATCCGCCTT CTAACCCACC TTCAACAGGT TAATACAGAA ACCCGCCTTT AAATACAGCA GGCCCATTAT
MS2CR   CAGGCCCTTT ATTTCTATAT TAATTTTTAT TTCCCTCCTT CTAACACACC TCTGACAGAT TAATACAGAA AGTCCCCTTA AAATACAGCA GGCCCTTTAT
MS3CR   CAGGCCCATT ATTTCTATAT TATTTATTTT TATCCGCCGT CTAAACCCTC TCCAAGAGGT TAATACAAAA ACCCGCCTTT AAACACAGCA GGCCCATTAT
                                                                          Repeat copy 3
Ms1CRL  CAGGCCCATT ATTTCTATAT TAATTATTTT TATCCGCCTT CTAACCCACC TTCAACAGGT TAATACAGAA ACCCGCCTTT AAATACAGCA GGCCCATTAT
```

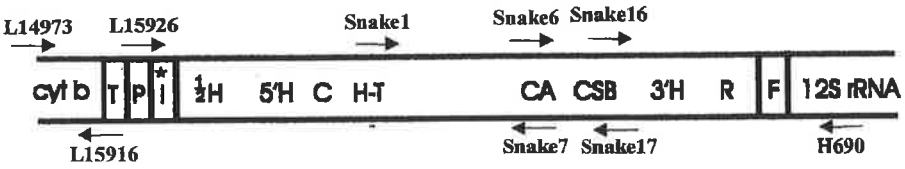

1300
Ms1CR TTCTATATTA ATTATTTTTA TCCGCCTTCT AACCCACCTT CAACAGGTTA ATACAGAAAC CCGCCTTTAA ATACAGCAGG CCCATTATTT CTATATTAAT
MS2CR TTCTATATTA ATTTTTATTT CCCTCCTTCT AACACACCTC TGACAGATTA ATACAGAAAG CCCCyTTAAA ATACAGCAGG CCCTTTATTT CTATATTAAT
MS3CR TTCTATATTA TTATTTCTAT ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----ATTATT
Repeat copy 4
Ms1CRL TTCTATATTA ACCGATCTAC C--------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
1400
Ms1CR TATTTTTATC CACCTTC--- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------
MS2CR TTTTATTTCC CTCCTTCTAA CACACTTCTG ACAGATTAAT ACAGAAAGCC CCCTTAAAAT ACAGCAGGCC CTTTATTTCT ATATTAATTT TATTTCCCCC
MS3CR TATTTTTATC CACCGG---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------
Repeat copy 5
Ms1CRL ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Ms1CR ---AAACAGG TTAATACAAG GCCCCTTGTT ATTGTAGCTT ACATTACTT AAAGCACAGC
MS2CR AACAAAATGA TACAGCAGAC CCCCCCTGTT ATTGTAGCTT ACATTACTT AAAGCACAGC
MS3CR ----TAATTT AACACAGAAG ACCCTCTGTT ATTGTAGCTT ACAA????? ??????????
tRNALeu
Ms1CRL ---------- ---------- ----ATTAAGGTA GCAAAGCAAG GCCATGCAAG AGGCTTAAAA CCTTGACACA GATGTTCAAA TCATCTCCTT AAT
TRNAPhe

from approximately 1450 to 1650 bp. PCR products were sequenced using a series of nested primers to obtain over-lapping sequences. Examination of the sequences showed a partial $tRNA^{Thr}$ gene 3' to the priming site, and a functional $tRNA^{Pro}$ gene, directly followed by 20 bp of the amino acid acceptor stem and the TψC arm of the $tRNA^{Ile}$ gene (henceforth called a paralogue). A comparison of the mitochondrial gene arrangements between ND1 and CO1 for pythons, viperid and colubrid snakes is shown in Fig. 3.6. The beginning of the *CR* sequence contained a 15 bp hairpin (the consensus sequence of which was CAGCCATAGTC TGTAAATACAGACTATGGCTG). There was a C-rich region consisting of two runs of nine cytosines separated by a thymine and an adenine. A C-rich region is also found at the 5' end of the CR in some birds (Marshall and Baker, 1997). There was a five-base hairpin which also incorporated a termination-associated sequence (TAS), 60 bp 3' to the C-rich region and the conserved sequence block CSB-1, common to mammalian control regions, was also present. There was no evidence of the mammalian CSB-2 or CSB-3. There was a CA di-nucleotide series of five repeats ~70 bp 5' of the CSB-1 and a 5 bp hairpin, 29 bp 3' of the CSB-1. At the 3' end of the *CR* there was a series of tandem repeats of 87 bp in four *M. viridis* individuals and 88 bp in two *M. viridis* and in three *M. spilota* individuals (Fig. 3.7). The sequence of each repeat within an individual remained the same as the first repeat directly 3' to the non-repeat region, except at the end of the repeat region adjacent to the $tRNA^{Phe}$ gene. However, the sequence, copy number and endpoint of the repeat region were variable between individuals (Fig. 3.5). Repeat sequence copy number ranged from 4½ to 6½ repeats and the last partial repeat was separated from the $tRNA^{Phe}$ gene by 3-29 bp of spacer DNA.

To compare the *CR* of *Morelia* with other snakes, single individuals of *Acrochordus arafurae*, *Antaresia childreni*, and *Python reticulatus* were sequenced for the whole *CR* region. Sequence are shown in Appendix 3.1. Most of the *CR* of *Candoia aspera* was also sequenced, except for completely sequencing through the repeat region (see below). All the complete *CR* sequences of the taxa examined, with the exception of *Acrochordus arafurae*, had a series of 2-6 tandem repeats 3' to the CSB-1. *Antaresia childreni* had only two



a)



b)

Fig. 3.6 a) Schematic of mitochondrial gene arrangements in colubrids, viperids and pythons for the region between the *ND6* and *COI* genes, which includes the *control region* and *control region-like* genes. * taken from Kumazawa *et al.* 1996. b) Secondary structures present in pythons (corresponding to Table 3.1) and PCR primer locations. E=glutamic acid, T=threonine, P=proline, F=phenylalanine, V=valine, L=leucine, I=isoleucine, Q=glutamine, M=methionine, W=tryptophan, A=alanine, N=asparagine, C=cysteine, Y=tyrosine, I*=isoleucine pseudogene, ½H=17bp partial repeat of hairpin, 5'H=hairpin, C=C-rich region, H-T=hairpin+associated termination-associated-sequence2, CA=CA di-nucleotide repeat series, CSB=conserved sequence block-1, 3'H=haripin 3' to CSB, R=repeats. Primers are marked with arrows indicating direction. Lines link associated regions between the common vertebrate gene arrangement and the viperid, colubrid and boid taxa considered in this study.

Fig. 3.7 Sequence alignment of the *CR* from the CSB-1 to the end of the first repeat unit for selected python genera and C. *aspera* showing the differences in the starting point of the repeats. The boundaries of the repeat units are marked by *. The 11bp deletion is shown and the G-A transition site is marked with an arrow.

imperfect repeats. The colubrid and viperid snake *CRs* examined by Kumazawa *et.al.* (1996; 1998) did not have tandem repeats at the 3' end of the *CR*. Comparison of the repeat regions between python species showed considerable sequence variation and the point from which the repeats were generated was different for each taxon (Figs. 3.5 and 3.7). The size of the repeats was conserved, being mostly 86-88 bp, with the exception of *C. aspera* (86-96 bp).

The *control region* PCR product amplified for *C. aspera* was ~3500 bp long. At the 5' end of the *CR*, *C. aspera* showed none of the structures between the $tRNA^{Pro}$ and the C-rich region present in pythons, except for 12 nucleotides that could be aligned with an imperfect CT di-nucleotide repeat series also present in pythons (Table 3.1). At the 3' end of the *CR* there was a combination of tandem repeats of either 85 or 96 bp (Fig. 3.7). There were four different versions of the repeat sequence as follows: i) a 96 bp sequence, designated A, ii) A with a G-T transversion, designated B iii) A with an 11 bp indel or iv) B with an 11 bp indel (designated C and D respectively). Sequencing all the way through the repeat region could not be achieved, with the readable endpoint of the sequencing from either direction terminating within the repeat region. Consequently, the point of overlap of the light and heavy strands was not determined, suggesting that the unusually large size of this fragment was due to a large number of repeats. Single stranded sequence was obtained for eight repeats downstream of the non-repeat region and for four and a half repeats adjacent to the $tRNA^{Phe}$ gene. The order of the repeat types at the 5' end is as follows: non-repeat region, A, B, C, C, D, A, C, C,..... and at the 3' end is: .....C, D, A, C, D-partial, $tRNA^{Phe}$.

Some of the structural regions of the *CR* reported in the colubrid snake genus, *Dinodon* and the viperid genera, *Crotalus* and *Ovophis* (Kumazawa *et al.*, 1996; 1998) were not present in the python *CRs* examined here. The python sequences corresponding to hairpins 1 and 2 and the termination sequence TAS1 in *Dinodon* were very variable, indicating that these structures are not conserved in pythons. In contrast, TAS 2 and hairpin 3, which are often associated in vertebrate *CRs*, were both present in the python *CRs* studied. In the python *Antaresia childreni* there was a 19 bp repeat of the 5' portion of the first hairpin

Table 3.1. Table showing primary and secondary structures that are present (O) or absent (X) in the snakes examined in this study. Ile* = isoleucine pseudogene, ½ H = 19bp partial repeat of hairpin, 5' H = hairpin adjacent to Ile*, C = C-rich region, H-T = hairpin + associated TAS2, CA = CA di-nucleotide repeat series, CSB = Conserved Sequence Block-1, 3' H = hairpin 3' to the CSB-1, R= repeats, P = partial structure.

Species	Ile*	½ H	5' H	C	H-T	CA	CSB	3' H	R
Acrochordus arafurae	X	X	X	O	O	X	O	O	X
Antaresia childreni	O	O	O	O	O	O	O		O
Antaresia maculosa	O	X	O	O	O	O			
Antaresia perthensis	O	X	O	O	O	O			
Antaresia stimsoni	O	O	O	O	O	O			
Apodora papuana	O	X	O	O	O	O			
Aspidites melanocephalus	P	X	O	O	O	O			
Aspidites ramsayi	P	X	O	O	O	O	O	O	O
Bothrochilus boa	O	X	O	O	O	O			
Candoia aspera	X	X	X	O	O	O	O	O	O
Leiopython albertisii	O	X	O	O	O	O			
Morelia amethistina	O	X	O	O	O	O			
Morelia bredli	O	X	O	O	O	O			
Morelia boeleni	O	X	O	O	O	O			
Morelia oenpelliensis	O	X	O	O	O	O			
Morelia spilota	O	X	O	O	O	O	O	O	O
Morelia viridis	O	X	O	O	O	O	O	O	O
Python curtus	X	X	X	O	O	O			
Python sebae	X	X	X	O	O	O			
Python reticulatus	O	X	O	O	O	O	O	O	O
Python timoriensis	P	X	O	O	O	O			

(see Table 3.1).

To further investigate the *CR* within the python lineage, a portion of the *CR* from the $tRNA^{Pro}$ gene to CSB-1 was sequenced for 14 other python species. (For sequence alignments see Appendix 3.2). Up to 720 nucleotides were aligned and compared. Whilst in general it was not difficult to align these sequences, there were major differences in the nucleotide sequences in some regions. All of the Australian-New Guinean pythons (the Moreliini of Underwood and Stimson, 1990) have a similar *CR* structure for the ~720 bp region compared as described above (Table 3.1). The 19 bp repeat of the 5' portion of the first hairpin found in *Antaresia childreni* was also present in *A. stimsoni,* but was not present in the congenerics, *A. maculosa* and *A. perthensis* (Table 3.1). To optimise the sequence alignment of all 20 taxa, many small indels (1-4 bp) were necessary and large indels of 6 bp-25 bp needed to be incorporated into the sequences 3' of the C-rich region of *B. boa, L. albertisii* and *M. boeleni* (see sequence alignment in Appendix 3.1). The C-rich region, hairpin 3/TAS2 structures and CA di-nucleotide repeat series were conserved across all boid taxa examined (Table 3.1).

The region of the *CR* between the $tRNA^{Thr}$ and the C-rich region was more variable than the region 3' to the C-rich region. The 19 bp isoleucine paralogue found in *Morelia* was not present in *Acrochordus arafurae, C. aspera , P. curtus* or *P. sebae* (Table 3.1) and only 10 bp and 14 bp of the paralogue was present in *P. timoriensis* and the two *Aspidites* species respectively. The isoleucine paralogue was present in all other taxa compared. The hairpin that is adjacent to the isoleucine paralogue in *Morelia* was not present in *Acrochordus arafurae, C. aspera , P. curtus* or *P. sebae.* This hairpin was present in all other taxa compared but the number of Watson-Crick pairs involved and the number of nucleotides in the loop varied from 9-15 bp and 0-5 nucleotides, respectively.

3.3.4 Comparison of short and long PCR products for Acrochordus arafurae CR.

Two bands amplified with the *CR* primers for *Acrochordus arafurea* were gel purified (chapter 2.5) and sequenced. (The short *CRL* gel-purified PCR products did not

produce clean, readable sequence). The sequences of the shorter *CR* products were identical for the 620 bp in common to the longer *CR* products. A second H690 priming site was found nested within both longer sequences, which allowed the amplification of two different sized products from the same template. (A comparison of the long and short sequences and the positions of the proposed second priming site are shown in Appendix 3.1). It is therefore most likely that *A. arafurae* contains only a single *CR* and *CRL*. The H690 primer sequence was only partially present in the corresponding position of python *CR*s and there was no similar sequence in the colubrid snake, *Dinodon semicarinatus* published by Kumazawa *et al.* (1998).

3.3.4 Comparison of CR *and* CRL

The mitochondrial *CRL* region from the 3' end of the *ND1* gene to the $tRNA^{Met}$ gene was PCR amplified from caesium-chloride purified mitochondrial DNA for *Morelia spilota* SAMA R29677 and *M. viridis* AMS R115348. The product was completely sequenced for three *M. spilota* (SAMA R29677, ABTC 68308 and ABTC 68280) and one *M. viridis* AMS R115348 and partially sequenced for the 5' and 3' ends of *Acrochordus arafurae, Antaresia childreni, C. aspera* and *P. reticulatus*. The gene order for *M. spilota* and *M. viridis* was as previously reported for pythons (Kumazawa *et al.*, 1996) with the *CRL* sequence being translocated between the $tRNA^{Ile}$ gene and the $tRNA^{Glu}$ gene. The $tRNA^{Leu}$ gene is also translocated between the *CRL* and the $tRNA^{Glu}$ gene.

A product of approximately 900 bp from the *ND1* gene to the CSB-1 was also sequenced for seven other individuals of *M. spilota*. As previously found in other snakes (Kumazawa *et al.*, 1996; 1998), for the python taxa examined in this study, the *CRL* sequences commencing downstream from the $tRNA^{Ile}$ gene to the CSB-1 were indistinguishable from the *CR* sequence. However, in *M. spilota* ABTC 68308, there was a C-T base substitution between the *CR* and *CRL* in one of the 5 *CR* (3¾ *CRL*) 88 bp repeat blocks situated between the CSB-1 and the $tRNA^{Leu}$ gene (Appendix 3.2). The *CR* and *CRL*

differed further in the number of repeats between the CSB-1 and the *tRNA^{Leu}* gene. For *Acrochordus arafurae, Antaresia childreni, C. aspera* and *P. reticulatus* that were partially sequenced for the flanking regions of the *CR* and *CRL*, sequence similarity was found for ~ 400 bp beyond the tRNAs and for the repeat blocks. For the eight ~800 bp partial *CR* and *CRL* sequences of *M. spilota* examined, the sequences between the *tRNA^{Ile}* gene and the non-repeat region were indistinguishable.

3.4 Discussion

3.4.1 Translocation of the Control Region into the IQM tRNA cluster

The present study has confirmed earlier reports of a second copy of the *CR* found in the *IQM transfer RNA* region of the snake mitochondrial genome (Kumazawa *et al.*, 1996). The range of taxa that do not have the insertion was expanded to incorporate skinks, monitors, helodermatid lizards and ramphotyphlopid snakes and the range of those with the insertion was expanded to include file snakes and elapids. The results of the present study support the hypothesis that the *CRL* duplication/insertion event occurred after the divergence of the major snake lineages from the more ancestral Scolecophidian lineage, but prior to the divergence of lineages of boids, colubrids, viperids and elapids from their common ancestor (Fig. 3.1a).

3.4.2 Sequence conservation between the CR *and* CRL

An interesting feature of the *CR* and *CRL* sequence in snakes is the maintenance of sequence fidelity between the two regions (Kumazawa *et al.*, 1996; 1998). The *CR* is non-coding and whilst there are domains that are highly conserved between quite divergent lineages (Brown, 1986; Clayton, 1992), it is likely that very little of the sequence has functional constraints. These more rapidly evolving, variable domains have proved to be useful for the study of evolutionary relationships at the population level (Avise, 1992; Brown *et al.*, 1992; Moritz *et al.*, 1987). A second copy of the *CR* should have a similar rate of

nucleotide mutation to the *CR*, which should be independent of the sequence evolution evidenced in the *CR*. However, it has been shown that the *CR* and *CRL* sequences in snakes are evolving in a concerted fashion (Kumazawa et al. 1996; the present study). From the present study of the *CR* and *CRL* in 11 genera of snakes from five families, it can be concluded that the incidence of recombination between the *CR* and *CRL* at least matches the sequence mutation rate in snakes, for only three unambiguous nucleotide differences (in two samples) were detected out of more than 11,000 nucleotides sequenced. For concerted evolution to occur, the sequences need to come into contact in order to facilitate an exchange of nucleotides. If the frequency of this exchange is low, a lag time might be expected where there would be differences between the *CR* and *CRL*. The extremely high level of sequence similarity between the non-repeat portions of the *CR* and *CRL* within individuals of *M. spilota* with *CR* sequence divergences of 0-6% between individuals, suggests that homogenisation of the two regions is keeping pace with nucleotide substitution rate. The boundaries of the recombination are marked by $tRNA^{Phe}$ and $tRNA^{Leu}$ at the 3' ends of the *CR* and *CRL* respectively and by $tRNA^{Ile}$ at the 5' end. Whilst it could not be definitively determined from this study whether or not the $tRNA^{Ile}$ paralogue is part of the homogenising event, the sequence similarity for the portion shared between the isoleucine paralogue and the functional gene suggests that homogenisation is also occurring between these regions. The repeat units are involved in the homogenisation, as the repeat sequences are conserved between the two regions, except for a single individual that had a transition in one of the repeats of the *CRL*. This may be evidence of a new substitution that has not yet been homogenised by recombination between the two regions. However, the generation of copy number of the repeats appears to be independent of the homogenisation of the two regions, suggesting that replication strand-slippage may be occurring separately at both regions or that different mechanisms are involved in the homogenisation phenomenon.

A recombination model for the mitochondrial genome now has support from studies that show homologous recombination activity in the mitochondria of mammals (Thyagarajan

et al., 1996) and nematodes (Lunt and Hyman, 1997). Kumazawa *et al.* (1998) proposed two models to account for this concerted evolution, a tandem duplication model or the crossing of nicked DNA strands through a Holliday structure. The preferred model of Kumazawa *et al.* (1998) suggested a tandem duplication of the *CR* through replication slippage, with light-strand synthesis then producing two mitochondrial types, one with two identical *CR*s and one with three copies. Under this model, if replication was initiated from the *CRL*, the $tRNA^{Ile}$ paralogue could also have been introduced into the *CR* region. The alternative model proposed recombination by crossing over and resolution of a Holliday structure to produce homogeneous sequences in both regions. Neither model can be discounted on the basis of my data.

3.4.3 CR *structure*

Within the python group it was shown that the *CR* is quite dynamic in its evolution of primary and secondary features. The *CR*s of pythons contain many elements that are important to the replication and transcription of the mitochondrial genome which are common to most vertebrates, such as the Conserved Sequence Block 1 (CSB-1), hairpins and the termination associated sequence, TAS2. However, elements such as CSB-2 and 3 and multiple TAS elements that were also thought to play a significant role in the function of the *CR* based on sequence conservation amongst reported vertebrate lineages, are not present in pythons. This suggests that either the functions of the regions that are absent have been taken over by some other sequences within the genome or there are other, as yet undetermined, regulatory mechanisms involved. There is also considerable flexibility for the *CR* to accommodate other secondary structures as evidenced from the presence of a large hairpin in the 5' region of some pythons. As this structure is absent from other snake lineages and from some pythons, if it has a function at all, it is dispensable. Similarly, whilst the C-rich region is present in all snake lineages examined to date, its variable position within the *CR* suggests that its location is not crucial to any function it may have.

The present study found that the gene arrangement for $tRNA^{Ile}$, $tRNA^{Leu}$ and $tRNA^{Gln}$ in *P. regius* reported by Kumazawa *et al.* (1996) was common to the file snake (*Acrochordus arafurae*) and the four boid genera examined, *Antaresia*, *Candoia*, *Morelia* and *Python*. The $tRNA^{Pro}$ gene was in the usual vertebrate arrangement and not in the *CRL* nor was there evidence of a $tRNA^{Pro}$ paralogue as was found in colubrids and viperids (Fig. 3.6). The presence of the *CRL* and the incorporation of the $tRNA^{Leu}$ gene into the *IQM* region in all the snake lineages except Scolecophideans examined to date suggests that the duplication/translocation occurred before the major snake lineages diverged (Kumazawa *et al.* 1996). If the duplication/translocation occurred as a single event, it is most likely that the region between the $tRNA^{Thr}$ and *12S rRNA* genes were involved, with subsequent differential loss of copied tRNAs or parts thereof. If the *CR* and associated flanking tRNAs did not translocate into the *IQM* region as a single event, subsequent translocation events have inserted genes into an order comparable to the usual arrangement for the vertebrate *cytochrome b* to *12S rRNA* region.

Unlike the *CRL* that is ubiquitous to all snakes except the Scolecophidians, the $tRNA^{Ile}$ paralogue is present in the *CR* of the Moreliini pythons (pythons of New Guinea and Australia, (Underwood and Stimson, 1990) and *Python reticulatus* and *P. timoriensis*, but is not present in *P. sebae,* and *P. curtus*, *Candoia aspera*, *A. arafurae*, *Acanthophis,* colubrids and viperids (Kumazawa *et al.*, 1996) (Table 3.1). This suggests that the duplication/translocation event that produced the $tRNA^{Ile}$ paralogue occurred within the python lineage after the African and Asian Pythonini had diverged but prior to the divergence of the Indonesian and Australo-New Guinea pythons. The Moreliini pythons and *P. reticulatus* and *P. timoriensis* have a large hairpin structure at the commencement of the *CR* and *CRL*. It has been suggested that hairpin structures and tRNA pseudogenes that can form stable stem-loop structures may cause the stalling of heavy-strand replication to aid the formation of the D-loop (Doda *et al.*, 1981; Kumazawa *et al.*, 1998). These structures have also been proposed as termination signals that regulate transcription levels of some

polycistronic light-strand RNA transcripts, leading to a higher production of shorter transcripts of the ribosomal sub-units only. A short sequence acting as a differential terminator of transcription (or transcriptional attenuator-like signal) contained within the $tRNA^{Leu}$ that is situated 3' to the *16S rRNA* gene regulates transcription levels in humans (Christianson and Clayton, 1988). A DNA binding protein is suspected of using this sequence to block elongation by mitochondrial RNA polymerase (Hess *et al.*, 1991). In snakes, the $tRNA^{Leu}$ gene has been translocated to the 3' end of the *CRL* and therefore may not be functional as a terminator of transcription (Kumazawa *et al.*, 1996). If the hairpin sequences in the 5' region of the *CR* were acting as transcriptional terminators, the short polycistronic transcript would be translated for 12S rRNA, 16S rRNA and ND1 gene products. Therefore, as well as elevated levels of the ribosomal sub-units, there would also be an increase in ND1. Because all of the elements for the initiation of transcription are present in both the *CR* and *CRL*, if these structures function as transcription terminators, the mitochondrial polycistronic transcripts in snakes may have at least two forms. A short transcript of the ribosomal sub-units and *ND1* and a long transcript of the genes between *ND2* and the *CR*. If the 5' hairpin structures have an attenuator function, there may be three or more forms of polycistronic transcript, which would increase the amount of 12S rRNA and 16S rRNA present in the cell. Alternatively, to have a differential transcription of the ribosomal genes, there would need to be a transcriptional enhancer of some sort associated with initiation of transcription from the *CR* and not from the *CRL*. There is no evidence to support the existence of such an enhancer in *CR*.

At the 3' end of the *CR* sequence in pythons there are a series of tandem repeats of approximately 88 nucleotides with the size of the repeats being mostly conserved across all the python species (Fig. 3.7). There is moderate sequence conservation of the repeat region between genera, but the repeat units themselves are mostly perfect except for the final unit which sometimes has only a portion of the repeat perfectly conserved. Between individuals of the same species, the sequences of the repeat units vary, reflecting population level sequence

mutation (Figs. 3.4 and 3.5). However, the 5' boundary of the region of repeat units is conserved within a species (Figs. 3.4 and 3.5), but is variable between species (Fig. 3.7). The fact that repeat regions reflect sequence that is unique to a particular individual suggests that the generation of the repeats happened after the sequences had evolved and the generation/deletion of repeats is perhaps an ongoing phenomenon. However, the presence of a series of similar-sized repeats commencing around a common point in most of the taxa considered in this study (Fig. 3.7), suggests a common element in the primary or secondary structure 3' to the CSB-1 that facilitates the generation of repeats from a particular region of sequence.

The presence of tandemly-repeated sequences within the *CR* is not uncommon. There are large (82 bp) and short (10 bp) tandem repeats associated with the TAS sequences at the 5' end of the *CR* in fish (Brown *et al.*, 1992; Buroker *et al.*, 1990; Faber and Stepien, 1998) and 4-32 bp tandem repeats at the 3' end in some birds (Berg *et al.*, 1995). Frequently, there is length heteroplasmy caused by mitochondria having different copy numbers of the repeats (Brown *et al.*, 1992; Buroker *et al.*, 1990). It has been suggested that repeated DNA sequence is often generated by slipped-strand mis-pairing occurring during strand replication (Macey *et al.*, 1997; Stanton *et al.*, 1994). This frequently occurs where there are runs of similar bases and may be the mechanism that accounts for the many short (5-9 bp) repeats found scattered throughout the python *CR*. However, random slipped-strand mis-pairing by itself would not explain the maintenance of repeat boundaries such that the repeat blocks were of perfect sequence similarity. Stem and loop or palindromic sequences have been implicated in slipped-strand mis-pairing (Glickman and Ripley, 1984; Ripley and Glickman, 1983; Stanton *et al.*, 1994). Buroker *et al.* (1990) proposed a model to generate perfect-sequence tandem repeats that used a secondary structure in the displaced strand of the D-loop. This model relies on the proximity of the TAS sequences to the repeat and also on the internal base pairing of the repeat to form stable hairpin structures that cause misalignment during replication of the heavy strand. Whilst the repeat region of pythons is not associated with the

5' TAS elements, the involvement of stem-loop structures in slipped-strand mis-pairing is supported in our study by the presence of a run of 5-6 thymines in the repeat units that would form a stable stem and loop with a run of adenines in the region downstream of the CSB-1. Slipped-strand mis-pairing to produce the 19 bp repeat insert in the two *Antaresia* species could have been initiated from a hairpin formed in the $tRNA^{Ile}$ paralogue.

The present study showed that the copy number of repeats was variable between individuals of the same species and varied between the *CR* and *CRL* within an individual (Appendix 3.1). If the mechanism of repeat generation/deletion occurs with strand replication, which is separate from the concerted evolution that is maintaining sequence fidelity between the *CR* and *CRL*, the repeat copy number for the *CR* and *CRL* may be specific to each region and may be constantly changing around an optimal mean. In *C. aspera*, the repeat units reflect the maintenance of a series of mutations (a transition, an 11 bp indel and a combination of both) in the repeat unit and there is no pattern of repeat type. This suggests that if strand slippage is involved in changing the copy number of repeats, it does not always commence at the same point, nor does it occur sequentially along the repeat units.

Throughout the *CR* of pythons there are a number of isolated repeated sequences of 5-12 bp that are often associated with runs of adenines or thymines. This finding supports earlier reports of sequences with a high AC/TG-content being more commonly involved in replication slippages than other sequence motifs (Levinson and Gutman, 1987).

Therefore, it can be concluded from the present study that the *CR* and *CRL* regions of pythons are evolving in a concerted manner, with sequences within individuals becoming identical at a rate similar to the rate of nucleotide substitution. In only two individuals was there evidence of different sequences for the *CR* and the *CRL*. As each individual was only sequenced once there is a possibility that the nucleotide differences are a result of polymerase substitution error. The question could be settled by sequencing multiple clones.

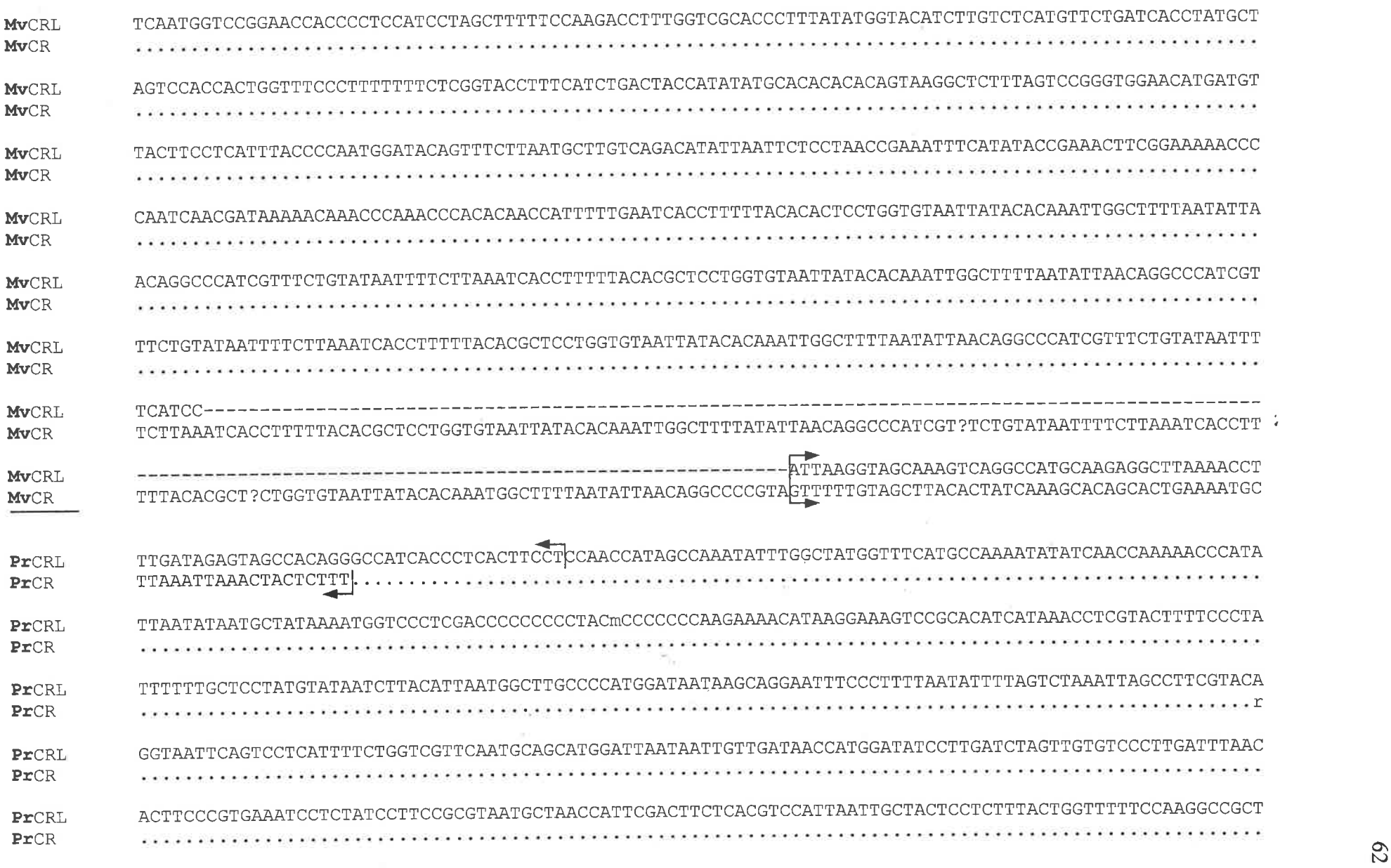
Appendix 3.1

Pair-wise comparisons between *control region* and *control region-like* nucleotide sequences. **Aa**=*Acrochordus arafurae*, SAMA R23999; (SCR=short *CR* product, LCR= long CR product, AaCRL is incomplete); **Ac**=*Antaresia childreni*, SAMA R21411; **Ms**=*Morelia spilota*, **1**=SAMA R26878, **2**=ABTC 68308, **3**=ABTC 68280, **4**=ABTC 66175, **5**=ABTC 66267, **6**=ABTC 66295, **7**=ABTC 68278, **8**= ABTC 68279, **9**=ABTC 68281, **10**=ABTC 68288; **Mv**=*Morelia viridis*, AMS R115348; **Pr**=*Python reticulatus*, SAMA R28533; ?= missing data; **.**= similarity with first sequence; -=gap; ↲=tRNA-CR/CRL junction. Position of primer-like sequences in Aa are shown - reverse primer sequence for *CR* is marked in red; reverse primer sequence for *CRL* is marked in blue; :=similarity with primer sequence; *=difference from primer sequence.

```
AaCRL   CCAATAACACTAAGCCATCTGCCTACTAAACATTTCCACCACATTAACATTCAACGGAATTCCCCCACAATGGAAGCGTGCCTGAGCAAGGATTACTTTG
AaLCR   ?????????????????????????????????????????????????????????????????????????????ATCAATTCATCTCTGGCCCCCAA
AaSCR   ?????????????????????????????????????????????????????????????????????????????.....C.................

AaCRL   ATATAGTAAACACGGAGACCAACACTCCCACTTCCCCTTGAAAAATAAGCCCTCCTGGACCCCCCCTTACCCCCCCCAAAACCTCTATGTATTATTATT
AaLCR   AACCAGCATTTTAACATTAAACTACTTTCTGCATAACATA............................................................
AaSCR   ....................................................................................................

AaCRL   ATACTTTATGTCTACATTTATTGCTATGTATAATAATACATTAACGTTTTGCCTCATGGATAATAAACCAGTATTCTATATTAATTAATTAAGCCTATTA
AaLCR   ......................................................................................  .............
AaSCR   ....................................................................................................

AaCRL   TTGCCTATGTGCAAAAGAAAAGTACCTCATTTCTTGGTCGTTCTATGCAGCACGGATTAATGTAATTCTTAATCAGTATGGATATTAAGAACCTAGTGGT
AaLCR   ....................................................................................................
AaSCR   ....................................................................................................

AaCRL   GTCCCATGGTTATGCCCTTCCCGTGAAATCCTCTATCCTTTCACCACACGCTAATTAACCGGCTTTTCAGGTCCATATTACTGTTGTCCCTCATATAGTG
AaLCR   ...................................................k...............m................................
AaSCR   ...................................................G...............A................................

AaCRL   TACTTTAAGAGGCCACTGGTTACACCTTCAAGATCATTTCGACGGCCCGGAACCATCCCTCCCTACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTA
AaLCR   ....................................................................................................
AaSCR   ....................................................................................................

AaCRL   TAATGGTACATGTCTCTCATGTTCTGATCACGTATGCTCGTTCCACCACTGGTAGCTCTTTTTTATCTCTACCTTTCACCTGACTACCATATATGCACGG
AaLCR   ....................................................................................................
AaSCR   .......................................................................................*:::*:::::*:*
                                                                          H690 primer  3'-ATATGTACGTTCG
```


AaCRL CTCCGCGAGAGCGGGAGGTAGGTTATCTATTCCGGGTTGAGCCCTGTTCTTGGCTTCGCATATTCCCTAAGTGGCTAAGTTTCCTAATGCTTGTTAGACA
AaLCR
AaSCR
H690 GAGTTG-5′
H4433 5′- TATGGGCTTTTACAA-CCAA-3′
AaCRL TATATATATTGTTTACCGATAATCTTAATTGTTGTTTTAACCTAAATT
AaLCR .....T.....CACCTACATTAAAATTTATTTTATTAAAAAAATTAATAAAAATAAATTTTTT
AaCRL
AaLCR TCCAAAAATGTCACAGGTCACAAACAAAATAATTTATTTTTTATTCCAAAAAATACAACCAATATCAACCATAACCTTATTATCTCAAATTTTTTATTAC
AaCRL
AaLCR CATAAATCACAATTGATTTCTAACACAGTCAGTGCTGAAATTAAACCTAACCTCTATAAGCAACCACACTAAAATTACCTAGAATTAATCAAAAGAACC.
AaCRL ATTTTCCATAAAAAGCCATATATTACCCCCTATCCGGGGGACCAGGTTTGCTATATATTTCAT----ATTAAGGTAGCAAAGTAAGGCTATGCTGTAGGC
AaLCR .....-ATTTGGGGGGCCCGGTTTCCCATGTATTTTTTATCACTGTTATTATAGCTTAATACACAAAGCATAGCA
AcCRL CCCCCACAATGGAAGCGTGCCCGAGCCAGGGGCTACCTTGATAGAGTAGACACAGGGCCACAACCCTCACTTCCTTCAGCCATAGTCTGTAAATACAGCC
AcCR ArGAmACTTCATCCCTGGTCCCCAAAACCAGAATTTTTAATTTAAACTACTCTTT.....
AcCRL ATAGTCTGTAAATACAGACTATGGCTGATGCCGCCATATAGAGCCGAAAAACCATATAATATACCACACAATAAACTCTCTCCTCGGCCCCCCCCCTACC
AcCR .....m
AcCRL CCCCCCCACTAAAACATATGGGAAAGCAGCACAAATACATATTAAAGAATGTCCAATTAATCCTCCTATGTATAATCTTACATTAATGGTTTGCCCCATG
AcCR .....y.....
AcCRL AATATTAAGCAGGAATTCCCCTTTAAATATTTTAGTCTAAAATGTCCTTTGTACAGAATATTTAGCCCTCATTCTCTGGTCGTTCAATGCAATCGGGATT
AcCR .....
AcCRL AATCATTCTTAATAACCATGACTATCCTTGATCTAGTGGTGTCCCTTGATTTAGTACTTCCCGTGAAATCCTCTATCCTTTCATAGAATGCTAACCATTC
AcCR .....
AcCRL GACTTCTCACGTCCATATAGTGTTGATCCTCTTAATAGCTTTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCC
AcCR .....


AcCRL
AcCR
ATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCTTTATATGGTTCATTTAGTCTCATGTTCTGATCAGCTATGTCAATCCACCACTGGTAGCTCTTTTT
TCTCTGTACCTTTCATCTGACTACCATATATGCACACACACAGTAAGGCTTTTAGTCCAGGTGGAGCACCATGTTTCTCCyTCATATTCCCCCCTwGGGA
TACATCTCTTAATGCTTGGAAGACATACAAATTACCCCTACCGAAATTTCATATACCAAAACTTCGGAAAAACCCCAATCTAACAAATAACAAATCCACC
ACACTTTTATCCTCCATCCAACACACCTAAAGTGAAACTATACACAAATCAGCTTTTAAAACCGCCAAGCCCATCGTTTTTGTATAAAACACTTTTATCC
TCCATCCAACACACCTGAGGCAGAACTATACACAAATCAGCTTTTAAAACCGCCAAGCCCATCGTTTTTGTATAAAACACTTTCAC----------
TATCCTCCACCCAACAC
ATTAAGG?AGCAAAACCAGGCCATGCAAGAAGCTTAAAACCTTGACACAGATGTTCA
ACCTATAGGCGGAGCCATACACAATCGACCTTTAAAACCACAGTTATTGTAGCTTACACTATAAAGCACAGCACTGAAAATGCTGAGAAGGTACCACACC
MvCRL
MvCR
AATAAACGTTTCAACCACCACAGCATTCCACGGTACACCCCCACAATGGAAGCGTGCCCGAGATTTCAGGGATTACCTTGATAGAGTAAACACAGGGATC
??????????????????????CTTTTTAGCATCAAAGAGAAAGAmGCTTCATCCCTGGTCCCCAAAACCAGTATTTTAAATTAAACTACTCTTTGAATC
AACCCTCACTTCCTCCTAAGCCATAGTyTGTAAGTTACAGACTATGGCTCATGCCTTAATATATAAACCAAAAACCCATATAATCACTGAACAATAAAAy
.....C.....C
TyTyTCCTyGGCCCCCCCCCTACCCCCCCCCGGAAAACCATAGwAGAAGTCAGCACATAAATAAACCTACTAATCCCATTGCTTCCTCsTATGTATAATC
.C.C....C.....A.....C.....
TTACATTAATGGTTTGCCCCATGAATATTAAGCAGGAATTTCCCTTCAAATATTTTAGCCTAAATTAGCTTCCGTACAAAATATCTAGCCCTCATTTTCT
GGTCGTTCAATGCAATCGGGGTTAATAAATCTTACTAACCATGGATATCCTTGATCAGGTGGTGTCTCTTAATTTAGTACTTCCCGTGAAATCCTCTATC
CTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATTATGTCAATCCTCCCTTCTAGCTTTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATC


MvCRL TCAATGGTCCGGAACCACCCCTCCATCCTAGCTTTTTCCAAGACCTTTGGTCGCACCCTTTATATGGTACATCTTGTCTCATGTTCTGATCACCTATGCT
MvCR .....
MvCRL AGTCCACCACTGGTTTCCCTTTTTTTCTCGGTACCTTTCATCTGACTACCATATATGCACACACACAGTAAGGCTCTTTAGTCCGGGTGGAACATGATGT
MvCR .....
MvCRL TACTTCCTCATTTACCCCAATGGATACAGTTTCTTAATGCTTGTCAGACATATTAATTCTCCTAACCGAAATTTCATATACCGAAACTTCGGAAAAACCC
MvCR .....
MvCRL CAATCAACGATAAAAACAAACCCAAACCCACACAACCATTTTTGAATCACCTTTTTACACACTCCTGGTGTAATTATACACAAATTGGCTTTTAATATTA
MvCR .....
MvCRL ACAGGCCCATCGTTTCTGTATAATTTTCTTAAATCACCTTTTTACACGCTCCTGGTGTAATTATACACAAATTGGCTTTTAATATTAACAGGCCCATCGT
MvCR .....
MvCRL TTCTGTATAATTTTCTTAAATCACCTTTTTACACGCTCCTGGTGTAATTATACACAAATTGGCTTTTAATATTAACAGGCCCATCGTTTCTGTATAATTT
MvCR .....
MvCRL TCATCC----------
MvCR TCTTAAATCACCTTTTTACACGCTCCTGGTGTAATTATACACAAATTGGCTTTTATATTAACAGGCCCATCGT?TCTGTATAATTTTCTTAAATCACCTT
MvCRL ----------ATTAAGGTAGCAAAGTCAGGCCATGCAAGAGGCTTAAAACCT
MvCR TTTACACGCT?CTGGTGTAATTATACACAAATGGCTTTTAATATTAACAGGCCCCGTAGTTTTTGTAGCTTACACTATCAAAGCACAGCACTGAAAATGC
PrCRL TTGATAGAGTAGCCACAGGGCCATCACCCTCACTTCCTCCAACCATAGCCAAATATTTGGCTATGGTTTCATGCCAAAATATATCAACCAAAAACCCATA
PrCR TTAAATTAAACTACTCTTT.....
PrCRL TTAATATAATGCTATAAAATGGTCCCTCGACCCCCCCCCCTACmCCCCCCCAAGAAAACATAAGGAAAGTCCGCACATCATAAACCTCGTACTTTTCCCTA
PrCR .....
PrCRL TTTTTTGCTCCTATGTATAATCTTACATTAATGGCTTGCCCCATGGATAATAAGCAGGAATTTCCCTTTTAATATTTTAGTCTAAATTAGCCTTCGTACA
PrCR .....r
PrCRL GGTAATTCAGTCCTCATTTTCTGGTCGTTCAATGCAGCATGGATTAATAATTGTTGATAACCATGGATATCCTTGATCTAGTTGTGTCCCTTGATTTAAC
PrCR .....
PrCRL ACTTCCCGTGAAATCCTCTATCCTTCCGCGTAATGCTAACCATTCGACTTCTCACGTCCATTAATTGCTACTCCTCTTTACTGGTTTTTCCAAGGCCGCT
PrCR .....


PrCRL GGTTACACCTTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCACCT
PrCR .....
PrCRL CATGTTCTGATCACCTATGCTAGACCACCACTGGTAGCTCTTTTTTCTCTCTCCCTTTCACCTGACTACCATATATGCACACACACAGTAAGGTCCTTAG
PrCR .....
PrCRL TCCAGGTGGAGCAACATGCTTACGTTACACCTACCCCTAATGGGATACATCTCTTAATGCTTGTTAGACATAGAATTTTTTCCAAACACCCACTCAAGGA
PrCR .....
PrCRL AACCAAACCACCCACCACCCAAAACGAAATAATACACAAATAAACTTTAAATCACAACAAACCCAGCGCTTTATATAAAATGTATTTATTTTTCCTTTTT
PrCR .....
PrCRL TCACCCAAAACGAAATAATACACAAATAAACTTTAAATTACAACAGGCCCAGCGTTTTTATATAAAATGTATTTATTTTTCCTTTTTTCACCCAAAACGA
PrCR .....
PrCRL AATAATACACAAATAAACTTTAAATTACAACAGGCCCAGCGTTTTTATATAAAATGTATTTATTTTTCCTTTTTTCACCCAAAACGAAATAATACACAAA
PrCR .....
PrCRL TAAACTTTAAATTACAACAGGCCCAGCGTTTTTATATAAAATGTATTTATTTTTCCTTTTTTCACCCAAAACGAAATAATACACAAATAAACTTTAAATT
PrCR .....
PrCRL ACAACAGGCCCAGCGTTTTTATATAAACCTTACACA----------
PrCR .....ATGTATTTATTTTTCCTTTTTTCACCCAAAACGAAATAATACACAAATAAACTTTAAATTACAACAGGCCCAG
PrCRL ----------TTTAAGGTAGCAAAGCCAGGCCATGCAATAGGCTTAAAACCTTGACACAGATGTTCAAATCATCTCCTTAAT
PrCR CGTTTTTATATAAAATGTATCTACCAAGGTTATTGTAGCTTACACAAAGCACAGCACTGAAAATGCTGAGATGGTGATTCACCCATATAACACAAAGTCT
Ms1CRL CTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCCAACCATAGTCTGTAATTTACAGACTATGGTCCATGCCTTAATATAAAGCCAAAAATCC
Ms1CR TTAAATTAAACTACTCTTT.....
Ms1CRL ATATAATTTACCACAAAATAAAGyTyTyTyTCGGCCCCCCCCCTACCCCCCCCCAAGAACATTGGGGAGACCGGCACACAAAACCATTAGAAAACTCTTA
Ms1CR .....C.C.C.C.....
Ms1CRL ACAAACCTCTCTATGTATAATCTTACATTAATGGTTTGCCCCATGAATATTAAGCAGGAATTTCCCTTTTATTATTTTAGTCTAAAATGGCCTTTGTACA
Ms1CR .....
Ms1CRL AAATATTCTGTCCTCATTCTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTAATCATGGATATTCTCAACCTAAGGGTGTCTCTTAGTCTAG
Ms1CR .....


Ms1CRL
Ms1CR
CGCTTCCCGTGAAATCCTCTATCCTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATCATGCTAATCCTCCCTACTAGCTCTTTCCAAGGCCG
CTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCAC
CTCATGTTCTGATCATCTATGTCTATCCACCATTGGTAGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCACACACACAGTTAGGCATTC
AGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCCTTCCCGAAATTTCATATA
CCAAAACTTCGGAAAAACCCCAATCCTGACAAAAATCTCACACAACTTTTTACTTTTTCCTCTAACCCACCTTCAACAGGTTAATACAGAAACCCGCCTT
TAAATACAGCAGGCCCATTATTTCTATATTAATTATTTTTATCCGCCTTCTAACCCACCTTCAACAGGTTAATACAGAAACCCGCCTTTAAATACAGCAG
GCCCATTATTTCTATATTAATTATTTTTATCCGCCTTCTAACCCACCTTCAACAGGTTAATACAGAAACCCGCCTTTAAATACAGCAGGCCCATTATTTC
TATATTAACCGATCTACC
TTATTTTTATCCGCCTTCTAACCCACCTTCAACAGGTTAATACAGAAACCCGCCTTTAAATACAGCAGGCCCATTATTTCTATATTAATTAT
ATTAAGGTAGCAAAGCAAGGCCATGCAAGAGGCTTAAAACCTTGACACAGATG
TTTTATCCACCTTCAACAGGTTAATACAAGGCCCCTTGTTATTGTAGCTTACATTACTTAAAGCACAGCACT
Ms2CRL
Ms2CR
CTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCCAACCATAGTCTGTAATTTACAGACTATGGTCCATGCCTTAATATAAAGCCGAAAATCC
TTAAATTAAACTACTCTTT
ATATAATCTGCCACAAAATAAAGyTyTyTyTCGGCCCCCCCCCTACCCCCCCCAAAAGAACATTGGGGAAACCGGCACAAAAAACTATTGGAAAACTACT
AACAAACCCCTCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTAAGCAGGAATTTCCCTTTTATTATTTTAGTCTAAAATGGCCTTTGTAC
AAAATATTCTGTCCTCATTCTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTGATCATGGATATTCTTAACCTAAGGGTGTCTCTTAGTCTA
m

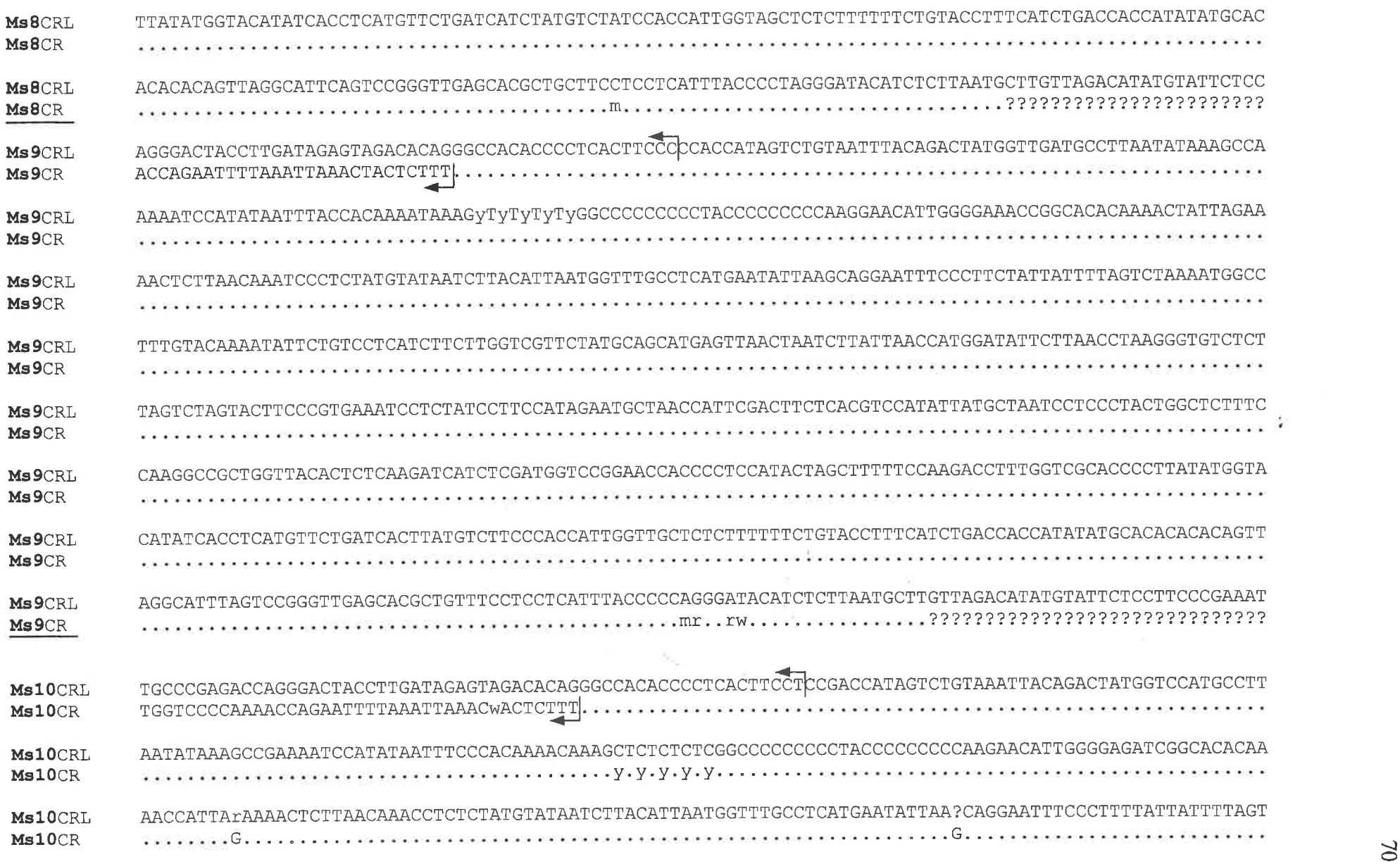

Ms2CRL GCACTTCCCGTGAAATCCTCTATCCTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTACTGGCTTTTTCCAAGGCC
Ms2CR .....
Ms2CRL GCTGGTTACACTCTCAAGATCATCTCGATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCA
Ms2CR .....
Ms2CRL CCTCATGTTCTGATCACTTATGTCAATCCACCACTGGTAGCCTCTTTTTTTCTGTACCTTTCATCTGACCTCCATATATGCACACACACAGTTAGGCACT
Ms2CR .....
Ms2CRL TAGTCCAGGTTGAGCACGCTGTTTCCTCCCCATTTACCCCTATGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCCTTCCCGAAATTTCATAT
Ms2CR .....
Ms2CRL ACCGAAACTTCGGAAAAACCCCAATCCTTTCAACAACCTCCCCACAGATTTTCATTTTTCCTCTAACACACCTCTGACAGATTAATACAGAAAGCCCCCT
Ms2CR .....
Ms2CRL TAAAATACAGCAGGCCCTTTATTTCTATATTAATTTTTATTTCCCTCCTTCTAACACACCTCTGACAGATTAATACAGAAAGTCCCCTTAAAATACAGCA
Ms2CR .....C.....
Ms2CRL GGCCCTTTATTTCTATATTAATTTTTATTTCCCTCCTTCTAACACACCTCTGACAGATTAATACAGAAAGCCCCCTTAAAATACAGCAGGCCCTTTATTT
Ms2CR .....y.....
Ms2CRL CTATATTAATTTTTATTTCCCTCCTTCTAACACACCTCTGACAGATTAATACAGAAAGCCCCCTTAAAATACAGCAGGCCCTTTATTTCTATATTAATCA
Ms2CR .....y.....TT
Ms2CRL ACCCACT----------
Ms2CR TTATTTCCCTCCTTCTAACACACTTCTGACAGATTAATACAGAAAGCCCCCTTAAAATACAGCAGGCCCTTTATTTCTATATTAATTTTATTTCCCCCAA
Ms2CRL ----------ATTAAGGTAGCAAAGCCAGGCCATGCAGAATGCTTAAAACCTTGACACAGATGTTCAAATCATCTCCTTAATACT
Ms2CR CAAAATGATACAGCAGACCCCCCCTGTTATTGTAGCTTACATTACTTAAAGCACAGCACTGAAAATG?TGAGATGGTACTCACCCAAATAACACAAAGTC
Ms3CRL CCCGAGACCAGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCTAACCATAGTCTGTGAATTACAGACTATGGTTGATGCCTTA
Ms3CR GCCCCCAAAACCAGAATTTTAAATTAAACTACTCTTT.....
Ms3CRL ATATAAAACCAAAAATCCATATAATTTACCACAAAACAAAGCTCTCTCTCGGCCCCCCCCCTACCCCCCCCCCAGGAACATTGGGGAGATCGGCACATAAA
Ms3CR .....y.y.y.y.y.....
Ms3CRL ACCATTAGAAAACTCTCAACAAACCTCGCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTTAGCAGGAATTTCCCTTTTATTATTTTAGTC
Ms3CR .....


Ms3CRL TAAAATGGCCTTTGTACAAAATATTCTGTCCTCATATTCTTGGTCGTTCTATGCAGCATGAGTTAACTAATCTTATTAACCATGGATATTCTTAACCTAA
Ms3CR .....
Ms3CRL GGGTGTCTCTTAGTCTAGCGCTTCCCGTGAAATCCTCTATCCTTCCATATAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTAC
Ms3CR .....
Ms3CRL TAGCTCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCC
Ms3CR .....
Ms3CRL TTATATGGTACATATCACCTCATGTTCTGATCATCTATGTCTATCCACCATTGGTAGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCAC
Ms3CR .....
Ms3CRL ACACACAGTTAGGCATTCAGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCC
Ms3CR .....
Ms3CRL TTCCCGAAATTTCATATACCGAAACTTCGGAAAAAACCCAATCCTGACAAAACCTCACACAACTTTTTTCCTTTTTCCTCTAACCCTCTCTCCAAGAGGT
Ms3CR .....
Ms3CRL TAATACAAAAACCCGCCTTTAAACACAGCAGGCCCATTATTTCTATATTATTTATTTTTATCCGCCGTCTAAACCCTCTCCAAGAGGTTAATACAAAAAC
Ms3CR .....
Ms3CRL CCGCCTTTAAACACAGCAGGCCCATTATTTCTATATTATTTATTTTTATCCGCCGTCTAAACCCTCTCCAAGAGGTTAATACAAAAACCCGCCTTTAAAC
Ms3CR .....
Ms3CRL ACAGCAGGCCCATTATTTCTAT----------
Ms3CR .....ATTATTTATTTTTATCCGCCGTCTAAACCCTCTCCAAGAGGTTAATACAAAAACCCGCCTTTAAACACAGCAGGCCCA
Ms3CRL ----------ATTAAGGTAGCAAAGCTAGGCCATGCAAGAGG
Ms3CR TTATTTCTATATTATTTATTTTTATCCACCGGTAATTTAACACAGAAGACCCTCTGTTATTGTAGCTTACAA??????????????
Ms4CRL TTCATCCACCAsTGCATTCCwCGGTACACCACCACAATGGAAGsGTGCCCGAGACCAGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACT
Ms4CR ????????????????????????????????????TTCATCCCTGGTCCCCAAAACCAGAATTTTAAATTAAACTACTCTTT.....
Ms4CRL TCCTCCCAACCwTAGTCTGTAATTTACAGACTATGGTCCATGCCTTAATATAAAGCCGAAAATCCATATAATCTGCCACAAAATAAAGCTyTyTyTyGGC
Ms4CR .....A.....y.....
Ms4CRL CCCCCCCCTACCCCCCCCCAAAGAACATTGGGGAAACCGGCACAAAAAACTATTGGAAAACTACTAACAAACCCCTCTAwGTwTAATCTTACATTAATGG
Ms4CR .....T..A.....


Ms4CRL TTTGCCTCATGAATATTAAGCAGGAATTTCCCTTTTATTATTTTAGTCTAAAATGGCCTTTGTACAAAATATTCTGTCCTCATTCTCTTGGTCGTTCTAT
Ms4CR .....
Ms4CRL GCAGCACGAGTTAACTAATCTTATTGATCATGGATATTCTTAACCTAAGGGTGTCTCTTAGTCTAGCACTTCCCGTGAAATCCTCTATCCTTCCATAGAA
Ms4CR .....
Ms4CRL TGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTACTGGCTTTTTCCAAGGCCGCTGGTTACACTCTCAAGAT?ATCTCGATGGTCCG
Ms4CR .....C.....
Ms4CRL GAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCACCTCATGTTCTGATCACCTATGTCAATCCACCACT
Ms4CR .....
Ms4CRL GGTAGCCTCTTTTTTTCTGTACCTTTCATCTGACCTCCATATATGCACACACACAGTTAGGCACTTAGTCCAGGTTGAGCACGCTGTTTCCTCCCCATTT
Ms4CR .....
Ms4CRL ACCCCTATGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCCT
Ms4CR .....??????????????
Ms5CRL CCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCCCCACCATAGTCTGTAATTTACAGACTATGGTTGATGCCTTAATATAAAGCCAAAAATCCA
Ms5CR TTTAAATTAAA?TAsTyTTT.....s.....y.....s.....
Ms5CRL TATAATTTACCACAAAATAAAGCTyTyTyTCGGCCCCCCCCCTACCCCCCCCCAAGGAACATTGGGGAAACCGGCACACAAAACTATTAGAAAACTCTTA
Ms5CR .....s.....y.....s.....
Ms5CRL ACAAATCCCTCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTAAGCAGGAATTTCCCTTCTATTATTTTAGTCTAAAATGGCCTTTGTACA
Ms5CR .....
Ms5CRL AAATATTCTGTCCTCATCTTCTTGGTCGTTCTATGCAGCATGAGTTAACTAATCTTATTAACCATGGATATTCTTAACCTAAGGGTGTCTCTTAGTCTAG
Ms5CR .....
Ms5CRL TACTTCCCGTGAAATCCTCTATCCTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTACTAGCTCTTTCCAAGGCCG
Ms5CR .....s.....
Ms5CRL CTGGTTACACTCTCAAGATCATCTCGATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCAC
Ms5CR .....
Ms5CRL CTCATGTTCTGATCACTTATGTCTTCCCACCATTGGTTGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCACACACACAGTTAGGCATTT
Ms5CR .....


Ms5CRL AGTCCGGGTTGAGCACGCTGTTTCCTCCTCATTTACCCCCAGGGATACATCTCTTAATGCTTGTTAGACAT
Ms5CR .....??????????
Ms6CRL CTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCCAACCATAGTCTGTGAATTACAGACTATGGTTGATGCCTTAATATAAAGCCAAAAATCC
Ms6CR TTAAATTAAACTACTCTTT.....s.....
Ms6CRL ATATAATTTACCACAAAACAAAGCTCTCTCTCGGCCCCCCCCCCTACCCCCCCCCCAAGGAACATTGGGGAAATCGGCACATAAAACCATTAGAAAACTCTC
Ms6CR .....y.y.y.y.....
Ms6CRL AACAAACCTCTCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTTAGCAGGAATTTCCCTTTTATTATTTTAGTsTAAAATGGCCTTTGTAC
Ms6CR .....C.....
Ms6CRL AAAATATTCTGTCCTCATATTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTAACCATGGATATTCTTAACCTAAGGGTGTCTCTTAGTCTA
Ms6CR .....
Ms6CRL GCGCTTCCCGTGAAATCCTCTATCCTTCCGTATAATGCTAACCATTsGACTTCTCACGTCCATATCATGCwAATCCTCTTTACTTGCTCTTTCCAAGGCC
Ms6CR .....C.....T.....
Ms6CRL GCTGGTTACACTCTCAAGATCATCTCGATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTACATATCA
Ms6CR .....
Ms6CRL CCTCATGTTCTGATCATCTATGTCTACCCACCATTGGTTGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCACACACACAGTTAGGCATT
Ms6CR .....
Ms6CRL TAGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCC
Ms6CR .....??????????????
Ms7CRL CCCGAGACCAGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCCAACCATAGTCTGTGAATTACAGACTATGGTTGATGCCTTA
Ms7CR GCCCCCAAAACCAGAATTTTAAATTAAACTACTCTTT.....
Ms7CRL ATATAAAACCAAAAATCCATATAATTTACCACAAAACAAAGyTyTyTyTyGGCCCCCCCCCCTACCCCCCCCCCAAGGAACATTGGGGAGATCGGCACATAA
Ms7CR .....s.....
Ms7CRL AACCATTAGAAAACTCTCAACAAACCTCGCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTTAGCAGGAATTTCCCTTTTATTATTTTAGT
Ms7CR .....

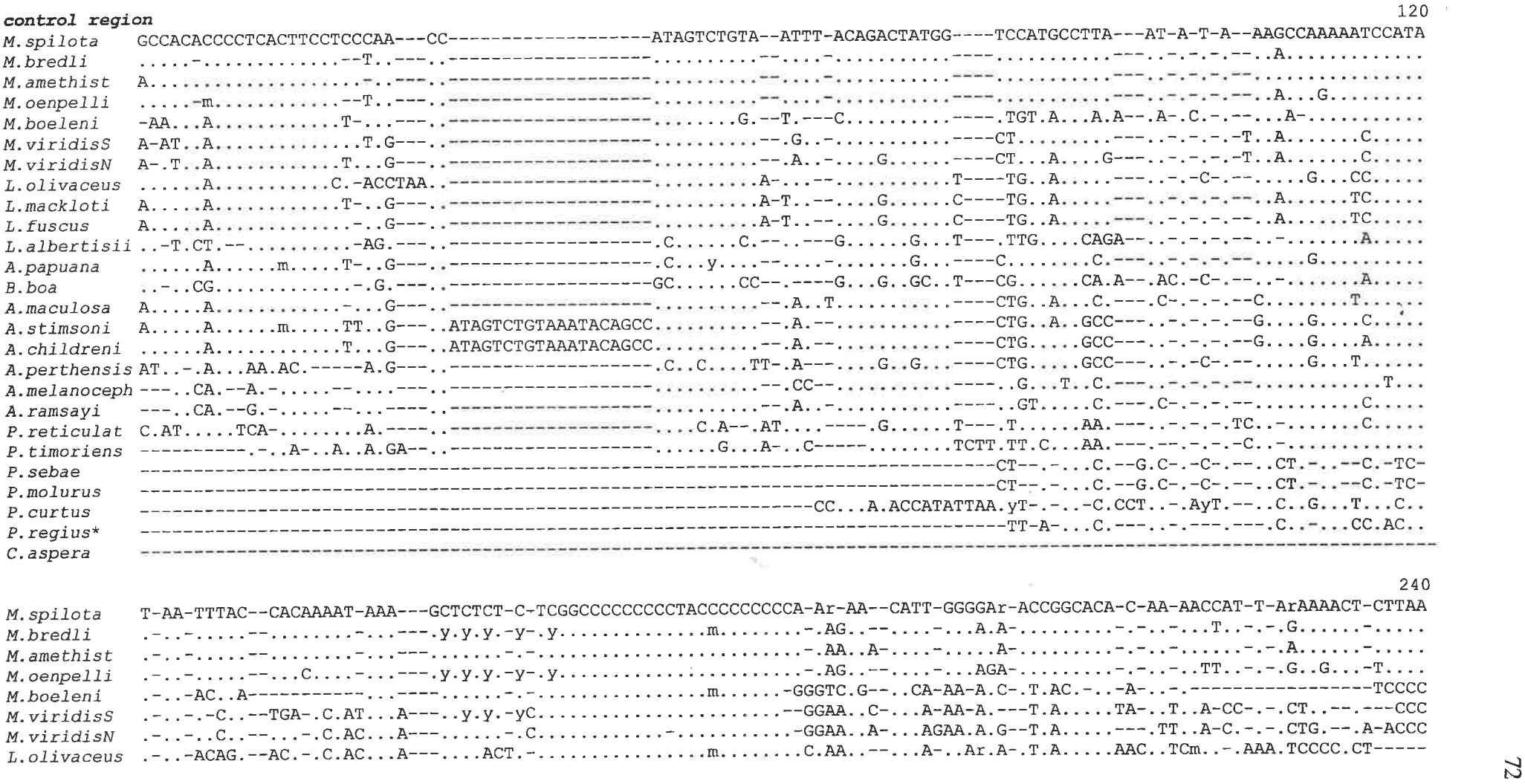
```
Ms7CRL    CTAAAATGGCCTTTGTACAAAATATTCTGTCCTCATATTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTAATCATGGATATTCTTAACCTA
Ms7CR     ....................................................................................................

Ms7CRL    AGGGTGTCTCTTAGCCTAGCGCTTCCCGTGAAATCCTCTATCCTTCCATATAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTA
Ms7CR     ....................................................................................................

Ms7CRL    CTGGCTCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCC
Ms7CR     ....................................................................................................

Ms7CRL    CTTATATGGTACATATCACCTCATGTTCTGATCATCTATGTCTATCCACCATTGGTAGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCA
Ms7CR     ....................................................................................................

Ms7CRL    CACACACAGTTAGGCATTCAGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATAT
Ms7CR     ..................................................................?????????????

Ms8CRL    CTCCTATGAGCCAATACCTCCCACTAACCCTAGCCATCTGCATACTAAACATCTCAGCCACCGCAGCATTCAATGGTACACCACCACAATGGAAGCGTGC
Ms8CR     ????????????????????????????????????????????????????????????????????????????????G.A.y.TC.TC.C..G

Ms8CRL    CCGAGACCAGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCCAACCATAGTCTGTGAATTACAGACTATGGTTGATGCCTTAA
Ms8CR     ..CCC.AA.CCAGAATTT.AA..TA.ACTACTCTTT.....................................................................

Ms8CRL    TATAAAACCAAAAATCCATATAATTTACCACAAAACAAAGCTCTCTCTCGGCCCCCCCCCTACCCCCCCCCCAAGGAACATTGGGGAGATCGGCACATAAA
Ms8CR     ..........................................y.y.y.y.y.....................................................

Ms8CRL    ACCATTArAAAACTCTCAACAAACCTCGCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTTAGCAGGAATTTCCCTTTTATTATTTTAGTC
Ms8CR     .......G.....................................................................................

Ms8CRL    TAAAATGGCCTTTGTACAAAATATTCTGTCCTCATATTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTAATCATGGATATTCTTAACCTAA
Ms8CR     ....................................................................................................

Ms8CRL    GGGTGTCTCTTAGCCTAGCGCTTCCCGTGAAATCCTCTATCCTTCCATATAATGCTAACCATTCGACTTCTCACGTCCATATTATGCwAATCCTCCCTAC
Ms8CR     ..............................................................................T.....................

Ms8CRL    TGGCTCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCC
Ms8CR     ....................................................................................................
```

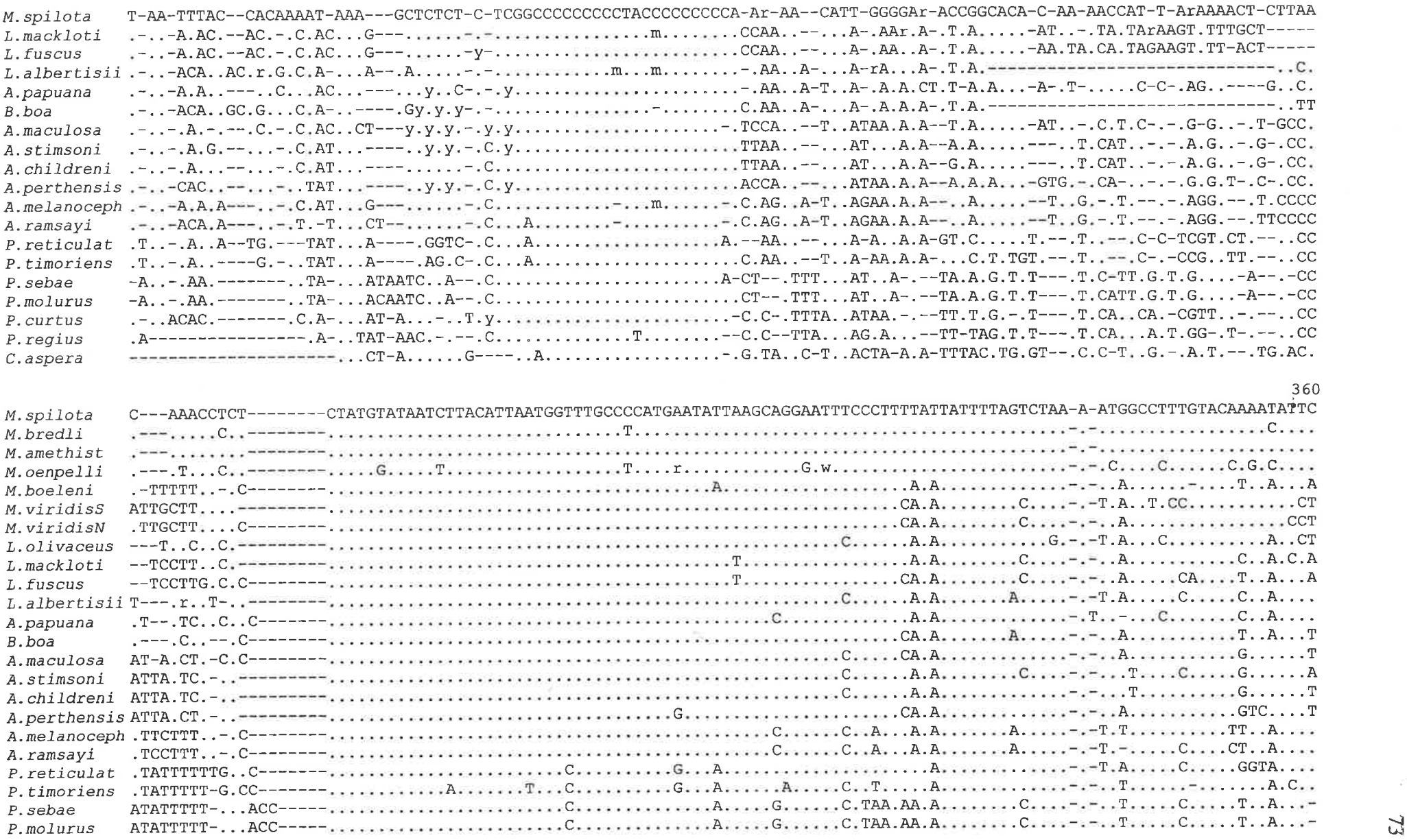

Ms8CRL TTATATGGTACATATCACCTCATGTTCTGATCATCTATGTCTATCCACCATTGGTAGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCAC
Ms8CR .....
Ms8CRL ACACACAGTTAGGCATTCAGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCC
Ms8CR .....m.....??????????????????????
Ms9CRL AGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCCCCACCATAGTCTGTAATTTACAGACTATGGTTGATGCCTTAATATAAAGCCA
Ms9CR ACCAGAATTTTAAATTAAACTACTCTTT.....
Ms9CRL AAAATCCATATAATTTACCACAAAATAAAGyTyTyTyTyGGCCCCCCCCCTACCCCCCCCCCAAGGAACATTGGGGAAACCGGCACACAAAACTATTAGAA
Ms9CR .....
Ms9CRL AACTCTTAACAAATCCCTCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTAAGCAGGAATTTCCCTTCTATTATTTTAGTCTAAAATGGCC
Ms9CR .....
Ms9CRL TTTGTACAAAATATTCTGTCCTCATCTTCTTGGTCGTTCTATGCAGCATGAGTTAACTAATCTTATTAACCATGGATATTCTTAACCTAAGGGTGTCTCT
Ms9CR .....
Ms9CRL TAGTCTAGTACTTCCCGTGAAATCCTCTATCCTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTACTGGCTCTTTC
Ms9CR .....
Ms9CRL CAAGGCCGCTGGTTACACTCTCAAGATCATCTCGATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATATGGTA
Ms9CR .....
Ms9CRL CATATCACCTCATGTTCTGATCACTTATGTCTTCCCACCATTGGTTGCTCTCTTTTTTCTGTACCTTTCATCTGACCACCATATATGCACACACACAGTT
Ms9CR .....
Ms9CRL AGGCATTTAGTCCGGGTTGAGCACGCTGTTTCCTCCTCATTTACCCCCAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTCCTTCCCGAAAT
Ms9CR .....mr..rw.....??????????????????????????????
Ms10CRL TGCCCGAGACCAGGGACTACCTTGATAGAGTAGACACAGGGCCACACCCCTCACTTCCTCCGACCATAGTCTGTAAATTACAGACTATGGTCCATGCCTT
Ms10CR TGGTCCCCAAAACCAGAATTTTAAATTAAACwACTCTTT.....
Ms10CRL AATATAAAGCCGAAAATCCATATAATTTCCCACAAAACAAAGCTCTCTCTCGGCCCCCCCCCTACCCCCCCCCCAAGAACATTGGGGAGATCGGCACACAA
Ms10CR .....y.y.y.y.y.....
Ms10CRL AACCATTArAAAACTCTTAACAAACCTCTCTATGTATAATCTTACATTAATGGTTTGCCTCATGAATATTAA?CAGGAATTTCCCTTTTATTATTTTAGT
Ms10CR .....G.....G.....

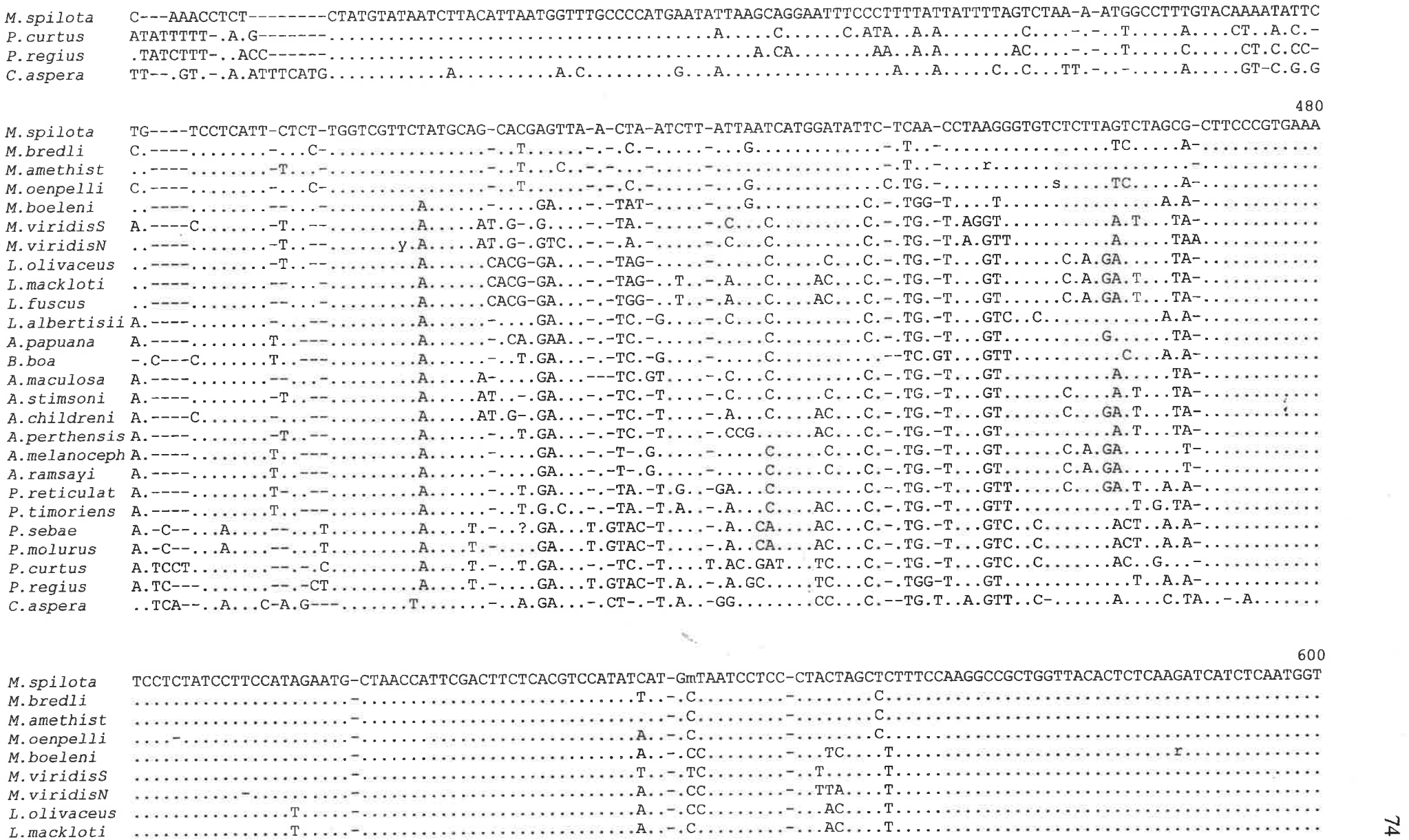
```
Ms10CRL     CTAAAATGGCCTTTGTACAAAATATTCTGTCCTCATCTTCTTGGTCGTTCTATGCAGCACGAGTTAACTAATCTTATTGATCATGGATATTCTTAACCTA
Ms10CR      ....................................................................................................

Ms10CRL     AGGGTGTyTCTTAGTCTAGCGCTTCCCGTGAAATCCTCTATCCTTCCATAGAATGCTAACCATTCGACTTCTCACGTCCATATTATGCTAATCCTCCCTA
Ms10CR      .......C............................................................................................

Ms10CRL     CTAGCCCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCGATGGTCCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCC
Ms10CR      ....................................................................................................

Ms10CRL     CTTATATGGTACATATCACCTCATGTTCTGATCACTTATGTCTATCCACCATTGGTTGCCTCTCTTTTTCTGTACCTTTCATCTGACCACCATATATGCA
Ms10CR      ....................................................................................................

Ms10CRL     CACACACAGTTAGGCATTTAGTCCGGGTTGAGCACGCTGCTTCCTCCTCATTTACCCCTAGGGATACATCTCTTAATGCTTGTTAGACATATGTATTCTC
Ms10CR      ....................................................??????????????????????
```

Appendix 3.2.

Nucleotide sequences for python taxa used in this study for partial *control region* sequences. ?=missing data; -=gap. Python regius sequence is from GenBank, see text for GenBank Accession No. Binary characters at the end of the sequence represent presnece (1) or absence (0) of the $tRNA^{Ile}$ paralogue, hairpin and ½ hairpin indel respectively.




M.spilota T-AA-TTTAC--CACAAAAT-AAA---GCTCTCT-C-TCGGCCCCCCCCCTACCCCCCCCCA-Ar-AA--CATT-GGGGAr-ACCGGCACA-C-AA-AACCAT-T-ArAAAACT-CTTAA
L.mackloti .-..-A.AC.--AC.-.C.AC...G---.....-.-.....m.....CCAA..--...A-.AAr.A-.T.A.....-AT..-.TA.TArAAGT.TTTGCT-----
L.fuscus .-..-A.AC.--AC.-.C.AC...G---.....-y-.....CCAA..--...A-.AA..A-.T.A.....-AA.TA.CA.TAGAAGT.TT-ACT-----
L.albertisii .-..-ACA..AC.r.G.C.A-...A--.A.....-.-.....m...m.....-.AA..A-...A-rA...A-.T.A.----------..C.
A.papuana .-..-A.A..---..C...AC...---...y..C-.-.y.....-.AA..A-T..A-.A.A.CT.T-A.A...-A-.T-.....C-C-.AG..----G..C.
B.boa .-..-ACA..GC.G...C.A-..----.Gy.y.y-.-.....-.....C.AA..A-...A-.A.A.A-.T.A.----------..TT
A.maculosa .-..-.A.-.--.C.-.C.AC..CT---y.y.y.-.y.y.....-.TCCA..--T..ATAA.A.A--T.A.....-AT..-.C.T.C-.-.G-G..-.T-GCC.
A.stimsoni .-..-.A.G.--...-.C.AT...----..y.y.-.C.y.....TTAA..--...AT...A.A--A.A.....---..T.CAT..-.-.A.G..-.G-.CC.
A.childreni .-..-.A...--...-.C.AT...----.....-.C.....TTAA..--...AT...A.A--G.A.....---..T.CAT..-.-.A.G..-.G-.CC.
A.perthensis .-..-CAC..--...-..TAT...----..y.y.-.C.y.....ACCA..--...ATAA.A.A--A.A.A...-GTG.-.CA-..-.-.G.G.T-.C-.CC.
A.melanoceph .-..-A.A.A---..-.C.AT...G---.....-.C.....-...m.....-.C.AG..A-T..AGAA.A.A--..A.....--T..G.-.T.--.-.AGG.--.T.CCCC
A.ramsayi .-..-ACA.A---..-.T.-T...CT--.....-.C...A.....-.....-.C.AG..A-T..AGAA.A.A--..A.....--T..G.-.T.--.-.AGG.--.TTCCCC
P.reticulat .T..-.A..A--TG.---TAT...A----.GGTC-.C...A.....A.--AA..--...A-A..A.A-GT.C.....T.--.T..--.C-C-TCGT.CT.--..CC
P.timoriens .T..-.A..----G.-..TAT...A----.AG.C-.C...A.....C.AA..--T..A-AA.A.A-...C.T.TGT.--.T..--.C-.-CCG..TT.--..CC
P.sebae -A..-.AA.-------..TA-...ATAATC..A--.C.....A-CT--.TTT...AT..A-.--TA.A.G.T.T---.T.C-TT.G.T.G....-A--.-CC
P.molurus -A..-.AA.-------..TA-...ACAATC..A--.C.....CT--.TTT...AT..A-.--TA.A.G.T.T---.T.CATT.G.T.G....-A--.-CC
P.curtus .-..ACAC.-------.C.A-...AT-A...-..T.y.....-C.C-.TTTA..ATAA.-.--TT.T.G.-.T---.T.CA..CA.-CGTT..-.--..CC
P.regius .A----------.A-..TAT-AAC.-.--.C.....T.....--C.C--TTA...AG.A...--TT-TAG.T.T---.T.CA...A.T.GG-.T-.--..CC
C.aspera ----------...CT-A.....G----..A.....-.G.TA..C-T..ACTA-A.A-TTTAC.TG.GT--.C.C-T..G.-.A.T.--.TG.AC.
360
M.spilota C---AAACCTCT--------CTATGTATAATCTTACATTAATGGTTTGCCCCATGAATATTAAGCAGGAATTTCCCTTTTATTATTTTAGTCTAA-A-ATGGCCTTTGTACAAAATATTC
M.bredli .---.....C..--------.....T.....-.-.....C....
M.amethist .---.....-------.....-.-.....
M.oenpelli .---.T...C..--------.....G.....T.....T....r.....G.w.....-.-.C....C.....C.G.C....
M.boeleni .-TTTTT..-.C--------.....A.....A.A.....-.-..A.....-....T..A...A
M.viridisS ATTGCTT....---------.....CA.A.....C....-.-T.A..T.CC.....CT
M.viridisN .TTGCTT....C--------.....CA.A.....C....-.-..A.....CCT
L.olivaceus ---T..C..C.---------.....C.....A.A.....G.-.-T.A...C.....A..CT
L.mackloti --TCCTT..C.---------.....T.....A.A.....C....-.-..A.....C..A.C.A
L.fuscus --TCCTTG.C.C--------.....T.....CA.A.....C....-.-..A.....CA....T..A...A
L.albertisii T---.r..T-..--------.....C.....A.A.....A.....-.-T.A.....C.....C..A....
A.papuana .T--.TC..C..C-------.....C.....A.A.....-.T..-...C.....C..A....
B.boa .---.C..--.C--------.....CA.A.....A.....-.-..A.....T..A...T
A.maculosa AT-A.CT.-C.C--------.....C.....CA.A.....-.-..A.....G.....T
A.stimsoni ATTA.TC.-..---------.....C.....A.A.....C....-.-...T....C.....G.....A
A.childreni ATTA.TC.-..---------.....C.....A.A.....-.-...T.....G.....T
A.perthensis ATTA.CT.-..---------.....G.....CA.A.....-.-..A.....GTC....T
A.melanoceph .TTCTTT..-.C--------.....C.....C..A...A.A.....A.....-.-T.T.....TT..A....
A.ramsayi .TCCTTT..-.C--------.....C.....C..A...A.A.....A.....-.-T.-.....C....CT..A....
P.reticulat .TATTTTTTG..C-------.....C.....G...A.....A.....-.-T.A.....C.....GGTA....
P.timoriens .TATTTTT-G.CC-------.....A.....T...C.....G...A.....A.....C..T.....A.....-.-..T.....-.....A.C..
P.sebae ATATTTTT-...ACC-----.....C.....A.....G.....C.TAA.AA.A.....C....-.-..T.....C.....T..A...-
P.molurus ATATTTTT-...ACC-----.....C.....A.....G.....C.TAA.AA.A.....C....-.-..T.....C.....T..A...-


M.spilota C---AAACCTCT--------CTATGTATAATCTTACATTAATGGTTTGCCCCATGAATATTAAGCAGGAATTTCCCTTTTATTATTTTAGTCTAA-A-ATGGCCTTTGTACAAAATATTC
P.curtus ATATTTTT-.A.G-------.....A.....C......C.ATA..A.A.....C....-.-..T.....A....CT..A.C.-
P.regius .TATCTTT-..ACC------.....A.CA.....AA..A.A.....AC....-.-..T.....C.....CT.C.CC-
C.aspera TT--.GT.-.A.ATTTCATG.....A.....A.C.....G...A.....A...A.....C..C...TT.-..-.....A.....GT-C.G.G
480
M.spilota TG----TCCTCATT-CTCT-TGGTCGTTCTATGCAG-CACGAGTTA-A-CTA-ATCTT-ATTAATCATGGATATTC-TCAA-CCTAAGGGTGTCTCTTAGTCTAGCG-CTTCCCGTGAAA
M.bredli C.----.....-...C-.....-..T.....-.-.C.-.....-...G.....-.T..-.....TC.....A-.....
M.amethist ..----.....-T...-.....-..T...C..-.-...-.....-.....-.T..-....r.....-.....
M.oenpelli C.----.....-...C-.....-..T.....-.-.C.-.....-...G.....C.TG.-.....s.....TC.....A-.....
M.boeleni ..----.....--...-.....A.....-....GA...-.-TAT-.....-...G.....C.-.TGG-T....T.....A.A-.....
M.viridisS A.----C.....-T..--.....A.....AT.G-.G....-.-TA.-.....-.C...C.....C.-.TG.-T.AGGT.....A.T...TA-.....
M.viridisN ..----.....-T..--.....y.A.....AT.G-.GTC..-.-.A.-.....-.C...C.....C.-.TG.-T.A.GTT.....A.....TAA.....
L.olivaceus ..----.....-T..--.....A.....CACG-GA...-.-TAG-.....-.....C.....C...C.-.TG.-T...GT.....C.A.GA.....TA-.....
L.mackloti ..----.....--...-.....A.....CACG-GA...-.-TAG-..T..-.A...C....AC...C.-.TG.-T...GT.....C.A.GA.T...TA-.....
L.fuscus ..----.....--...-.....A.....CACG-GA...-.-TGG-..T..-.A...C....AC...C.-.TG.-T...GT.....C.A.GA.T...TA-.....
L.albertisii A.----.....-...--.....A.....-....GA...-.-TC.-G....-.C...C.....C.-.TG.-T...GTC..C.....A.A-.....
A.papuana A.----.....T..---.....A.....-.CA.GAA..-.-TC.-.....-.....C.....C.-.TG.-T...GT.....G.....TA-.....
B.boa -.C---C.....T..---.....A.....-..T.GA...-.-TC.-G....-.....C.....--TC.GT...GTT.....C...A.A-.....
A.maculosa A.----.....--...-.....A.....A-....GA...---TC.GT....-.C...C.....C.-.TG.-T...GT.....A.....TA-.....
A.stimsoni A.----.....-T..--.....A.....AT..-.GA...-.-TC.-T....-.C...C.....C...C.-.TG.-T...GT.....C....A.T...TA-.....
A.childreni A.----C.....-...--.....A.....AT.G-.GA...-.-TC.-T....-.A...C....AC...C.-.TG.-T...GT.....C...GA.T...TA-.....
A.perthensis A.----.....-T..--.....A.....-..T.GA...-.-TC.-T....-.CCG.....AC...C.-.TG.-T...GT.....A.T...TA-.....
A.melanoceph A.----.....T..---.....A.....-....GA...-.-T-.G.....-.....C.....C...C.-.TG.-T...GT.....C.A.GA.....T-.....
A.ramsayi A.----.....T..---.....A.....-....GA...-.-T-.G.....-.....C.....C...C.-.TG.-T...GT.....C.A.GA.....T-.....
P.reticulat A.----.....T-..--.....A.....-..T.GA...-.-TA.-T.G..-GA...C.....C.-.TG.-T...GTT.....C...GA.T..A.A-.....
P.timoriens A.----.....T..---.....A.....-..T.G.C..-.-TA.-T.A..-.A...C....AC...C.-.TG.-T...GTT.....T.G.TA-.....
P.sebae A.-C--...A....--...T.....A....T.-..?.GA...T.GTAC-T....-.A..CA....AC...C.-.TG.-T...GTC..C.....ACT..A.A-.....
P.molurus A.-C--...A....--...T.....A....T.-....GA...T.GTAC-T....-.A..CA....AC...C.-.TG.-T...GTC..C.....ACT..A.A-.....
P.curtus A.TCCT.....--.-.C.....A....T.-..T.GA...-.-TC.-T....T.AC.GAT..TC...C.-.TG.-T...GTC..C.....AC..G...-.....
P.regius A.TC---.....--.-CT.....A....T.-....GA...T.GTAC-T.A..-.A.GC.....TC...C.-.TGG-T...GT.....T..A.A-.....
C.aspera ..TCA--..A...C-A.G---.....T.....-..A.GA...-.CT-.-T.A..-GG.....CC...C.--TG.T..A.GTT..C-.....A....C.TA..-.A.....
600
M.spilota TCCTCTATCCTTCCATAGAATG-CTAACCATTCGACTTCTCACGTCCATATCAT-GmTAATCCTCC-CTACTAGCTCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGT
M.bredli .....-.....T..-.C.....-.....C.....
M.amethist .....-.....-.C.....-.....C.....
M.oenpelli ....-.....-.....A..-.C.....-.....C.....
M.boeleni .....-.....A..-.CC.....-...TC....T.....r.....
M.viridisS .....-.....T..-.TC.....-..T.....T.....
M.viridisN .....-.....-.....A..-.CC.....-..TTA....T.....
L.olivaceus .....T.....-.....A..-.CC.....-...AC....T.....
L.mackloti .....T.....-.....A..-.C.....-...AC....T.....

```
M.spilota     TCCTCTATCCTTCCATAGAATG-CTAACCATTCGACTTCTCACGTCCATATCAT-GmTAATCCTCC-CTACTAGCTCTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGT
L.fuscus      ..............T.....-.............................A..-.C.........-...AC....T.....................................
L.albertisii  C.............CT.....-.............................A..-.C........T-.....G.............s.....................................
A.papuana     ..............T.....-.............................G.AT.-.CC........-...AC....T.....................................
B.boa         ..............CT.....-.............................A..-.C.........-.....G.-.T.....................................
A.maculosa    ..................-.............................A..-.CC........-...TC....T.....................................
A.stimsoni    ..................-.............................AG.-.T.G......T--..A.....T.....................................
A.childreni   ...........T.......-.............................AG.-.T.G......T--..A.....T.....................................
A.perthensis  ............C.....-.............................A..-.C.........-...AA....T.....................................
A.melanoceph  ............TCT...-.............................T..-.CC........-...AC....T.....................................
A.ramsayi     ............TCT...-.............................CG.T..-.CC........-...AC....T.....................................
P.reticulat   ............GCGT..-.............................-.A..T.C..C.....T-T....G.-.T.....................CT.....................
P.timoriens   ............GCGT..-.............................-.A..T.C..T.....T-T....G.-.T.....................CT.....................
P.sebae       ...........T....C.-..-.............................-AG.-.C..CC....TT..-..T.....................................G.C...
P.molurus     ...........T....C.-..-.............................-AG.-.C..CC....TT..-..T.....................................G.C...
P.curtus      ...........T....C.-..-.............................C.-TT.-.C..CC....TT..-..T.....................................G.C...
P.regius      ...........T....C.-..-.............................C.-AG.-.C..CC.....T..-..T............A.....................G.....
C.aspera      ..............-......G.....AG.C.TG...T............-T..-.C..CC.....TT..-.T..A............A........T..T.....TC........C...

                                                                                                                                  720
M.spilota     CCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATAT-GGTACATATCACCTCATGTTCTGATCATCTATGTCTAT-CCACCATTGGTAGCTCTCTTTTT
M.bredli      ..................................................T.....-.............................CT.........-...........T.........
M.amethist    ..................................................-.....-.............................C-...........T.........
M.oenpelli    ..................................................T.....-.............................CT.........-...........T.........
M.boeleni     ................C.................................T.....-.......C.....................A.C-......C..........T.....
M.viridisS    ................C.................................T.....-.......C.TGT.............C.....CTAG.-......C....TT.C..T.....
M.viridisN    ................C.................................T.....-.......C.TGT.............C......TAG.-..G...C....TT.C..T.....
L.olivaceus   ..................................................T.....-.......T.................G.......A..-......C..........T....-
L.mackloti    ..................................................T.....-.......T.................G......TAC.-......C.....T.C..T....-
L.fuscus      ..................................................T.....-.......T.................G......TAT.-......C.....T.C..T....-
L.albertisii  ..................................................T.....-......CT.................G.......A..-..G..T.....TT....T.....
A.papuana     ................C.................................T.....-......CT.................G.......A..-......C..............-
B.boa         ..................................................T.....-.......T.TGT.....................C...C-..T.........TTG...T....-
A.maculosa    ...............T..................................T.....-...T...T.TGT.............C.....CT...-......C..........T....-
A.stimsoni    ..................................................T.....-...T...T.AGT.............G.......A..-......C..........T....-
A.childreni   ..................................................T.....-...T...T.AGT.............G.......A..-......C..........T....-
A.perthensis  ..................................................T.....-...T...TCTTT.............G.......A..-...........T..C..T....-
A.melanoceph  ..................................................T.....-.......T.TG..............G......TAT.-......C.....T.C..T....-
A.ramsayi     .........T............C.......C.........T......T.....-......CTCTG..............G......TAT.-......C.....T.C..T....-
P.reticulat   ..................................................-.....-.....................C.....CTAGA-......C..........T....-
P.timoriens   ..................................................-.....-.......T.TG............C.....CTAGA-.....TC..........T....-
P.sebae       ..................................................T.....-.......T.TG..G..........C.....CTAG.-.A....C.....T.C..T....-
P.molurus     ..................................................T.....-.......T.TG..G..........C.....CTAG.-.A....C.....T.C..T....-
P.curtus      ......................C...........................-.....-.......T.TG.T.........G...C.....CTAG.-.A....C....TT.C..T.....
```

```
M.spilota    CCGGAACCACCCCTCCATACTAGCTTTTTCCAAGACCTTTGGTCGCACCCCTTATAT-GGTACATATCACCTCATGTTCTGATCATCTATGTCTAT-CCACCATTGGTAGCTCTCTTTTT
P.regius     .....................................................T......-......AT.....................G.....C.AG.-......C.....T.C..T....-
C.aspera     .........T.....T.....................................T......T..CG..C-.G.T......A........C.....C.AG.T.A....C.....A.-..T.....

                                                                                                                          768
M.spilota    -TCTGTACCTTTCATCTGACCACCATATATGCACACACACAGTTACAC 101
M.bredli     -..................................................... ...
M.amethist   -..................................................... ...
M.oenpelli   -..................................................... ...
M.boeleni    C.................T...........T..................... ...
M.viridisS   C..G..............T.......................A.... ...
M.viridisN   C.................T.......................A.... ...
L.olivaceus  C.................T.......................A.... ...
L.mackloti   C.................T........................... ...
L.fuscus     C.................T.......................A.... ...
L.albertisii -C................T.......................-... ...
A.papuana    C.................T.......................A.... ...
B.boa        C.................T.......................A-... ...
A.maculosa   C.................T.......................A.... ...
A.stimsoni   C.................T.......................A..C. .1.
A.childreni  C.................T.......................A..C. .1.
A.perthensis C.................T.......................A.... ...
A.melanoceph C.................T.......................A.... ...
A.ramsayi    C....-...........?T.......................A.... ...
P.reticulat  C...C.C.......C.....T.......................A.... ...
P.timoriens  C...C.C.......C.....T.......................A.... ...
P.sebae      C...-.C.......C.....T.........................A.A 0.0
P.molurus    C...C.C.......C.....T.........................A.A 0.0
P.curtus     C.............................................A.A 0.0
P.regius     C...C.C.......C.....T.........................A.A 0.0
C.aspera     C...C.........C.....AC.........TTG...GGGG.GG.A.A 0.0
```

Chapter 4.

Python phylogenies: comparisons from morphology and mtDNA.

4.1 Introduction

Snakes can be traced in the fossil record for about 130 MYA, extending to the early Cretaceous (Barremian) period (Rage and Richter, 1994) and it is generally accepted that snakes originated from lizards. However, the precise relationship of snakes to other squamates has been much debated. Recent morphological (Caldwell, 1998; Lee *et al.*, 1999; Lee *et al.*, 1998) and molecular evidence (Saint *et al.*, 1998) suggest that the varanoid lizards, especially the extinct marine mosasaurs, are the likely sister group to snakes. From their lizard-like beginnings, snakes diversified into several major lineages. In recent times, many workers have used a three-way subdivision of snakes into the Scolecophidia (blind snakes), the Henophidia ("primitive" snakes) and the Caenophidia ("advanced" snakes). Henophidia, based mainly on the boas and pythons, is clearly a paraphyletic taxon, as it has been generally agreed that the caenophidian snakes are derived from a 'henophidian' ancestor (Rage, 1987).

Whilst acknowledging that a well-corroborated phylogenetic hypothesis for all snakes does not yet exist, McDiarmid *et al.* (1999) proposed a classification that can be regarded as a consensus arrangement of the major snake taxa. They subdivided the "Henophidia" by removing fossorial taxa to a redefined Alethinophidia and placing the remaining Henophidia, including boas and pythons, in the Macrostomata. Scanlon and Lee (2000) proposed that the Macrostomata comprises a group of relatively plesiomorphic families that lie at the base of the Caenophidian radiation. McDiarmid *et al.* (1999) placed pythons in their own family, Pythonidae, while Scanlon and Lee (2000) treated pythons as a subfamily, Pythoninae, of the Boidae. I use the informal terms 'pythons' and 'boas' for Pythonidae and Boidae respectively

(Boidae, McDiarmid *et al.*, 1999; or Boinae, Scanlon *et al.*, 1999).

Pythons range from western and southern Africa, through the tropical rainforests of Southeast Asia, eastwards as far as New Guinea and into the cooler climates of southern Australia (Fig. 4.1). They differ from the generally similar boas in reproductive mode and anatomically by the presence in pythons of a novel bone, the supraorbital, on the dorsal margin of the orbit. Pythons are oviparous (egg layers) and the females coil around the eggs to incubate them until they hatch (McDowell, 1987). Pythons can be terrestrial, arboreal, fossorial or semi-aquatic and vary in size from 0.5m *e.g. Antaresia perthensis* (Smith and Plane, 1985) to approximately 10m in length *e.g. Python reticulatus* (Minton and Minton, 1973).

4.1.2 Taxonomy

Most python genera and species are restricted to the Australian region. Of the eight genera recognised by Kluge (1993), only the genus *Python* is not found in Australia or New Guinea. Of the other seven genera, three (*Leiopython, Liasis* and *Morelia),* are found in Australia and New Guinea and two each are restricted to Australia (*Antaresia* and *Aspidites*) or New Guinea (*Apodora* and *Bothrochilus*).

The species composition of the subfamily Pythoninae has also been controversial (Cogger *et al.*, 1983; Kluge, 1993; McDowell, 1975; Smith, 1985; Storr *et al.*, 1986; Stuhl, 1932; Underwood, 1976; Underwood and Stimson, 1990 – see Fig. 4.2). In particular, *Liasis fuscus* and *Morelia bredli* have not been recognised as species by some authors (Fyfe, 1990; Smith, 1985). I follow Barker & Barker (1994) in recognising 27 described extant species with the inclusion of *Liasis fuscus* and *Morelia bredli* and an unnamed sibling taxon of *M. viridis* from northern New Guinea (see chapter 5).

The comprehensive morphological studies of McDowell (1975), Underwood & Stimson (Underwood and Stimson, 1990) and Kluge (1993), who each used an array of different characters in their analyses, each reached different conclusions about the

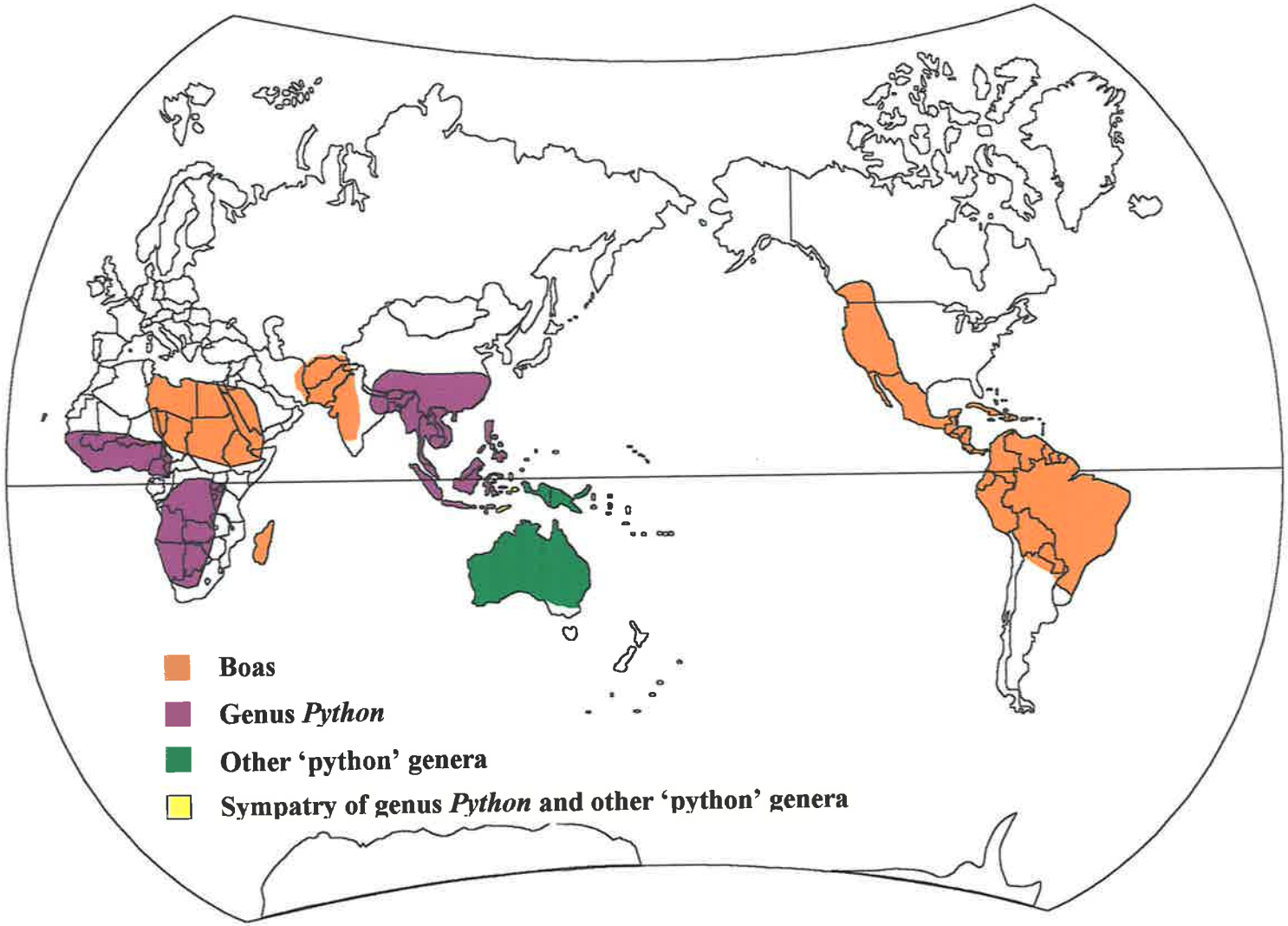


Fig. 4.1. World map showing the distributions of boas and pythons. The regions of sympatry shown in yellow are Timor for Python *reticulatus*, P. *timoriensis* and *Liasis mackloti* and Ambon, the Banda Isalnds, Halmahera, Seram and Tanimbar for *P. reticulatus* and *Morelia amethistina*.

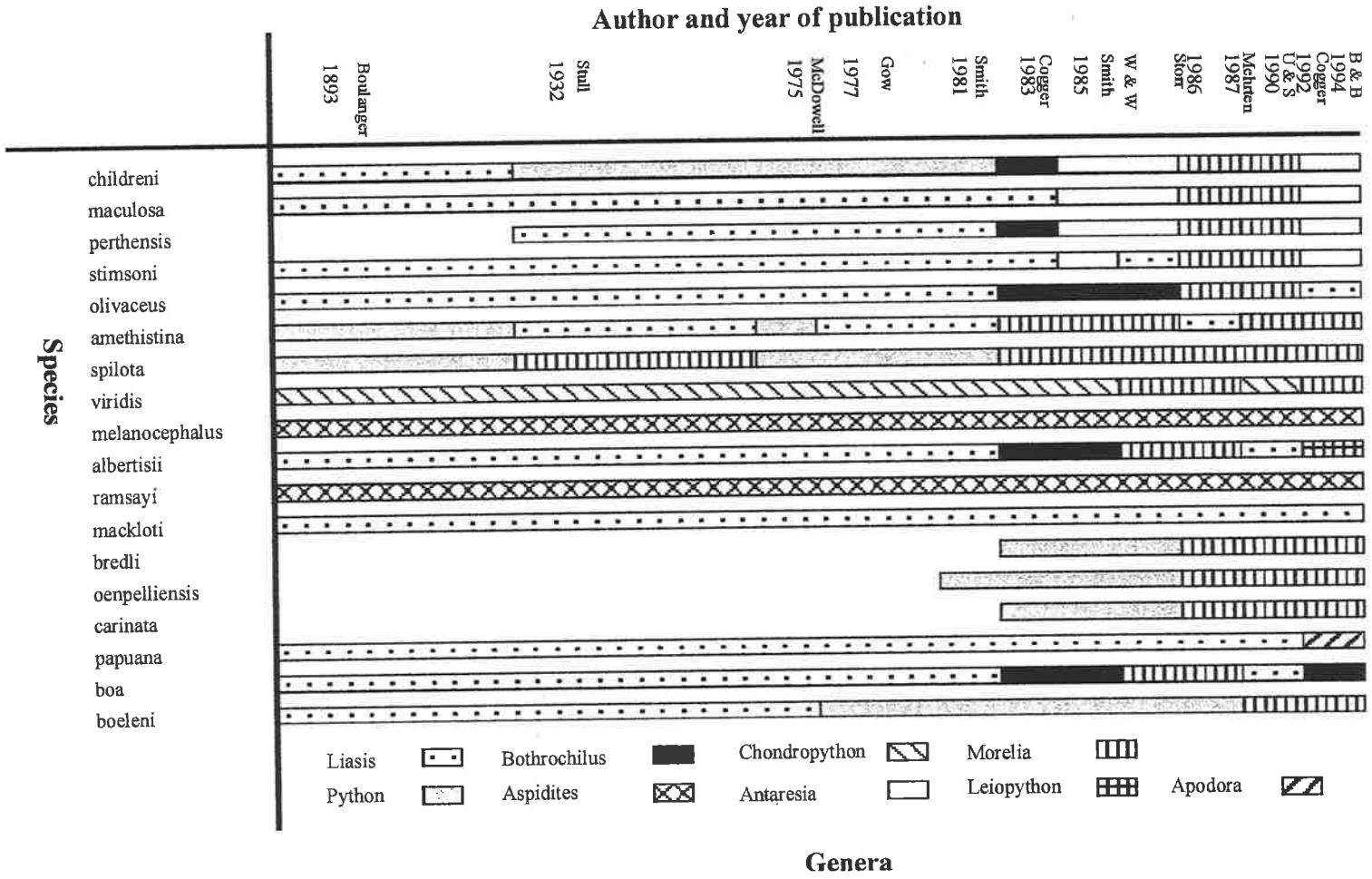


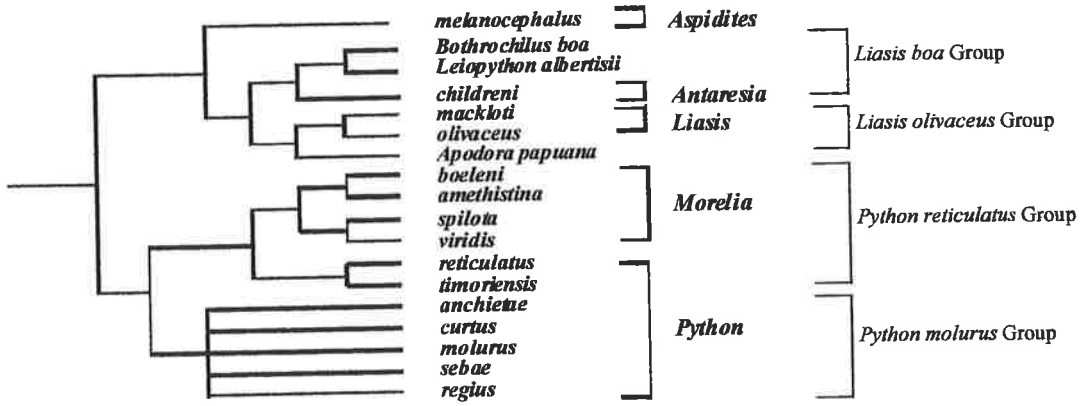
Fig. 4.2 A graphic representation of the generic allocations of Australo-New Guinean python species from their first description to the present. Author abbreviations are: B & B = Barker & Barker; U & S = Underwood & Stimson; W & W = Wells & Wellington.

relationships among species (Fig. 4.3). The cladogram of python relationships determined from McDowell's discussion of species affinities shown in Fig. 4.3a has been adapted from Kluge (1993). Strikingly, from these three studies, there is concordance only in the monophyly of the genus *Morelia*, as outlined below.

McDowell (1975) classified the pythons into two major lineages (Fig. 4.3a). In one group, *Aspidites* is a sister lineage to a clade supporting two distinct groups within the genus *Liasis*, one group comprising *mackloti*, *olivaceus* and *papuana* (called the *Liasis olivaceus* Group) and the other group comprising *albertisii*, *boa* and *childreni* (called the *Liasis boa* Group). A second lineage contains two clades, one comprising *Python reticulatus* and *P. timoriensis* plus *Morelia amethistina, M. boeleni, M. spilota* and *M. viridis* and the other comprising *Python anchietae, P. curtus, P. molurus, P. regius* and *P. sebae.*

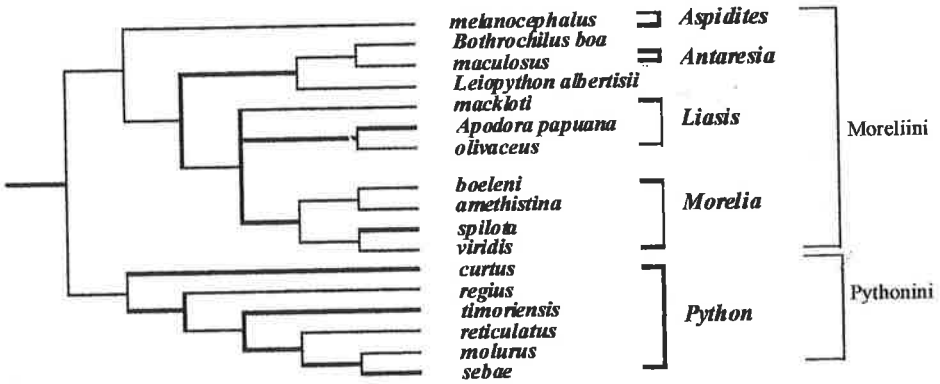
Underwood & Stimson (1990) divided the pythons into two major tribes, the Pythonini from Afro-Asia, limited to the genus *Python*, and the Moreliini from Australia and New Guinea (Fig. 4.3b). Characters used to define these two groups were the number of neck scales being greater than 48 as a synapomorphy for *Python* and intercostal arteries arising in pairs rather than singly as a synapomorphy for the Australo-Papuan pythons. Furthermore, among the Moreliini, *Aspidites* was the sister lineage to the remaining taxa of that tribe. There is agreement between McDowell's and Underwood & Stimson's studies in the relationships within *Morelia* with the species pairs *boeleni/amethistina* and *spilota/viridis* being each other's closest relatives.

Kluge (1993) used a set of 121 morphological characters, incorporating some of the characters used by McDowell and Underwood & Stimson, to understand further the relationships of pythons. Kluge's (1993) analysis differed from the other analyses in placing *Aspidites* as a sister taxon to all other pythons and in placing *Python* well within the Australo-Papuan clade (Fig. 4.3c). Kluge placed *Morelia* as the sister group to *Python*, but was not able to define the relationships within *Python.* This differed from McDowell who placed *Python* paraphyletic with *Morelia* (Fig. 4.3a). Underwood & Stimson (1990) placed *Python*



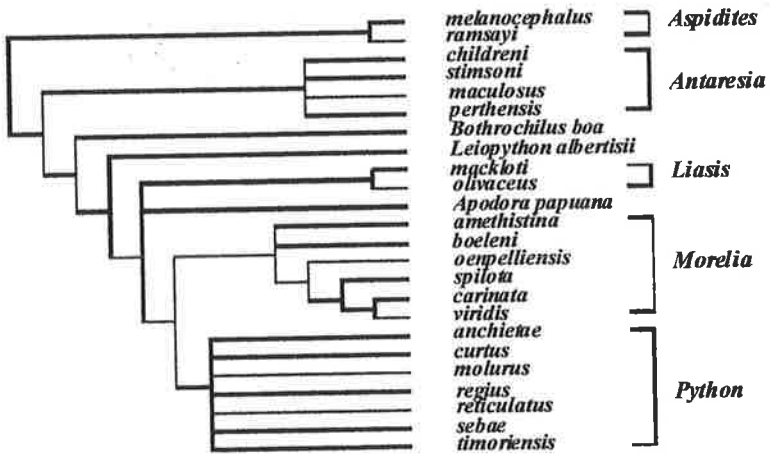
a

(McDowell, 1975)



b

(Underwood & Stimson, 1990)



c

(Kluge, 1993)

Fig. 4.3. a) Python relationships determined from McDowell's (1975) discussion of species groups. Brackets on right show McDowell's assignment of Species Groups, with the exception of *viridis*, which is assigned to the genus *Chondropython*. b) Python relationships determined from Underwood & Stimson's (1990) morphological analysis of within and between Afro-Asian and Australo-New Guinea pythons, showing the division between the Moreliini and the Pythonini. c) Python relationships based on morphological analysis of Kluge (1993).

as the sister group to all other pythons (Fig. 4.3b). Kluge's argument with McDowell's (1975) and Underwood & Stimson's (1990) placement of *Aspidites* is based on the latters' *a priori* assumption that the absence in *Aspidites* of some characteristics (*e.g.* thermoreceptive labial pits) is due to secondary loss as opposed to never having them in the first place, as well as a secondary acquisition of the horizontal part of the nasal bone lying above the nostrils.

A general conclusion from investigators of python systematics is that there is little morphological divergence within the python group (Schwaner and Dessauer, 1981; Smith, 1981a; Smith, 1981b; Smith, 1985; Storr *et al.*, 1986; Underwood and Stimson, 1990), however, Kluge (1993) felt that the lack of differentiation was most likely due to the incorporation of too few characters and non-parsimony methods of analysis. The ambiguous generic allocation of some species provides a further indication of the relatively low level of morphological divergence within the lineage. For example, McDowell (1975) synonymised *Bothrochilus* with *Liasis*, whilst Storr *et al.* (1986) synonymised *Liasis* and *Chondropython* with *Morelia*. McDowell (1975) felt that whilst *Aspidites* could be distinguished well from other pythons by its lack of labial pits, the forward extension of the horizontal processes of the nasal bone and entire subcaudal scales, morphology could only weakly define most other genera. McDowell (1975) had difficulty in making a generic distinction between *amethistina* and *reticulatus* and commented that *timoriensis* appeared to be an intermediate. McDowell (1975) recognised *Morelia* (*spilota*) and *Chondropython* (*viridis*) as monotypic genera but felt that that there was little justification for recognising *spilota* as non-*Python*. Qualitative comparisons of transferrins by Schwaner and Dessauer (1981) also found that there was little detectable divergence among the Australian taxa compared, although there was a pronounced drop in cross-reactivity when Australian taxa were compared with *Python*. Such morphological and molecular conservatism may be due to recent origins for many extant python taxa, or may reflect a slow rate of evolution, molecular and morphological. Kluge (1993) suggested that the inability to distinguish between subgroups of pythons might be due to insufficient number of characters examined.

Biogeographic implications that flow from hypotheses suggesting slow rates of evolution are: 1) that pythons arose in Laurasia and moved southwards through the Asian land bridge into Australasia (supported by McDowell's, 1975 and Underwood and Stimson's, 1990 systematic hypotheses) or 2) that pythons arose in Gondwana and moved northwards into Eurasia. Python fossils may be present in Eocene deposits in Europe (Szyndlar and Böhme, 1993) and are known from early-middle Miocene deposits in Europe, the Middle East and central Asia (Hoffstetter, 1964; Rage, 1987; Rage, 1976). Gondwanan fossils of *Liasis* and *Morelia,* are known from Queensland, Australia in deposits of Miocene age (Kluge, 1993; Scanlon *et al.*, 1999; Smith and Plane, 1985). Thus, the fossil record does not unequivocally support either hypothesis.

Given the very comprehensive nature of Kluge's (1993) morphological study, it is appropriate to look for other independent measures of relatedness that could be used to determine evolutionary relationships between python species. Analyses of nucleotide sequences and relative evolutionary rate tests provide tools with which alternative evolutionary scenarios can be tested. I present sequence data from two mitochondrial genes with differing rates of molecular evolution, the faster-evolving *control region* (*CR*) and the more medium- to fast-evolving coding gene, *cytochrome b* (*cytb*). I also considered Kluge's (1993) data to evaluate the characters that were used to distinguish *Aspidites* from other pythons and reanalysed the data with an expanded set of outgroups.

4.2 Materials & Methods

4.2.1 Specimens examined

Genomic DNA was extracted as per chapter 2.2. Individuals included in this study are: ($^{+}$ = extractions enriched for mitochondrial DNA). *Antaresia childreni* SAMA R21411^{+}; *A. maculosa* ABTC 68227; *A. perthensis* ABTC 68276; *A. stimsoni* SAMA R38794; *Apodora papuana* ABTC 68240; *Aspidites melanocephalus* ABTC 68246; *A. ramsayi* SAMA R19831^{+}; *Bothrochilus boa* AMS R129533; *Leiopython albertisii* AMS R124481^{+}; *Liasis*

mackloti SAMA R21422^{+}; *Liasis fuscus* ABTC 68263; *Liasis olivaceus* ABTC 6503; *Morelia amethistina* AMS R115347^{+}; *M. boeleni* BPBM 11611; *M. bredli* ABTC 68339; *M. oenpelliensis* ABTC 68277; *M. spilota* SAMA R26878^{+}; *M. viridis* AMS R115348^{+}, (Southern PNG), BPBM 11617, (Northern PNG); *Python curtus* ABTC 24797; *P. molurus* ABTC 67159; *P. reticulatus* SAMA R28533; *P. sebae* SAMA R26137 and *P. timoriensis* ABTC 68326. Boinae (outgroup) *Candoia aspera* AMS R115337^{+}. Because of the unavailability of material, *Morelia carinata* and *Python anchietae* were not included in the present study. *Python regius* sequences were retrieved from GenBank (see below).

4.2.2 PCR primers and conditions

The Polymerase Chain Reaction (PCR) was used to amplify partial *transfer RNA^{Thr}/control region* and *cytochrome b* gene products. All primer sequences are listed in Table 2.1. To preferentially PCR amplify the *CR* instead of the *control region-like* gene (Kumazawa *et al.*, 1996), nested PCR was used as per chapter 2.8. Two overlapping partial *cytb* products of approximately 300 bp and 900 bp were amplified using primers L14841 and H15149 (Kocher *et al.*, 1989) for the short product and either L14973 or Snake12 (L) with H15916 (Kumazawa *et al.*, 1996) for the longer product. PCR products for *cytb* and *CR* were sequenced for both strands with the PCR primers and also with nested primers Snake 1 (L), Snake 6 (L) and Snake 7 (H) for the *CR* products. The potential for each of the mitochondrial primer pairs to amplify mitochondrial genes rather than nuclear paralogues was tested as per chapter 2.4. Enriched mitochondrial DNA (mtDNA) and total cellular DNA of *Liasis mackloti* SAMA R21422, *Morelia spilota* SAMA R26878 and *M. viridis* AMS R115348 were used to test the whether the *CR* primers amplified mitochondrial genes. *M. spilota* SAMA R26878, *M. viridis* AMS R115348 and *C. aspera* AMS R115337 were used to test the *cytb* primers.

CR sequences were initially aligned in Clustal W (Thompson *et al.*, 1994) under varying gap penalties and insertions and deletions (indels) were incorporated to optimise the

alignments. Regions of sequence alignment that varied under differing gap penalties were considered to be of ambiguous alignment and were excluded from the final analyses.

Incorporated in this region of ambiguous sequence alignment are three structural features of the *CR* (see Chapter 3, Fig. 3.5,6) which were coded as binary characters and included in the analysis. The first feature is an indel at the 5' end of the region adjacent to the $tRNA^{Pro}$ gene, that is approximately 20 bp of the amino acid acceptor stem and the TψC arm of the $tRNA^{Ile}$ gene. The second feature is a 15 bp hairpin found adjacent to the isoleucine paralogue. The third feature is a 15 bp partial repeat of the hairpin in the region 5' to the hairpin that is present in *Antaresia childreni* and *A. stimsoni*. A parsimony analysis of character state change of the evolution of each of these characters was performed using McClade version 3.04 (Maddison and Maddison, 1992).

Complete *cytb* gene sequences were retrieved from GenBank for *Aspidites melanocephalus*, *Antaresia childreni*, *Apodora papuana*, *Liasis mackloti*, *L. olivaceus*, *Leiopython albertisii*, *Morelia amethistina*, *M. spilota*, *M. viridis*, *Python molurus, P. regius*, *P. reticulatus* and *P. sebae* (Acc. Nos. U69741, 751, 760, 835, 837, 839, 842, 843, 847, 851, 853, 857, 860, 863). The *M. spilota* sequence U69851 had a stop codon present due to an autapomorphy in a first codon position, this nucleotide position was coded as missing for the present study. Due to sequencing difficulties, only partial *cytb* sequence (337 bp) was used for *Liasis fuscus*. Because it has been shown that in snakes the *control region-like* sequence that is present between the *ND1* and *ND2* genes is typically indistinguishable from the *control region* sequence for the region of the *CR* sequenced here (Kumazawa *et al.*, 1996; Kumazawa *et al.*, 1998 see chapter 3), and *CR* sequence for *Python regius* was not available, I used the published *control region-like* sequence (GenBank Acc. No. D84258) in the *CR* data set. Aligned nucleotide sequences for *CR* and *cytb* are presented in Appendices 3.2 and 4.1 respectively.

4.2.3 Morphological characters

Kluge (1993) found 16 morphological characters to support the sister relationship of *Aspidites* to Group C (all other pythons). A summary of the 16 characters used by Kluge (1993) to define *Aspidites* can be found in Appendix 4.2.

Assistance in helping me to understand the morphological characters used by Kluge (1993) was kindly given by Dr. M. Hutchinson. I examined preserved python skulls in the South Australian Museum collection for 11 of these 16 characters in order to assess the reliability, *i.e.* non-ambiguity, with which the character states could be determined. Taxa that were examined from preserved skulls were: *Antaresia childreni*, *Aspidites ramsayi*, *A. melanocephalus*, *Boa constrictor*, *Candoia carinata*, *Liasis fuscus*, *Morelia amethistina, M. spilota, M. viridis, Python molurus,* and *P. reticulatus*. I considered each character with respect to four questions. 1. Is the character correctly scored? That is, can the character states be verified in other specimens? 2. Are the character states homologous and anatomically identical? 3. Is the character continuous and therefore have the character state definitions been subjectively determined? 4) Are the characters correlated in any way? That is, could the character state of one character be influencing the character state of another character because, for instance, they form part of a common structure?

Characters that answered yes to 1 and 2, but no to 3 and 4 were considered to be strong candidate characters to use in considering the relationships of *Aspidites* to other pythons. Phylogenetic support for those strong characters that define the *Aspidites* lineage could then be tested using the Bremer decay index.

I reanalysed Kluge's (1993) full data set with a modified outgroup comprising a set of five outgroup taxa each as an individually coded taxon (compiled by S. Donnellan) rather than using Kluge's common ancestor approach. The five outgroups (after Kluge) are: 1. boids; 2. erycines and "advanced snakes" which includes tropidophiines, bolyeriines, *Acrochordus* and "higher snakes = colubroids per Marx & Rabb (1970)"; 3. *Loxocemus*; 4. *Xenopeltis* and 5. anilioids which includes *Anilius*, *Cylindrophis* and the uropeltines. Reliability of the resulting

tree topology was evaluated by bootstrapping, from 2000 pseudoreplicates.

4.2.4 Phylogenetic Analysis

Data partitions were tested for combinability with the ILD test of Farris *et al.* (1995) by using the Partition Homogeneity Test in PAUP*4.0b2a, (Swofford, 1999) (see chapter 2.12.1). I adopted an α of 0.01 as the combinability criterion for the ILD test (Cunningham, 1997). Aligned sites were phylogenetically analysed with PAUP*4.0b2a.

Parameters such as transition/transversion (ts/tv) ratio, the shape of a gamma distribution describing rate variation between sites (Γ), the proportion of invariable sites (I) and substitution rates were estimated for the nucleotide data by maximum likelihood (ML) in PAUP*4.0b2a. Modeltest3.0 (Posada and Crandall, 1998) was used to determine the appropriate molecular evolution models for ML and distance analyses of the molecular data. For MP analysis, gaps were treated as a fifth character state. The robustness of the trees was evaluated by bootstrap analysis from 2000 "fast" pseudoreplicates for MP and NJ analyses and 200 for ML analysis. Support for the internal nodes of branches of MP trees were determined from decay indices (= branch support of Bremer, 1994). Decay indices for individual partitions were estimated in TreeRot v.2 (written by M.Sorenson *et al.*, 1999).

Saturation of nucleotide substitutions was examined by plotting the genetic distance as determined by the selected model of evolution against the uncorrected *p*-distance for *cytb* and *CR* for all substitutions, transitions only and transversions only and for first and third codon positions for *cytb*. There were too many comparisons with no transversions to obtain distances for second positions. A linear relationship between the two distances compared suggests that the nucleotide substitutions are not saturated for multiple hits and can be used in the analysis. Saturated substitution types should be excluded from the MP analysis but not distance and ML analyses.

4.2.5 Hypothesis testing

The Wilcoxon signed-ranks test was applied to examine statistical significance of the most parsimonious tree relative to alternative hypotheses (Templeton, 1983). To find the most parsimonious tree(s) under the alternate hypothesis, a heuristic search was performed under the constraint of the hypothesis to be tested. All the most parsimonious trees generated under the hypothesis constraint were tested (with the Templeton (1983) test) against all the most parsimonious trees generated under no constraint.

4.2.6 Hard vs. Soft Polytomies

Polytomies in a phylogeny may be the result of 1) a rapid radiation in a short space of evolutionary time, called a hard polytomy (Maddison, 1989) or 2) insufficient signal in the data to distinguish genetic relationships or homoplasy due to superimposed nucleotide substitutions at a single site, called a soft polytomy (Maddison, 1989). Jackman *et al.* (Jackman *et al.*, 1999) proposed that hard and soft polytomies could be distinguished by a series of statistical tests. They suggested that if long branches have been divided by many successive branching events, the appropriate subsets of species should reveal phylogenetic structure that is obscured when all species are analysed together. Also, if simultaneous branching occurred, analyses using subsets should lack phylogenetic structure. The statistical test first establishes a critical value for bootstrapping pseudoreplicates by calculating the 95^{th} percentile of a null distribution of bootstrap values. This is done by selecting a subset of taxa representative of branches in a polytomy and creating 100 randomised data sets. A maximum parsimony heuristic search is then performed for each randomised data set and two sets of four taxa are selected on the basis of the tree recovered (see Jackman *et al.* 1999 for selection criteria). For each set of four taxa, a bootstrap analysis is done and the bootstrap values recorded. These 200 bootstrap values provide a null distribution of values from which the 95^{th} percentile is calculated. Next, a bootstrap search is done for every permutation of four taxa from the selected subset of taxa for the real non-randomised data set. All observed

bootstrap proportions greater than or equal to the critical value then indicate that simultaneous branching can be rejected for the taxon set in question. I selected 10 in-group taxa that represented unresolved clades in the weighted maximum parsimony bootstrap analysis and used the shuffle option in MacClade 3.04 (Maddison and Maddison, 1992) to create the 100 randomised data sets.

4.3 Results

4.3.1 Test for paralogous sequences

All the mitochondrial primer pairs amplified PCR products at a dilution of $\geq 10^{-4}$ while the nuclear primers amplified products to a dilution of 10^{-2}. Partial *CR* and *cytb* sequences amplified from enriched mitochondrial DNA and from total cellular DNA for *C. aspera* AMS R115337, *Liasis mackloti* SAMA R21422, *Morelia spilota* SAMA R26878 and *M. viridis* AMS R115348 were indistinguishable, providing no evidence that the primers may be amplifying paralogous sequences.

4.3.2 Partition homogeneity test

With no evidence to suggest the amplification of paralogous sequences, 26 aligned partial *CR* and *cytb* sequences and Kluge's 121 morphological characters were analysed using PAUP*4.0b2a. For the *CR* sequences, there were a total of 768 nucleotides, 230 of which were excluded due to ambiguous alignment, and including three binary characters representing structural features. For the *cytb* sequences, there were 1114 nucleotides. To determine if the morphological and two molecular data partitions could be combined into a single analysis, the Incongruence Length Difference (ILD) Test (Farris *et al.*, 1995) with 1000 replications was performed. The test was significant (p = 0.001). The test was then performed on the molecular data partitions only and the result was also significant. Experimentation with this test, using sequential addition of taxa, on the effect of various

combinations and input order of taxa in the data set, produced peculiar results. For the combination of *CR* and *cytb* partitions, I could include up to 20 taxa and have a non-significant test, but this depended on the addition order of the taxa. There were six taxa that could not be added without making the test result significant.

4.3.3 Single partition analysis

Factors contributing to the heterogeneity among data partitions could include different gene histories (gene trees), the partitions are not representative samples of the character population due to sampling error or there are differing underlying models of nucleotide substitution. Since the two mitochondrial partitions in the present study are samples of a single non-recombining, maternally inherited "supergene", each has an identical gene history. Conflict between partitions could occur due to sampling error such that the gene portions chosen for analysis may contain insufficient phylogenetic signal to recover the real tree or they may contain homoplasy that confounds the search for the most-parsimonious tree.

4.3.3.1 Signal or noise?

Commonly used indicators of the level of phylogenetic signal to noise in sequence data are the proportions of parsimony-informative sites and the g_1 statistic which measures the skewed-ness of tree-length distributions. (Hillis, 1991; Hillis and Huelsenbeck, 1992). The proportion of parsimony informative sites was similar for each partition, being 161/539 for *CR* (~30%) and 332/1114 for *cytb* (~30%). The g_1 statistic for each partition was –1.015 and –0.429 respectively, suggesting that there is adequate signal for parsimony analysis.

4.3.3.2 Test for model of sequence evolution

Although the genes of the mitochondrial genome have identical histories of descent, each gene may be subjected to different selection pressures, leading to substantially different underlying modes of nucleotide substitution, particularly as the sequence partitions include different classes of functional genes in the mitochondrial genome.

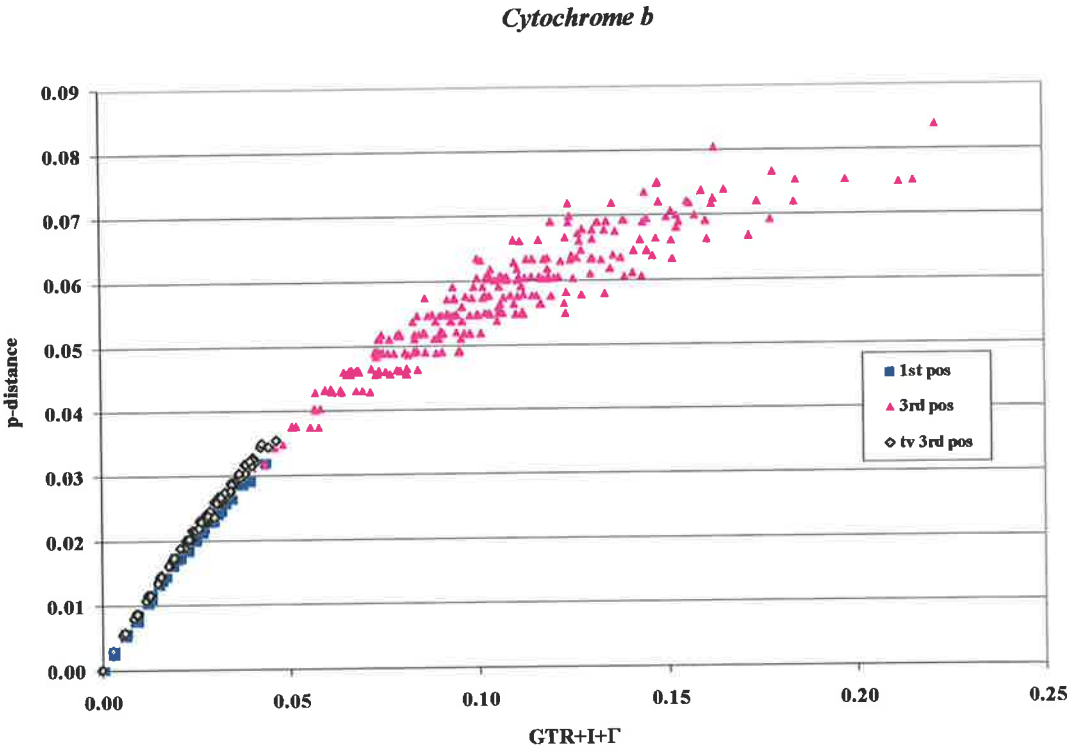
To determine the mode of nucleotide substitution for each data set, I used Modeltest3.0 (Posada and Crandall, 1998) to estimate the molecular evolutionary model best suited to each partition. The General Time Reversible plus proportion invariable sites plus gamma shape parameter (GTR+I+Γ) model (Rodriguez *et al.*, 1990) was found to be the most suitable for each partition. The GTR+I+Γ parameters were estimated for *CR* and *cytb* using ML. The parameters estimated for *CR* and cytb are listed in Table 4.1.

Table 4.1 Maximum likelihood parameters for the GTR+I+Γ evolutionary model as estimated by PAUP*4.0b2a.

		control region	*cytochrome b*	**Combined**
nucleotide frequencies	(A)	0.36088	0.36590	0.32374
	(C)	0.35906	0.39273	0.30097
	(G)	0.07624	0.03875	0.13238
	(T)	0.20383	0.20262	0.24291
substitution rates	A	0.5159	0.2974	1.52284
	A	7.7015	21.3083	7.06710
	A	0.8848	0.58837	1.78174
	C	0.2157	1.2980	0.41394
	C	9.8119	7.2182	12.80278
	G	1	1	1
proportion of invariable sites	(I)	0.49425	0.45073	0.47532
gamma shape parameter	(Γ)	0.97158	1.11574	0.61675

4.3.3.3 Test for nucleotide saturation

The *CR* sequences were examined for saturation of transitions and transversions by plotting pair-wise uncorrected *p*-distances against GTR+I+Γ distances. As third codon positions of mitochondrial coding genes are frequently saturated for transitions, I analysed all three codon positions of *cytb* and all sites for transition and transversion saturation. There was evidence of saturation across all sites for transitions in the *CR* and for transitions at third codon positions in *cytb* (Fig. 4.4). Therefore, transitions in the *CR* and at third codon positions in *cytb* were excluded from the MP analysis by down-weighting transitions to 0 in a



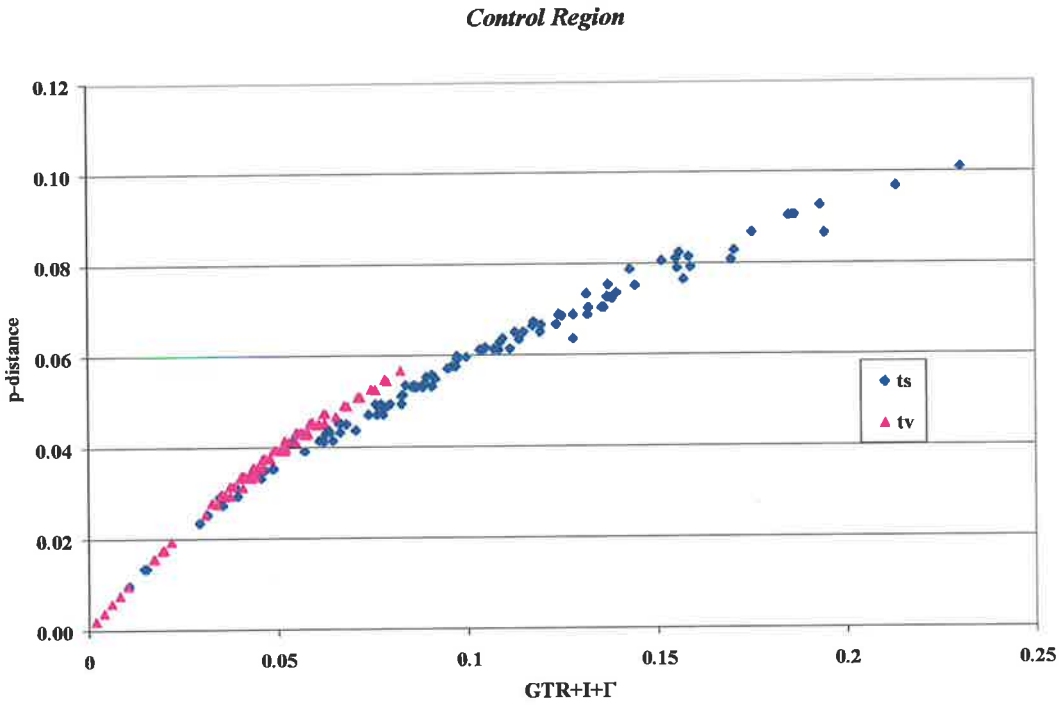


Fig. 4.4 Graphs of uncorrected p-distance against the GTR+I+Γ distance for 1st and 3rd codon positions for transitions and 3rd position transversions for *cytochrome b* (top) and for transitions and transversions for *Control Region*.

step-matrix in PAUP*4.0b2a.

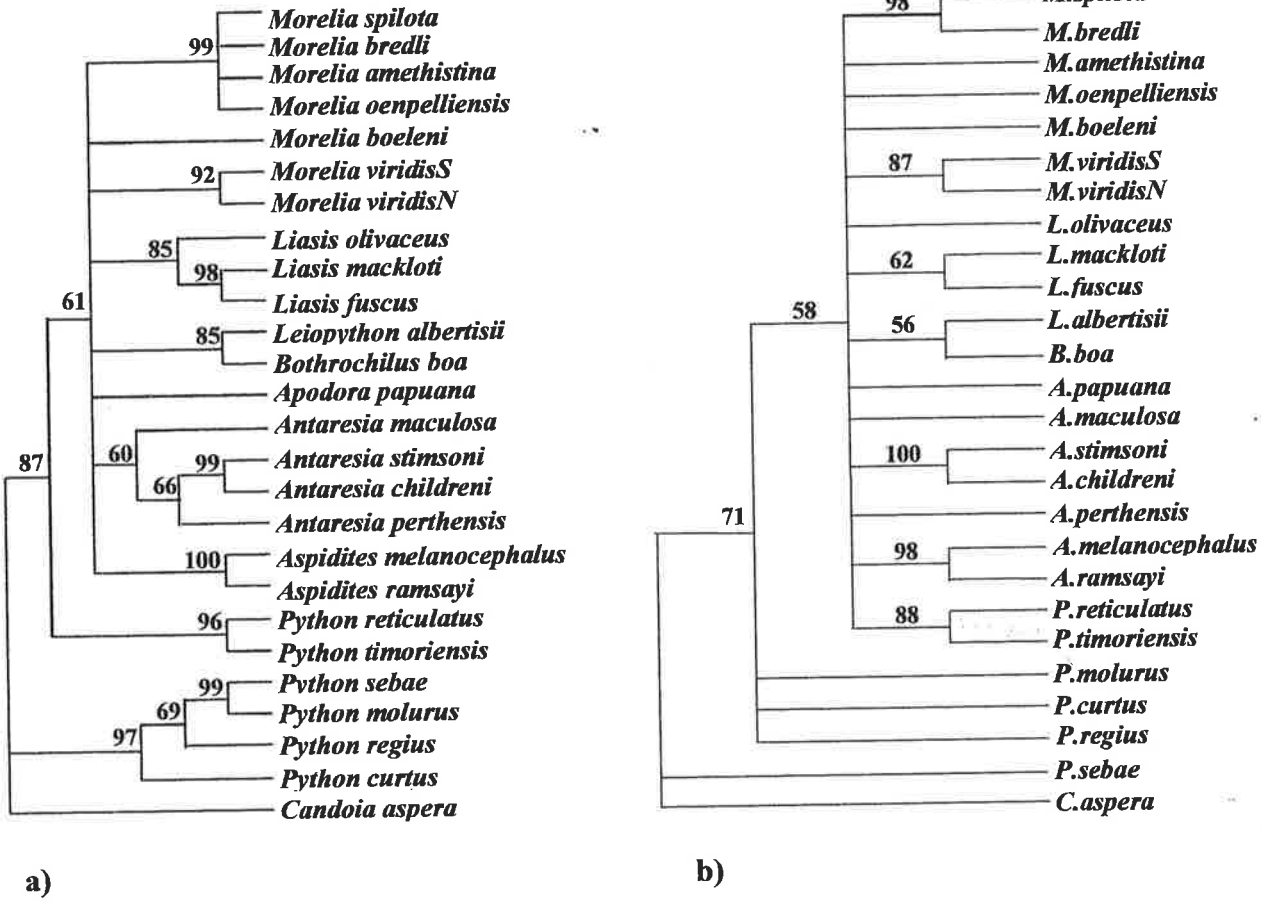
4.3.3.4 Phylogenetic analysis

For *CR* data in a weighted MP heuristic search, there were 21 equally most-parsimonious trees. (A strict consensus tree of 21 equally most-parsimonious trees for *CR* is presented in Fig. 4.5a. Proportions of bootstrap pseudoreplicates are also shown). For *cytb*, the weighted MP analysis, produced three equally most-parsimonious trees, which are shown in Figs. 4.5b and c. Much of the structure present in the three *cytb* trees collapses when bootstrapped. For both partitions, there was very little resolution at deeper levels of the trees.

Analysis of the three structural features present in the 5' region of the *CR* relative to the python phylogeny showed concordance between the presence of the hairpin and the partial isoleucene tRNA paralogous sequence. Both features are present in all the Australo-Papuan pythons although there are only 14 bp of the isoleucine tRNA sequence present in the two species of *Aspidites* and only 10 bp present in *P. timoriensis*. Neither feature is present in *P. curtus*, *P. sebae* or the outgroup, *Candoia aspera* and consequently the *CR* is considerably shorter for these taxa. A parsimony analysis of character state change of the evolution of each of these characters, performed with MacClade version 3.04 (Maddison and Maddison, 1992), showed the insertions of the isoleucine paralogue and hairpin to be derived along the lineage between the Afro-Asian pythons and the western Indonesian *Python reticulatus* and *P. timoriensis* (Fig. 4.6). The partial duplication of the hairpin appeared along the lineage separating A. *childreni* and A. *stimsoni* from the other *Antaresia* species.

4.3.4 Combined data partitions

Finally, on the basis that conclusions of single-partition analyses that: i) a similar mode of evolution could be applied to each partition, ii) there was very little strongly supported conflict in separate partition analyses, iii) there was no conflict in decay index support between partitions at the nodes leading to *Aspidites* and *Python* species (the



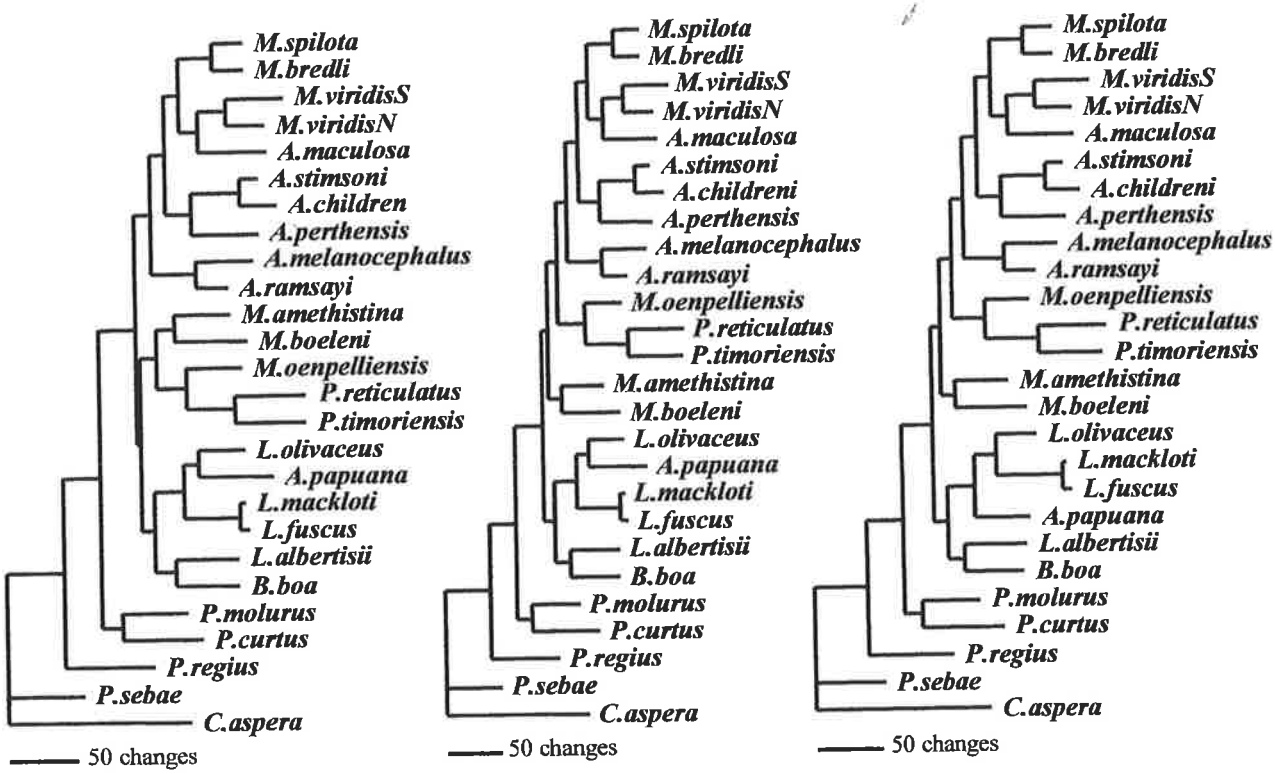
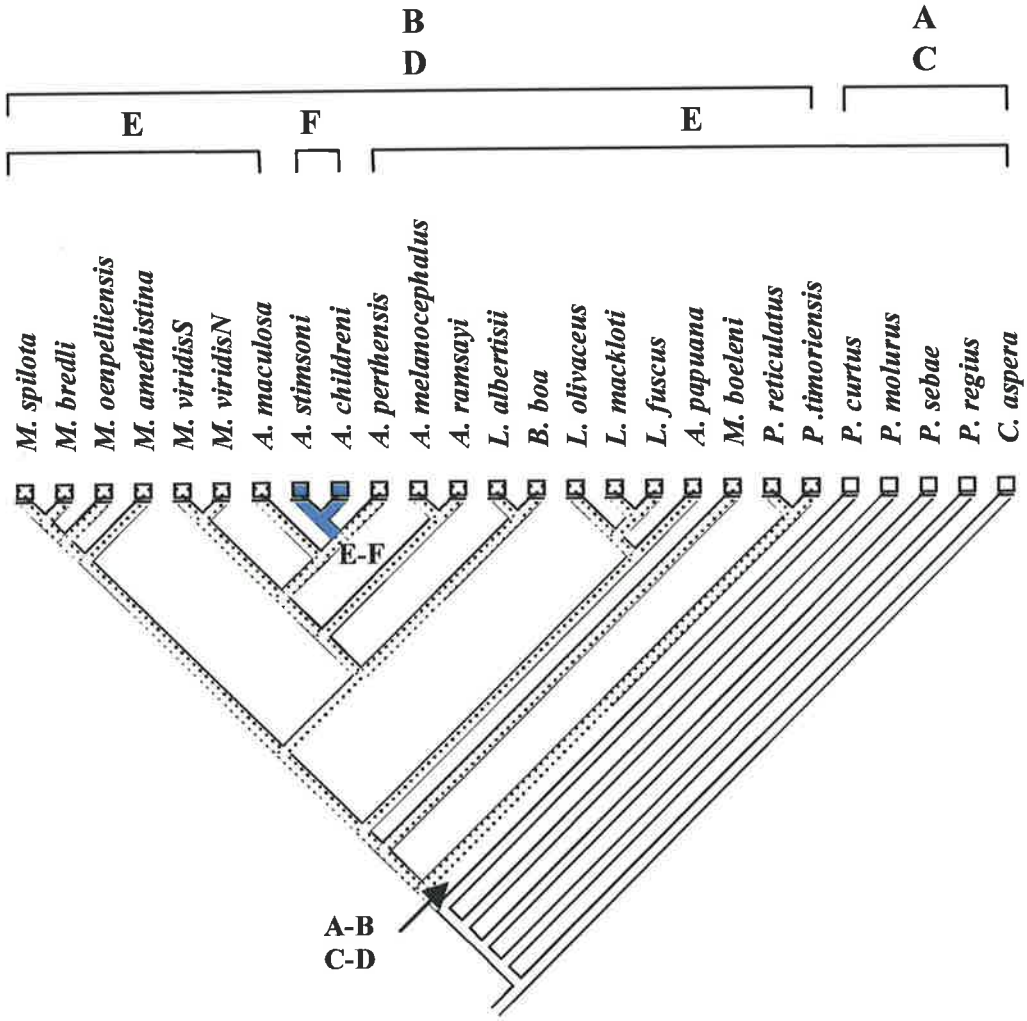


Fig. 4.5 Strict consenus MP trees of a) 21 and b) 3 equally most-parsimonious trees for *CR* and *cytb* partitions respectively. The numbers on each tree represent bootstrap pseudoreplicate proportions. c) Three equally most-parsimonious trees for the *cytb* partition.



$tRNA^{Ile}$ pseudogene & 5' hairpin

$tRNA^{Ile}$ pseudogene, 5' hairpin & ½ hairpin indel

Fig. 4.6 Presence/absence of three control region secondary structure elements, a $tRNA^{Ile}$ pseudogene, a 5' ~30bp hairpin and a 15bp ½-hairpin indel.

hypotheses being tested in this study) and iv) the gene partitions all have one evolutionary history because they are all mitochondrial sequences, I combined the nucleotide data for phylogenetic analysis.

4.3.4.1 Phylogenetic analysis

For the combined nucleotide sequence partitions, MP, ML and NJ methods were used. For the parsimony analysis, *control region* transitions and *cytb* 3rd codon transitions were excluded. The model of nucleotide substitution found for the combined data set, using Modeltest3, (Posada and Crandall, 1998) was GTR+I+Γ. This model was used for the ML and NJ analyses. Parameters for the combined data set estimated for this model using ML are listed in Table 4.1. The ML tree is shown in Fig. 4.7 with bootstrap proportions for MP, NJ and ML analyses indicated.

In all three methods of analysis, similar tree topologies were recovered, with internal branch lengths being very much shorter than terminal branches. There was also a high level of concordance for a significant lack of bootstrap support for inter-generic relationships among pythons, except for the sister relationship between *Bothrochilus boa* and *Leiopython albertisii* (100%, 96%, 100% pseudoreplicates for ML, MP and NJ respectively) (Fig. 4.7). In all analyses, the genus *Morelia* was not monophyletic. *Morelia amethistina*, *M. bredli*, *M. oenpelliensis* and *M. spilota* formed a clade. However, the two *M. viridis* species clustered as a sister clade to the *Antaresia* species complex and *M. boeleni* was a sister taxon to all other Moreliini pythons. None of these relationships are supported by strong bootstrap support in the MP or ML analyses but there was quite strong bootstrap support in the NJ analysis. The hypothesis of *Morelia* monophyly was tested using the Templeton (1983) test (see below). The *Antaresia* species (*childreni*, *maculosa*, *perthensis*, and *stimsoni*) and the *Liasis olivaceus* Group each formed a monophyletic clade, but with low (<50%) bootstrap support for each in the MP and ML analyses (Fig. 4.7). A sister relationship between *A. childreni* and *A. stimsoni*

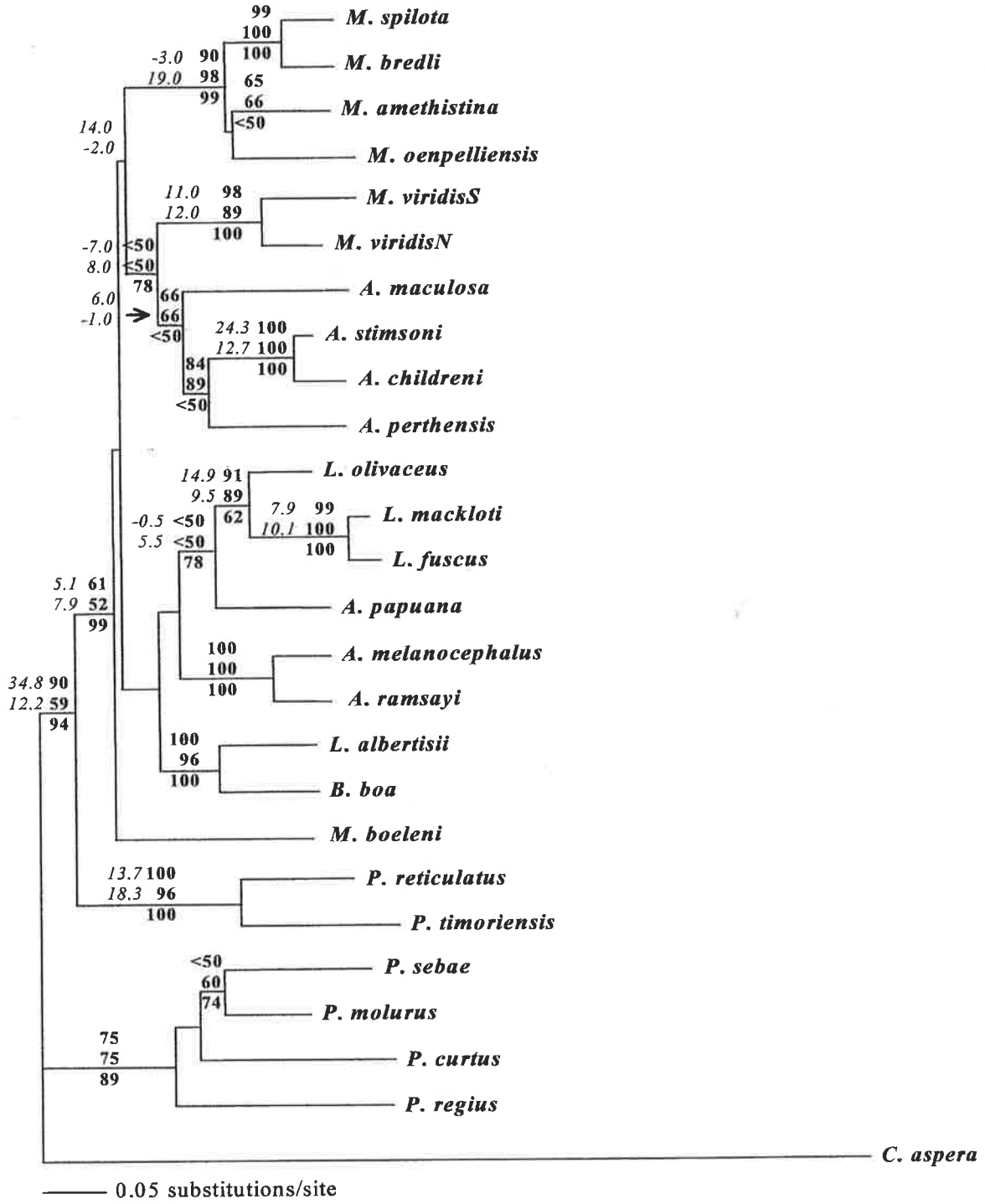


Fig. 4.7 Maximum likelihood phylogram of a weighted heuristic search showing proportion of bootstrap pseudoreplicates for MP, ML and NJ analyses in descending order. Bremer decay indices for the *CR* and *cytb* partitions are shown in unbolded italicized numbers in descending order respectively.

was strongly supported by bootstrapping (100%) in all three analyses. In each analysis, the *Aspidites* clade formed a sister group to the *Liasis olivaceus* Group.

The *Python* species were also paraphyletic, with *P. reticulatus* and *P. timoriensis* forming a sister clade to the Moreliini. Only the ML analysis had the other *Python* species forming a monophyletic group, with bootstrap support of 90% pseudoreplicates. In the weighted MP analysis, there were two equally most-parsimonious trees conflicting only in the placement of *P. curtus*, relative to *P. molurus*. In the MP and NJ analyses, there were 90% and 94% support for the clade comprising the Moreliini and *P. reticulatus*/*P. timoriensis* but there was only 59% bootstrap support in the ML analysis.

4.3.4.2 Bremer decay analysis of data partitions

To determine the support contributed by each data partition to the phylogenetic analysis, Bremer decay indices were determined using TreeRot v.2.0. (Sorenson, 1999) for each partition on a maximum parsimony strict consensus tree generated from the combined data. Decay indices were very low (-1.0 – 2.0) for most nodes leading to *Python* species, with the exception of the node separating *P. reticulatus* and *P. timoriensis* from the other *Python* species (Fig. 4.7). This node had decay indices of 34.8 and 12.2 for *CR* and *cytb* respectively. Decay indices of 5.1 and 7.9 supported the node that separated *P. reticulatus* and *P. timoriensis* from the Moreliini taxa. For the most part, the decay indices for both partitions were in agreement as to the measure of support for each node. Notably, there was disagreement at the node associating *Apodora papuana* with *Liasis*, at a node associating *M. viridis* with *Antaresia* species, and with nodes linking *Morelia* species.

4.3.5 Hard vs. soft polytomies

Ten in-group taxa were selected from a weighted MP tree to represent unresolved clades. The taxa included in this analysis were: *Antaresia maculosa*, *Apodora papuana*, *Aspidites melanocephalus*, *Candoia aspera*, *Leiopython albertisii*, *Liasis mackloti*, *Morelia*

spilota, *M. viridisS*, *Python molurus* and *P. reticulatus*. A weighted MP heuristic search was conducted on each of one hundred randomised pseudoreplicates of the data set. Using the criteria of Jackman *et al.* (1999), two subsets of four taxa were chosen from each of these trees, giving 200 subsets. One hundred bootstrap pseudoreplicates were performed for each of the two hundred subsets and the bootstrap values were used to provide a null distribution of bootstrap values (Fig. 4.8). The 95th percentile calculated for this null distribution set a critical bootstrap value at 89% pseudoreplicates. One hundred bootstraps were performed on the unshuffled data set for each of 204 permutations of 4 taxa and bootstrap values equal to or greater than 89% were indicative that simultaneous branching could be rejected for the taxon set in question. Almost twenty-seven (26.8) percent of the 204 permutations had significant bootstrap values. This was considerably higher than the expected number of significant sub-samples (that is, five per cent of the total number of sub-samples, *i.e.* 10.2 significant bootstrap values out of 204). However, when those significant sub-samples that included either *Python* species were removed, there were only two significant sub-samples. According to Jackman *et al.* (1999), if the observed number of times that a particular branch consistent with the most-parsimonious tree appeared in significant sub-samples was greater than predicted from their proportional representation in all sub-samples, that branch was a probable source of phylogenetic signal. One hundred and fifty-six out of 204 or 76.5% of sub-samples contained *P. molurus* or *P. reticulatus*. Of the significant results, there were 52 out of 54 or 96.3% that contained *P. molurus* or *P. reticulatus*, suggesting that there is strong phylogenetic signal in the branch supporting the *Python* species as sister taxa to other taxa in the selected subset. The two non-*Python*-containing significant subsets were in conflict with the phylogenies produced by all three analyses and therefore, can be rejected as real conflicts in the data. Therefore, the null hypothesis of a hard polytomy cannot be rejected for all the non-*Python* containing branches within the phylogenies recovered.

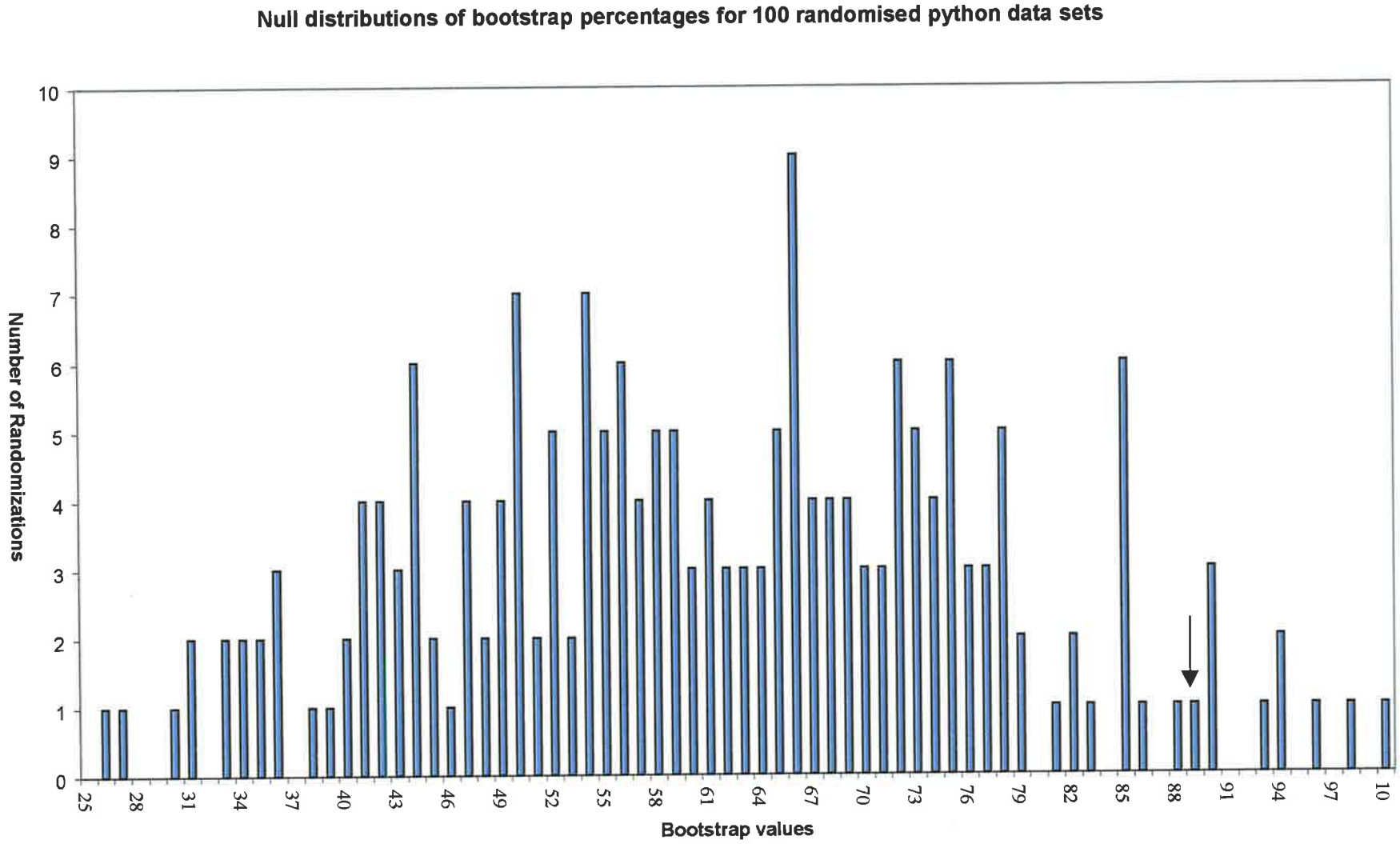


Fig. 4.8 Null distribution of bootstrap values for 100 randomised four-taxa python data sets. The x-axis represents the proportion of bootstrap pseudoreplicates supporting the relationships within the lineage of interest. The arrow marks the 95th percentile of 89%.

4.3.6 Hypothesis testing

Templeton's (1983) test and the Kishino-Hasegawa (Kishino and Hasegawa, 1989) test were applied to the data set to examine the statistical significance of the shortest MP trees relative to the hypothesis of Kluge (1993) that *Aspidites* rather than *Python* is the sister lineage to all pythons, *i.e.* (*Candoia*, (*Aspidites*, (*Antaresia*, *Apodora*, *Bothrochilus*, *Leiopython*, *Liasis*, *Morelia, Python*). In a weighted heuristic search under this constraint, there was only a single most-parsimonious tree generated. This single most-parsimonious tree was compared to each of the two equally most-parsimonious trees found with no constraint. The tree constrained to Kluge's hypothesis was not significantly different from the overall shortest tree in either test, with *P*-values between 0.1186 and 0.1236 for both tests. Therefore, the hypothesis of *Aspidites* as a sister group to all other pythons could not be rejected according to the Templeton or Kishino-Hasegawa tests.

I also tested the monophyly of the genus *Python* as suggested by the phylogeny of Underwood & Stimson (1990) *i.e.* (*Candoia*, *Aspidites*, *Antaresia*, *Apodora*, *Bothrochilus*, *Leiopython*, *Liasis*, *Morelia* (*Python*) and the monophyly of the genus *Morelia, i.e.* (*Candoia*, *Aspidites*, *Antaresia*, *Apodora*, *Bothrochilus*, *Leiopython*, *Liasis*, *Python, (Morelia*). There was only a single most-parsimonious tree generated under each constraint. *P*-values for the Templeton (1983) and Kisihino-Hasegawa (1989) tests were between 0.1144 and 0.1223 for the monophyly of *Python* and between 0.4041 and .06937 for the monophyly of *Morelia*. Therefore, neither hypothesis could be rejected according to the Templeton or Kishino-Hasegawa tests.

4.3.7 Morphological Data

4.3.7.1 Skull examinations.

With the assistance of Dr. M. Hutchinson, I examined 11 of the 16 characters defined by Kluge (1993) in the light of the four questions outlined above. Kluge examined many

specimens of each species. Intra-species variation of quantitative characters was represented by the modal or median condition. Frequently observed states in qualitative traits were included in Kluge's (1993) data matrix, but multiple states within a species were treated as unknown in his analysis. I could not examine character 48 (fenestra of the septomaxilla) as the *Aspidites* skull was damaged in this region, characters 91 and 95 which were scale characters and character 111 which was an arterial character.

Characters answering no to the question "Can the qualitative character states in a species observed by Kluge (1993) be verified in other specimens?" were excluded from the ongoing analysis. Character 46 was described as the ventral, horizontal, fluted wing of the vomer, posterior to the septomaxilla in two character states – as i) uniformly wide or variant in width or ii) uniformly narrow in width. In the skulls examined, I found that the ancestral state of character 46 (i) could not be singularly assigned to *Aspidites*, but that *Antaresia childreni* and *Morelia spilota* also had uniformly narrow wings of the vomer, similar to *Aspidites*. Thus, my findings differed from those of Kluge and I excluded character 46 from the next stage of analysis.

Characters answering no to the question "Are the character states homologous and anatomically identical?" were excluded from the ongoing analysis. In discussing character 5 (described as the ventral openings for the premaxillary channels as located posterior (0) or anterior (1) to the posterior margin of the premaxillary teeth), Kluge (1993) mentions character states that occur in single specimens, that he excludes from the analysis. For example, one specimen of *Bothrochilus boa* which had premaxillary teeth and no channels was excluded from the analysis, as the remaining *B. boa* had state 1. Also, one *Liasis olivaceus* specimen had state 0, but as all the remaining *L. olivaceus* had state 0, posterior was the character state used in the analysis. Similarly, the character could not be scored in the outgroups which lack adult teeth, but as state 0 was determined for *Anilius* and *Xenopeltis*, posterior was assumed to be the plesiomorphic state. For *Aspidites,* it was difficult to determine the character state of this character as premaxillary teeth are absent in this genus.

Kluge assumes the state to be 0 as the channels are located in a posterior position on the premaxilla. However, the channels may appear to be in a more posterior position because of the anterior growth (hypertrophy) of the rostrum of the premaxilla, a condition that is associated with burrowing specialisation. It would be more conservative to score this character state for *Aspidites* as unknown. Thus, I excluded character 5 from being a strong character to support the sister relationship of *Aspidites* to other pythons.

Character 41 was described as the lateral crests on the supraoccipital having two character states – i) being absent or weakly developed or ii) being well-developed and prominent. Absence of the lateral crests on the supraoccipital was found in *Aspidites* and *P. sebae* but the finding was equivocal in that the character state found in *Aspidites* and *P. sebae* was homologous to that of *Boa*.

Eight characters out of eleven could be examined for the above questions and answered in the affirmative. These characters were then examined under two further questions. If the questions "Is the character continuous and therefore have the character state definitions been objectively determined?" could be answered yes and no respectively, the character was excluded from ongoing analysis. The following characters answered no under this examination and were excluded from ongoing analysis: Character 66 was described as the height of the dentary teeth, with two character states, i) a slight change or ii) a marked change between the heights of teeth 4-6 and teeth 9-11. The change in teeth height is not very well determined for the taxa I examined and therefore I consider this character to be somewhat ambiguous.

Finally, I asked the question, "Are the characters correlated in any way?" If examination of the character answered yes, I felt that the character may not contribute as a unique, unlinked character providing information to the total analysis and therefore it was excluded. *Aspidites* differs from other pythons in that it is a burrowing species. Therefore, there may be correlated characters that are significant to the adaptation to burrowing. I considered that characters (particularly characters of the head) in *Aspidites* that showed a

possible linked pattern of distribution or that were similar to burrowing adaptations in other burrowing taxa may be correlated characters. The following characters were excluded on the basis of question 4. Character 1 describes the presence or absence of palatine teeth. Premaxillary teeth are present in neonate *Aspidites*, but are lost in adults as the rostral shield develops to form the spatulate shape of the snout that is suited to burrowing. Kluge (1993) leaves the debate on the significance of this character open until developmental series of samples can be examined. Character 73 was described as the width of the neck relative to the head and has three character states, with the neck i) as wide as, ii) slightly or iii) markedly narrower than the head. This is a continuous character that is influenced by ontogenetic variation (Kluge 1993) but Kluge compensated for this latter point by only using adult specimens. Neck width could be one of a group of head-associated characters that may evolve together to contribute to the overall reduction in head size found in some fossorial species burrow (Greer, 1985; Rieppel, 1984). Therefore, character 73 is ambiguous relative to question 4 and was excluded from the analysis. Character 86 describes the number of posterior sutures in the nasal scale and has three character states, i) none, ii) one and iii) two or more. To examine this character I used the taxon descriptions of Barker and Barker (1994). This character of the number of posterior sutures in the nasal scale is polymorphic in *Antaresia perthensis* (Barker and Barker, 1994), *M. spilota* and *P. anchietae* (Kluge 1993, Barker & Barker 1994) for states i) and ii), whilst *Aspidites* and *M. viridis* only have state i). As other burrowing reptiles also exhibit the fusion and enlargement of nasal scales (e.g. Ramphotyhplops, Ophioscincus, Cogger, 1996), state i) may be an adaptation to burrowing in *Aspidites*. Character 106 was described as the number of thermoreceptor pits in seven states. *Aspidites* does not have thermoreceptor pits, nor do the outgroups bolyerines, erycines, *Loxocemus* and *Xenopeltis*. *Boa manditra, Corallus* and some *Epicrates* (Kluge 1993) have some form of pits and there is some debate as to whether *Boa constrictor, B. dumerili, B. madagascariensis* and *Eunectes* have pits (Kluge 1993). *Aspidites* may have secondarily lost thermoreceptor pits as an adaptation to burrowing. Thus, I excluded character 106 from

ongoing analysis.

Therefore, on the basis of my criteria, for 11 out of 16 characters used by Kluge (1993) to separate *Aspidites* as a sister lineage to all other pythons, eight characters were rejected as being equivocally useful phylogenetically diagnostic characters. Four characters that were not excluded from the analysis on the basis of my criteria were character 31 (based on the separation of the supraorbital from the parietal), character 45 (based on the relative length of the quadrate), character 50 (based on the number of palatine teeth per ramus in an adult) and character 55 (based on the width of the maxillary process).

4.3.7.2 Phylogenetic analysis of morphological characters

Using MP, I re-analysed Kluge's (1993) data with an expanded set of five outgroups. A heuristic search of 121 characters (119 parsimony informative, two parsimony uninformative) with simple stepwise addition with tree bisection-reconnection found 60 equally most-parsimonious trees of length 464 steps. A strict consensus tree is shown in Fig. 4.9a. This compared with two equally most-parsimonious trees of length 410 steps found using a generalised 'boine' outgroup (consensus tree shown in Fig. 4.9b). Each search was also repeated with the eight characters mentioned above excluded from the analysis. For the analysis with five outgroups, there were 20 equally most-parsimonious trees of length 422 steps (Fig. 4.9c), compared to seven equally most-parsimonious trees of length 375 steps found with a single outgroup (Fig. 4.9d). In each of the analyses using all characters, *Aspidites* was placed as the sister taxon to all other pythons with weak bootstrap support (58%, decay index of 1) (Fig. 4.9). All of the polytypic genera were each monophyletic, but there was only strong bootstrap support for the *Antaresia* and *Aspidites* clades (93% [d=5] and 97% [d=4] respectively with expanded outgroups and 97% and 100% respectively with a generalised outgroup). *Python* and *Morelia* formed sister clades with bootstrap pseudoreplicate proportions of 66% and 54% with expanded outgroups and generalised outgroup respectively. Decay indices were all five or less with the strongest support for the

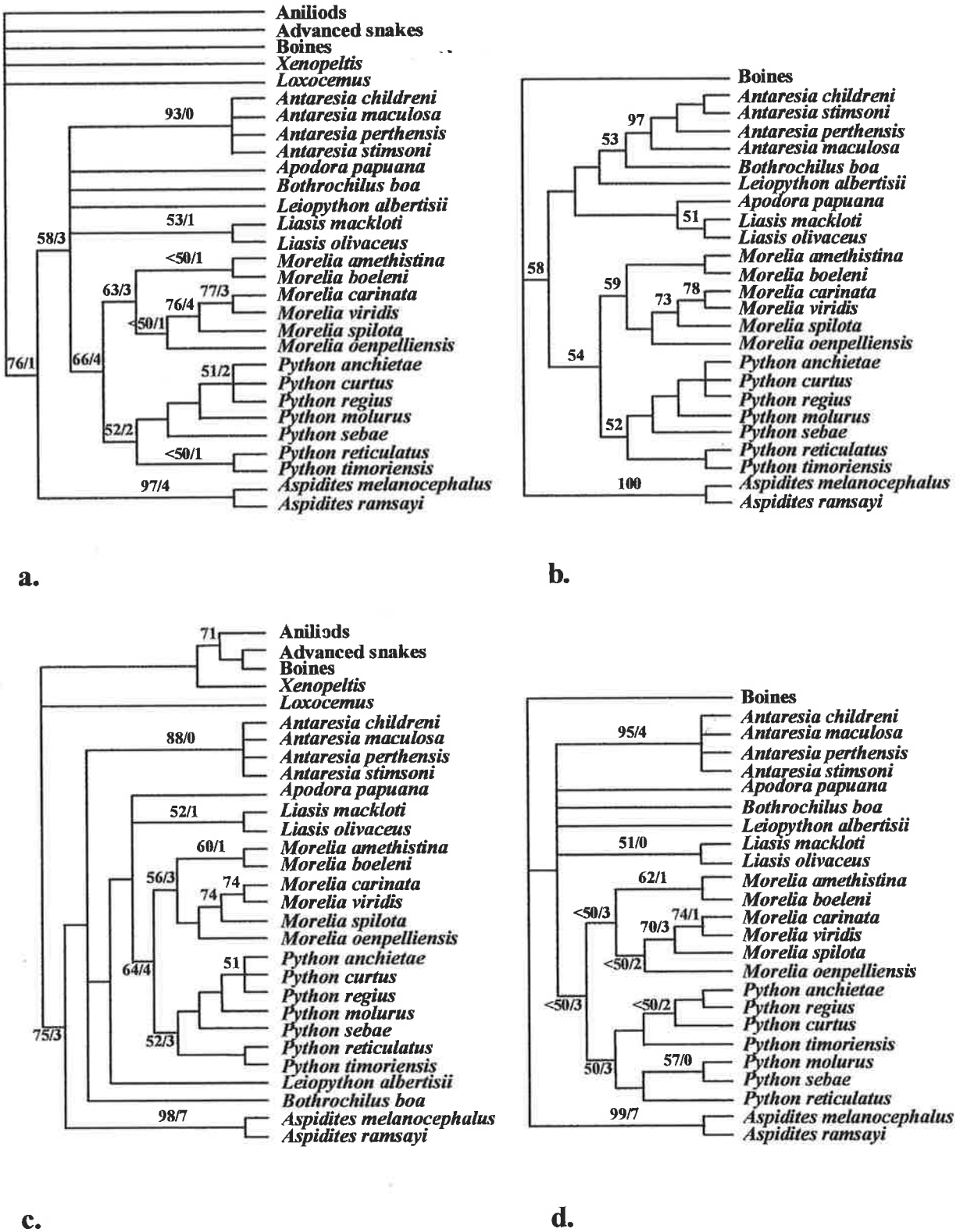


Fig. 4.9 Strict consensus trees of maximum parsimony analysis of Kluge's (1993) morphological data. Figs a and c include an expanded set of outgroups, b and d incorporate a single common ancestor as per Kluge's analysis. Figs c and d exclude eight potentially ambiguous morphological characters that Kluge considers significant to *Aspidites* characterization (see text). Numbers represent bootstrap pseudoreplicate proportions and Bremer decay indices respectively.

Antaresia clade.

The alternative hypothesis that *Python* may be the sister group to all other pythons was tested on the morphological data using the Templeton (1983) and Kishino-Hasegawa (1989) tests. When a constraint forcing the *Python* clade as the sister group to all other pythons (*e.g.* (*Python*, (*Antaresia*, *Apodora*, *Aspidites*, *Bothrochilus*, *Leiopython*, *Liasis*, *Morelia*))) was imposed on Kluge's (1993) data set, eight equally most-parsimonious trees were found. There were no trees generated under the constraint that were significantly different from the either of the two equally most-parsimonious trees generated without constraint for either test, *p*-values ranged from 0.1083-0.2931. This suggests that the hypothesis of *Python* as the sister to all other pythons cannot be rejected from Kluge's data.

4.4 Discussion

4.4.1 Morphology

The classification of pythons has been an area of frequent change for many years, even at the level of generic assignment, as morphological characters have been assessed and reassessed as to homology and polarity (McDowell, 1975, Kluge 1993). It has been suggested in the past that there is little divergence within the python group, with McDowell (1975) even suggesting that there was good argument for referring all python taxa to the genus *Python*.

The most controversial difference between previous python systematic arrangements is the placement of *Aspidites* relative to other python taxa. Kluge's (1993) argument against McDowell's (1975) and Underwood & Stimson's (1990) placement of *Aspidites* is based on the latters' *a priori* assumption that the presence or absence in *Aspidites* of some characteristics (*e.g.* thermoreceptive labial pits and the horizontal part of the nasal bone lying above the nostrils) is due to secondary loss or secondary acquisition of these traits respectively. Premaxillary teeth are present in juvenile *Aspidites* and are lost in adults, which suggests that the absence of premaxillary teeth in adults is more likely to be a secondary loss. However, Kluge (1993) will not consider a possibility of evolutionary reversal of character

states until an ontogenetic series can be examined.

Many of the characters that Kluge (1993) considered to be controversial because of Underwood and Stimson's (1990) *a priori* polarisation of character state changes were associated with fossoriality. While this may have been a procedural error on the part of Underwood and Stimson, it is nevertheless true that *Aspidites* is unique among pythons in its burrowing habits, and therefore characters which contribute to this mode of life are likely to co-evolve. Of Kluge's 16 synapomorphies which unite all pythons exclusive of *Aspidites*, at least seven (the functional significance of most of the remainder is unknown) can be argued as being typical of the anatomical changes which occur in the skulls and scalation of limb-reduced squamates which burrow (Greer, 1979; Greer, 1985; Rieppel, 1984). These characters are linked to the formation of a burrowing rostrum (characters 1 and 5), shortening and strengthening of the facial bones (characters 50 and 55), reduction of the cross-sectional area of the head (character 73), fusion of head shields (character 85) and in the specific case of pythons, absence (a hypothesis of secondary loss can be invoked) of the forward-opening thermoreceptive pits (character 106). A majority of the outgroup taxa used in the parsimony analysis (erycines, *Loxocemus*, *Xenopeltis*, anilioids) are also fossorial to variable degrees, and there is a possibility that *Aspidites* is 'attracted' to a basal position on the tree due to characters convergently shared by snakes with fossorial adaptations, rather than because it might retain symplesiomorphies.

In summary, phylogenetic analyses of Kluge's morphological data with and without the eight characters that may not be reliable autapomorphies of *Aspidites*, and with and without expanded outgroups, are consistent with the placement of *Aspidites* as the sister group to all other pythons. However, bootstrap support for this arrangement is only strong (76%) for the original Kluge (1993) data set. With expanded outgroups, the bootstrap support for *Aspidites* as a sister group drops to 58% and hypothesis testing shows that the data do not unequivocally support Kluge's hypothesis. The sister relationship of *Morelia* to *Python* is also supported in each analysis with low (50%-64% pseudoreplicates) bootstrap support.

4.4.2 Phylogenetic implications

The lack of divergence found in morphological analyses is also reflected at a molecular level. Genes such as *12S rRNA* and *16S rRNA*, often used to determine deeper evolutionary histories (e.g. Hedges and Poling, 1999) do not have sufficient phylogenetic signal to be useful in python systematics (Rawlings unpubl. data). On the basis of similar evolutionary histories and modes of evolution and limited topological conflict, I combined the two mitochondrial genes, *CR* and *cytb* into a single analysis.

The molecular data suggest that the primary split among pythons is between the genus *Python* and the other genera, with *Aspidites* included within the Australian radiation. This is in concordance with the findings of Underwood & Stimson (1990) and Schwaner and Dessauer (1981) who placed the genus *Python* as the sister lineage to all the Moreliini and is fundamentally different from the set of relationships proposed by Kluge (1993) which places the genus *Python* as a nested clade among the majority of the Australo-Papuan genera. The second finding is that the genus *Python* is paraphyletic with *P. reticulatus* and *P. timoriensis* forming the sister clade to the Moreliini. McDowell (1975) made the observation that *P. reticulatus* appeared to be more closely related to the Australo-Papuan genus *Liasis* (now *Morelia amethistina*) than to the other African and Asiatic python species. This relationship is also supported by the presence of the *CR* secondary structures, a $tRNA^{Ile}$ paralogue and a partial 5' hairpin that *Python reticulatus* and *P. timoriensis* have in common with the Moreliini.

However, surprisingly, analysis of the sequence data places *Aspidites* as nested within the Moreliini clade as the sister group to the *M. viridis/Antaresia* clade. There is no previous hypothesis that has placed *Aspidites* as a nested clade. Whilst the hypotheses of Kluge (1993), McDowell (1975) and Underwood & Stimson (1990) differ in the placement of *Aspidites*, each has *Aspidites* as the sister lineage to other major clades. Kluge places *Aspidites* as the sister lineage to all pythons, McDowell (1975) places *Aspidites* as the sister

lineage to a *Bothrochilus/Leiopython*, *Antaresia* and *Liasis* clade and Underwood & Stimson (1990) place *Aspidites* as the sister lineage to all the Moreliini. There is not strong bootstrap support for the relationship of *Aspidites* recovered in the present study.

The diversification of distinct genera is genetically supported, with evidence for the monophyly of *Antaresia* and the *Liasis olivaceus* Group. The present molecular analyses agree with the general finding of morphological studies that there seems to be very limited divergence between the eight genera of the pythons. The molecular analyses show short branch lengths at deeper divergences with much longer terminal branches between species within genera.

Although there is not strong statistical support for tree topologies from each analysis, the nucleotide data support the monophyly of most genera of pythons. Exceptions are *Morelia* and *Python*. Amongst the *Morelia* species, *M. boeleni* is the most divergent. This is contrary to the phylogenetic hypotheses of McDowell (1975), Underwood & Stimson (1990) and Kluge (1993) which all conclude that *Morelia* is monophyletic (Fig. 4.3). In the nucleotide analyses, *M. viridis* is placed in a surprising arrangement as a sister group to *Antaresia*. Previous taxonomists, unsure of any similarity of green pythons to other species of pythons, classified the green python in a monotypic genus *Chondropython*. Perhaps there is more support for classifying green pythons outside of the *Morelia* group than earlier thought. However, there is no *a priori* evidence to suggest this relationship of *M. viridis* to *Antaresia* and whilst there is not strong statistical support in the molecular data for *Morelia* monophyly, the hypothesis of *Morelia* monophyly could not be rejected by the Templeton (1983) test.

The results of the phylogenetic analyses are concordant with parts of each of the three morphological studies detailed above (Fig. 4.3). There was no resolution of deeper relationships except for the distinction of the Afro-Asian pythons from the Australo-Papuan pythons. This supports Underwood & Stimson's (1990) division of Pythonini and Moreliini (Fig. 4.3b) and the immunodiffusion results of Schwaner & Dessauer (1981). Also, the molecular analyses place *Python* as the sister group of the python lineage as per Underwood

& Stimson (1990) (Fig. 4.3b) and not as a derived lineage as found by Kluge (1993) (Fig. 4.3c). The splitting of *Python* in all analyses, with *P. reticulatus* and *P. timoriensis* separated from the remaining *Python* species, reflects the finding of McDowell (1975) (Fig.4.3a), although McDowell does not place *Python* as the sister clade to all other pythons and the present study does not place *P. reticulatus* and *P. timoriensis* as a sister group to *Morelia*.

The present study also concurs with McDowell (1975) in placing *Bothrochilus boa* and *Leiopython albertisii* as sister taxa. McDowell (1975) felt there were so many similarities between *B. boa* and *L. albertisii* that separate generic status was unwarranted. There is considerable pattern variation in *B. boa* with distinctive orange and black striped markings being the most striking and the most common in collections (Kluge, 1993; O'Shea, 1996). However, there is also a uniformly dark form similar to *L. albertisii* (Kluge, 1993; McDowell, 1975; O'Shea, 1996) and a range of variations in between (Kluge, 1993). The geographic range of each species does not overlap, with *B. boa* being found exclusively in the Bismarck Archipelago, and *L. albertisii* not being found in the Bismarck Archipelago, except on Massau Island, the island of the archipelago farthest from New Guinea. McDowell (1975) suggested that the Massau Island population of *L. albertisii* was a remnant population from a previously continuous distribution along the archipelago and that *B. boa* replaced *L. albertisii* across the rest of the archipelago.

The present study also concurs with McDowell (1975) on the monophyly of the *Liasis olivaceus* Group, with all analyses supporting (but with bootstraps of <50% pseudoreplicates for MP and ML) *Apodora* as the sister taxon to the three *Liasis* species. Kluge (1993) tentatively raised *papuana* to the monotypic genus *Apodora* designated as *sedis mutabilis* (of changeable position) due to the lack of resolution of the relationships between *Apodora, Liasis mackloti* and *L. olivaceus*.

The result of the hard *vs.* soft polytomy test that the hypothesis of a hard polytomy could not be rejected suggests that the lack of well-supported resolution in the data set is likely to be the result of a rapid radiation. It has been suggested that in a molecular

phylogeny, when internal branches are much shorter than terminal branches, rapid diversification in earlier evolutionary history is implied and that low sequence divergence between species suggests rapid diversification in the more recent past (Jackman *et al.* 1999). This is indicative of either long periods of species stability amongst just a few species, with a rapid, much more recent, radiation and dispersal, or python speciation occurred a long time ago and the rate of molecular evolution is very slow, maintaining low levels of divergence after speciation. However, this evaluation will also be dependent on the rate of sequence evolution. A rate of nucleotide substitution for *cytb* for pythons was calculated to be was 0.21% per MY respectively for all substitutions and 0.06% per MY for transversions only (see Chapter 7). These estimates were found to be consistent with other data supporting a slower rate of molecular evolution in heterotherms than in the homoeothermic placental mammals (Kocher *et al.*, 1989; Martin *et al.*, 1992; Rand, 1993; Rand, 1994). However, the rate of nucleotide substitution for pythons was found to be even slower than colubrid snakes, suggesting that the lack of phylogenetic resolution from this and other studies of pythons may not be exaggerated.

Phylogenetic analysis of pythons in the present study has shown both short internal branches relative to the terminal branches and low sequence divergence. Pythons similar to the genus *Python* have been around in the fossil record since Eocene times (Hoffstetter, 1964; Rage, 1984; Szyndlar and Böhme, 1993) and fossils similar to the Moreliini have been found only from Miocene deposits in Australia (Kluge, 1993; Scanlon and Lee, 2000; Scanlon *et al.*, 1999). This, coupled with the sequence data, could suggest that rapid diversification did occur during the late Miocene, possibly associated with python radiation into Australasia, but a slow rate of sequence evolution has maintained low levels of divergence in the time since.

The patterns of nucleotide substitutions in the mitochondrial genes examined in the present study would support the hypothesis that ancestral python species underwent little evolutionary change until recent times when they underwent a rapid radiation. The lack of python fossils predating the middle Miocene (Rage, 1987; Scanlon, 1996) could also be

explained by the recent-radiation hypothesis.

Therefore, the present study has not resolved the puzzle of python origins. Future genetic analysis needs to turn to nuclear introns in the search for nucleotide sequences that might be evolving at rate intermediate to the mitochondrial ribosomal RNA and coding genes and the non-coding, seemingly unconstrained sequences of the *CR*.

4.4.3 Secondary Structures

Whilst secondary structures of the *control region* generally have a role in the regulation of the region's function (Clayton, 1992; Doda *et al.*, 1981) and are, therefore, subject to some selective pressures, they are much less susceptible to selective pressures of environment that might determine morphological traits. Thus, patterns of insertions and deletions within the *control region* may reflect association by descent rather than be due to convergence. The pattern of evolution of the 5'secondary structures of the *CR* could be explained as follows. Pythons are the only snake lineages examined to date that have an extended 5' section of the *CR* (Kumazawa *et al.*, 1996; Kumazawa *et al.*, 1998 see Chapter 3). The isoleucine tRNA paralogue and 15 bp hairpin are not found in the boid outgroup *Candoia aspera* nor are they present in colubrid or viperid snakes (Kumazawa *et al.*, 1996; Kumazawa *et al.*, 1998; see Chapter 3). This presence/absence pattern of these structures is also found in the duplicated copy of the *CR* that is present elsewhere in the mitochondrial genome of snakes (Kumazawa *et al.*, 1996; Kumazawa *et al.*, 1998; see Chapter 3). Therefore, absence of the isoeleucine paralogue and hairpin is the plesiomorphic state, which is also found in four of the *Python* species (*P. molurus*, *P. sebae*, *P. curtus* and *P. regius*). *Python reticulatus* and *P. timoriensis* have the derived state in common with the Moreliini pythons, suggesting that the two eastern-most Indonesian pythons and the Moreliini probably shared a common ancestor in which the insertion of the isoleucine paralogue and hairpin occurred.

4.4.4 Biogeographic implications

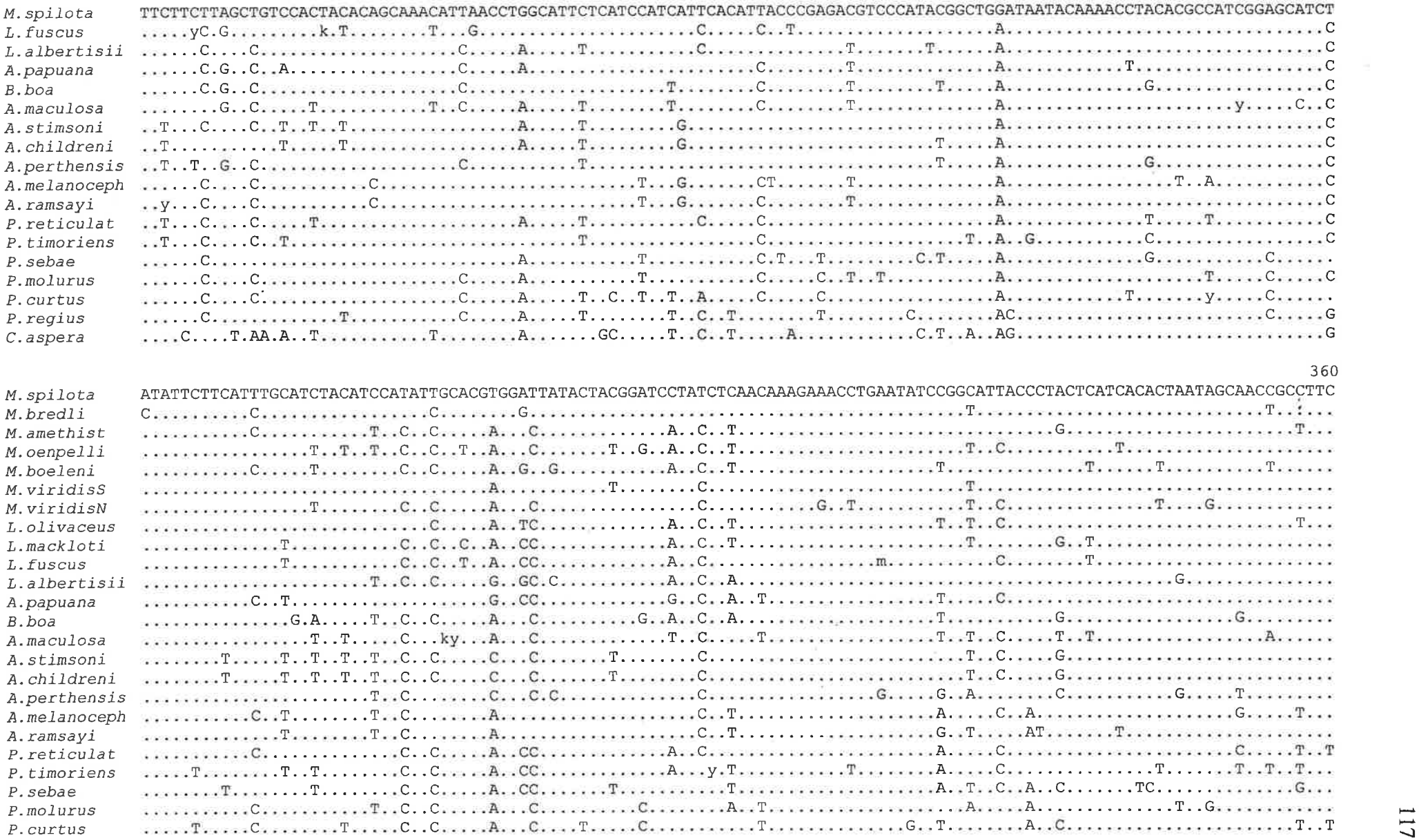
On the basis of the most ancestral taxon in each study, Underwood & Stimson (1990) and Kluge (1993) formulated biogeographic hypotheses for the pythons. Underwood & Stimson (1990) suggested that the Pythonini originated in south-east Asia and the Moreliini evolved from a south-east Asian stock and radiated into Australia and New Guinea. In contrast, Kluge (1993) concluded an Australia-New Guinea origin of pythons with subsequent radiation into south-east Asia and Africa.

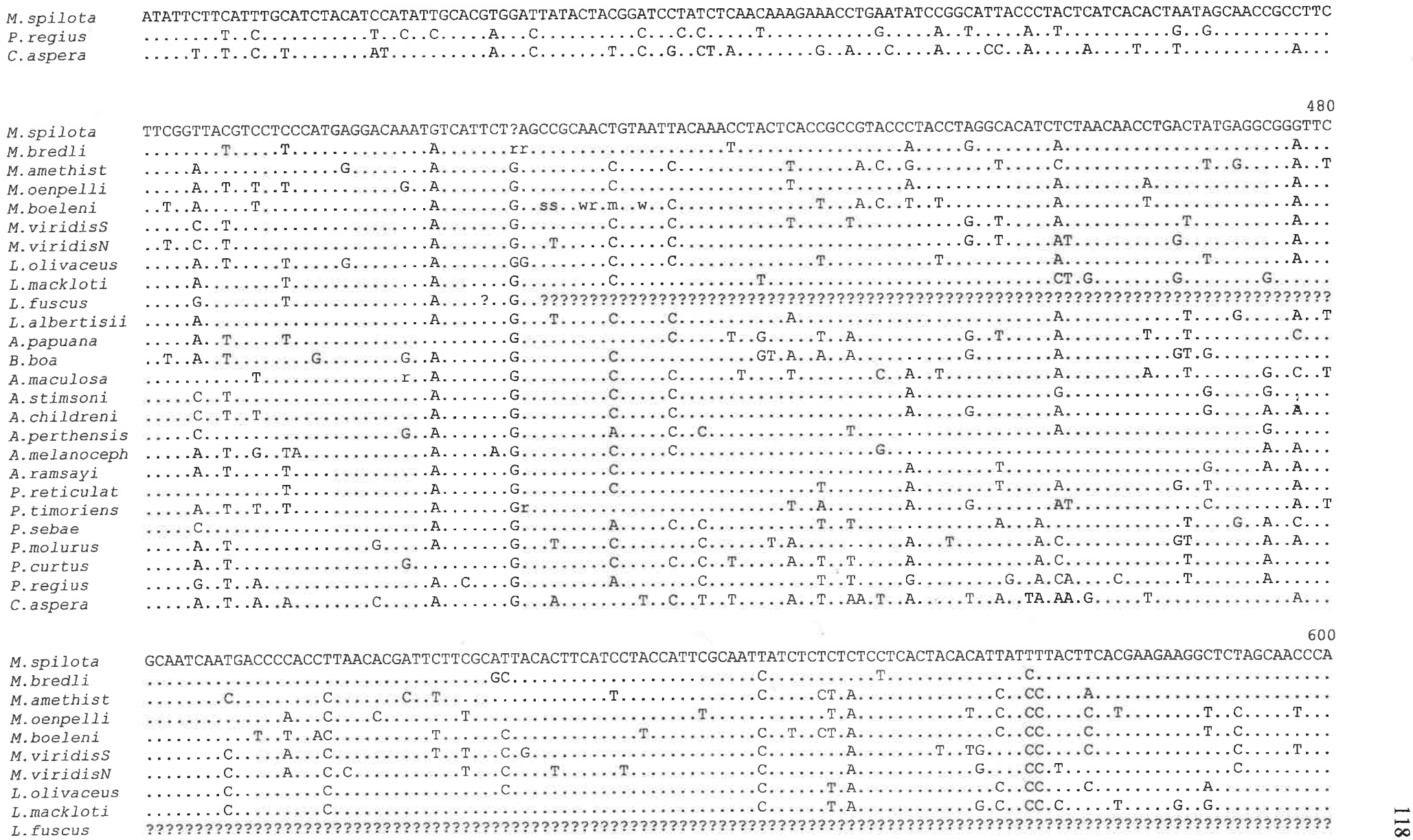
The phylogenetic results of the present study place the Afro-Asian pythons as the sister group to all other pythons. This suggests that pythons arose out of Africa or Asia and dispersed eastwards through Malaysia and Indonesia to Australia and New Guinea. The paraphyletic split within *Python* occurs amongst the Asian pythons with *P. reticulatus* and *P. timoriensis* being divergent from the Asian *P. curtus* and *P. molurus* and the African pythons. *Python curtus* has a distribution stretching from Malaysia, south-eastwards to Sumatra and Borneo. *Python molurus* is found in Pakistan, India and southern China and is sympatric with *P. reticulatus* through Thailand, Sumatra and into Java. *Python reticulatus* extends its range eastwards into Nusa Tenggara and the Moluccas. It is sympatric with *P. timoriensis* on the islands of Flores and Timor. Lack of resolution of relationships amongst the *Python* species precludes any speculation as where to the divergence of *P. reticulatus* may have occurred.

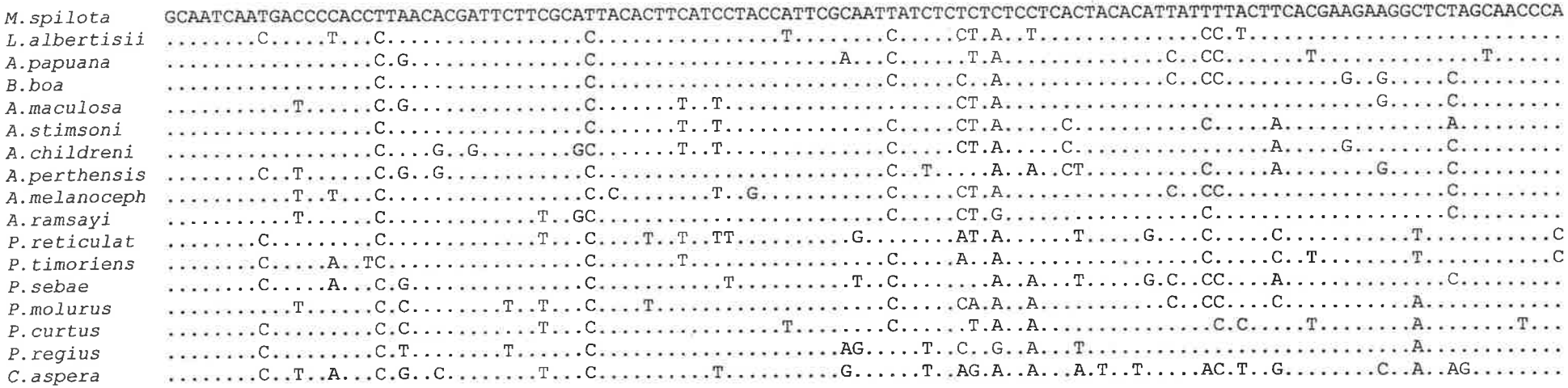
Appendix 4.1.

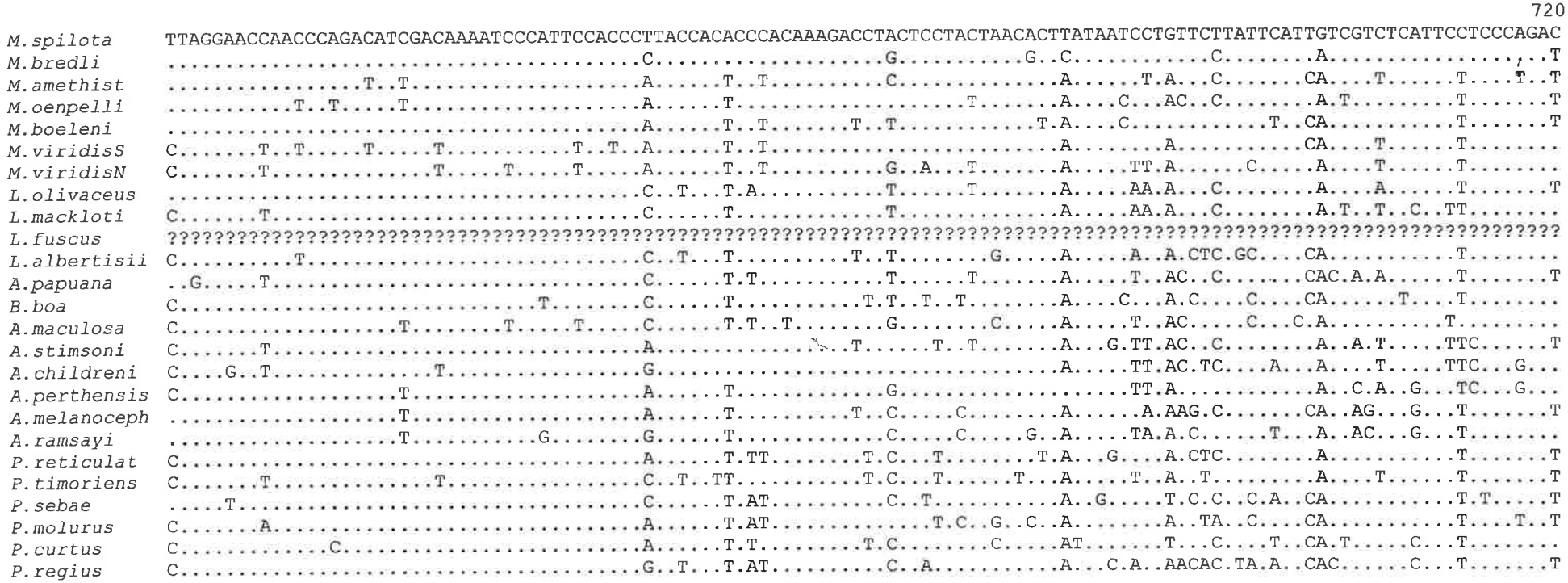
Nucleotide sequences for python taxa used in this study for partial *cytochrome b* sequences. ?=missing data; -=gap. Sequences marked with * are from GenBank. See text for GenBank Accession Nos.

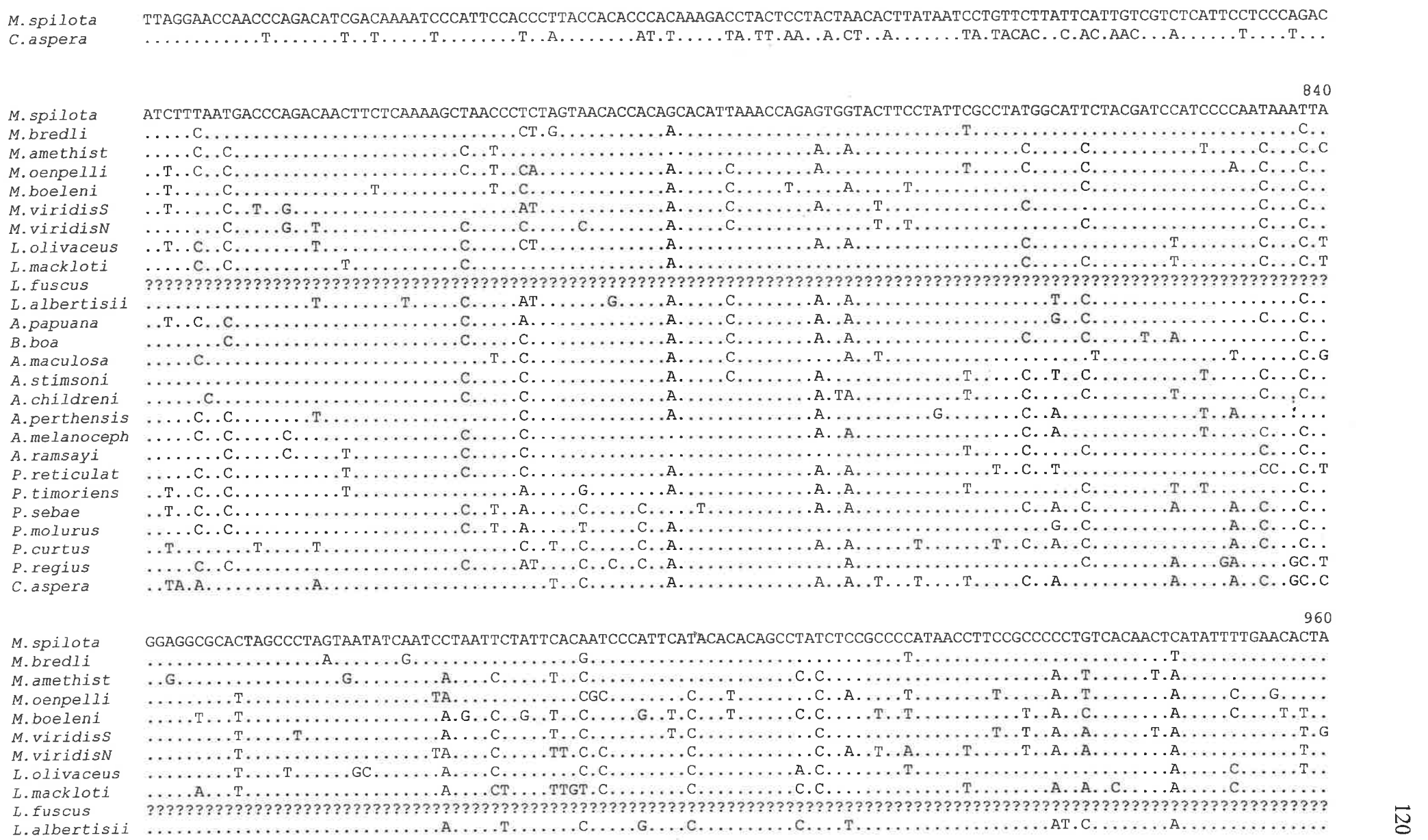
```
                                                                                                                                  120
M.spilota      *ATGCCCCACCACTACATCCTAACCTTATTTGGCCTTCTCCCCGTAGCAACCAATATCTCAACATGATGAAACTTCGGCTCAATACTATTAACATGCCTAGCCCTACAAGTTCTAACCGGC
M.bredli        ????????????????????????????????????????????????????????????????????????................................G...A.C.........
M.amethist     *..................T................C..A..G...........C..T..............................C..........T..A........A.....T...
M.oenpelli      ????????????????????????????????????????????????????????????????????????......................................A.........
M.boeleni       ????????????????????????????????????????????????????????????????????????........T...............T...G.T.......AA........
M.viridisS      ????????????????????????????????????????????????????????????????????????.....y................................AT........
M.viridisN      ????????????????????????????????????????????????????????????????????????............T..........TT.............A.........
L.olivaceus    *.................T......C.G..C.....CT.A..A...........C..T....................T.........C............T.........AT.......T
L.mackloti     *...............G.T......C.G.....T..CT.A..A..............T..............................C.......TT.............A........A
L.fuscus        ????????????????????????????????????????????????????????????????????????...............C.......TT.............ATw......A
L.albertisii   *.................TT.....C..........C..A..............C....................T..T..................T....T.....G..A.........
A.papuana      *...........T............C.G..C.....C..A..A...........C..T..............................C......................AT....T...
B.boa           ????????????????????????????????????????????????????????????????????????..T.........T.............G...........A.........
A.maculosa      ????????????????????????????????????????????????????????????????????????............T..........s..G.s...G.....AT.......A
A.stimsoni      ????????????????????????????????????????????????????????????????????????.......................T..............AT........
A.childreni    *.................T......C....C........G..T...........C.........................................T..............AT.......T
A.perthensis    ????????????????????????????????????????????????????????????????????????.....m.........C.......TT.....T.......A.........
A.melanoceph   *........................C.............G..T........T..C.................................C.......TT...G.........A.........
A.ramsayi       ????????????????????????????????????????????????????????????????????????...............C........T...GTy.......A.........
P.reticulat    *...........T..T.......................A..A...........C........C.....................T..C.......T......T.......A.........
P.timoriens     ????????????????????????????????????????????????????????????????????????.........C.............T..............AT....T..T
P.sebae        *.....A.....T..T...T.....C....C..A..C..A..A...........C..T..............T...............C.......TT.....T.....ACG..C..A...
P.molurus      *..............T....................C..A..A...........C....................................G.....T.............AT.......A
P.curtus        ????????????????????????????????????????????????????????????????????????.....T......T..C.C..T.......T.........A.........
P.regius       *..............T.........C.C..C........A..A...........C.................................C.......T...ATGT.......A..T......
C.aspera       *...........AC.A..A....TTC.......A..A..A..A.....T..T..C..T....................A......T..C....C..TTC...AT.......AA....A..A

                                                                                                                                  240
M.spilota       TTCTTCTTAGCTGTCCACTACACAGCAAACATTAACCTGGCATTCTCATCCATCATTCACATTACCCGAGACGTCCCATACGGCTGGATAATACAAAACCTACACGCCATCGGAGCATCT
M.bredli        .....T.....G....................C.....A.................C.............................A...........s....................C
M.amethist      ......C....C..A...........G...........A.......................C.................T.....A..............G.....T..T........C
M.oenpelli      ......C....C.................T..C.....A.....T........T..C.....C.....T..............T..A................................C
M.boeleni       ......C....C..A.................C...T.A.................C.....C.......................A............T.G.....T............
M.viridisS      ......C....C..T..........s.........T..A.................C.....CT.......T..T........T..A................................C
M.viridisN      ...........C....................C.....A..............T..C..T..C..T.....T..............A................................C
L.olivaceus     ......C.........................C..T..A...............G...............................A....................T............
L.mackloti      ......C.G....................T........................G.......C..T.....T..............A.................................
```

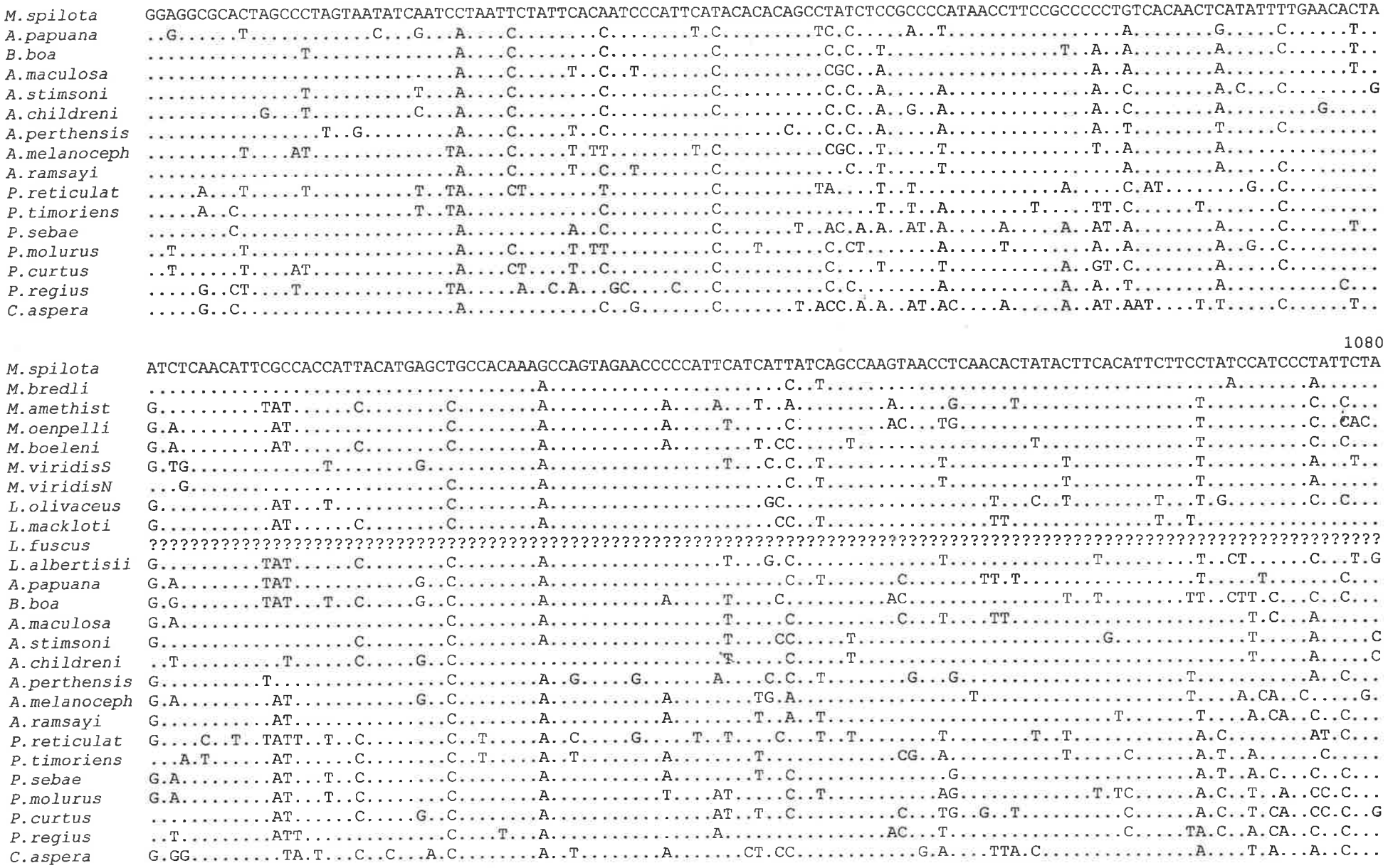

M.spilota TTCTTCTTAGCTGTCCACTACACAGCAAACATTAACCTGGCATTCTCATCCATCATTCACATTACCCGAGACGTCCCATACGGCTGGATAATACAAAACCTACACGCCATCGGAGCATCT
L.fuscus .....yC.G.....k.T.....T...G.....C.....C..T.....A.....C
L.albertisii .....C....C.....C.....A.....T.....C.....T.....T.....A.....C
A.papuana .....C.G..C..A.....C.....A.....C.....T.....A.....T.....C
B.boa .....C.G..C.....C.....T.....C.....T.....T....A.....G.....C
A.maculosa .....G..C.....T.....T..C.....A.....T.....T.....C.....T.....A.....y.....C..C
A.stimsoni ..T...C....C..T..T..T.....A.....T.....G.....A.....C
A.childreni ..T.....T.....T.....A.....T.....G.....T.....A.....C
A.perthensis ..T..T..G..C.....C.....T.....T.....A.....G.....C
A.melanoceph .....C....C.....C.....T...G.....CT.....T.....A.....T..A.....C
A.ramsayi ..y...C....C.....C.....T...G.....C.....T.....A.....C
P.reticulat ..T...C....C.....T.....A.....T.....C.....C.....A.....T.....T.....C
P.timoriens ..T...C....C..T.....T.....C.....T..A..G.....C.....C
P.sebae .....C.....A.....T.....C.T...T.....C.T.....A.....G.....C.....
P.molurus .....C....C.....C.....A.....T.....C.....C..T..T.....A.....T.....C.....C
P.curtus .....C....C.....C.....A.....T..C..T..T..A.....C.....C.....A.....T.....y.....C.....
P.regius .....C.....T.....C.....A.....T.....T..C..T.....T.....C.....AC.....C.....G
C.aspera ....C....T.AA.A..T.....T.....A.....GC.....T..C..T.....A.....C.T..A..AG.....G
360
M.spilota ATATTCTTCATTTGCATCTACATCCATATTGCACGTGGATTATACTACGGATCCTATCTCAACAAAGAAACCTGAATATCCGGCATTACCCTACTCATCACACTAATAGCAACCGCCTTC
M.bredli C.....C.....C.....G.....T.....T...
M.amethist .....C.....T..C..C.....A...C.....A..C..T.....G.....T...
M.oenpelli .....T..T..T..C..C..T..A...C.....T..G..A..C..T.....T..C.....T.....
M.boeleni .....C.....T.....C..C.....A..G..G.....A..C..T.....T.....T.....T.....T.....
M.viridisS .....A.....T.....C.....T.....
M.viridisN .....T.....C..C.....A...C.....C.....G..T.....T..C.....T....G.....
L.olivaceus .....C.....A..TC.....A..C..T.....T..T..C.....T...
L.mackloti .....T.....C..C..C..A..CC.....A..C..T.....T.....G..T.....
L.fuscus .....T.....C..C..T..A..CC.....A..C.....m.....C.....T.....
L.albertisii .....T..C..C.....G..GC.C.....A..C..A.....G.....
A.papuana .....C..T.....G..CC.....G..C..A..T.....T.....C.....
B.boa .....G.A.....T..C..C.....A...C.....G..A..C..A.....T.....G.....G.....
A.maculosa .....T..T.....C...ky..A...C.....T..C.....T.....T..T..C.....T..T.....A.....
A.stimsoni .....T.....T..T..T..T..C..C.....C...C.....T.....C.....T..C.....G.....
A.childreni .....T.....T..T..T..T..C..C.....C...C.....T.....C.....T..C.....G.....
A.perthensis .....T..C.....C...C.C.....C.....G.....G..A.....C.....G.....T.....
A.melanoceph .....C..T.....T..C.....A.....C..T.....A.....C..A.....G.....T...
A.ramsayi .....T.....T..C.....A.....C..T.....G..T.....AT.....T.....
P.reticulat .....C.....C..C.....A..CC.....A..C.....A.....C.....C.....T..T
P.timoriens .....T.....T..T.....C..C.....A..CC.....A...y.T.....T.....A.....C.....T.....T..T..T...
P.sebae .....T.....T.....C..C.....A..CC.....T.....T.....A..T..C..A..C.....TC.....G...
P.molurus .....C.....T..C..C.....A...C.....C.....A..T.....A.....A.....T..G.....
P.curtus .....T.....C.....T.....C..C.....A...C....T.....C.....T.....G..T.....A..C.....T..T


M.spilota ATATTCTTCATTTGCATCTACATCCATATTGCACGTGGATTATACTACGGATCCTATCTCAACAAAGAAACCTGAATATCCGGCATTACCCTACTCATCACACTAATAGCAACCGCCTTC
P.regius .....T..C.....T..C..C.....A...C.....C...C.C.....T.....G.....A..T.....A..T.....G..G.....
C.aspera .....T..T..C..T.....AT.....A...C.....T..C..G..CT.A.....G..A...C....A....CC..A.....A....T...T.....A...
480
M.spilota TTCGGTTACGTCCTCCCATGAGGACAAATGTCATTCT?AGCCGCAACTGTAATTACAAACCTACTCACCGCCGTACCCTACCTAGGCACATCTCTAACAACCTGACTATGAGGCGGGTTC
M.bredli .....T.....T.....A.....rr.....T.....A.....G.....A.....A...
M.amethist .....A.....G.....A.....G.....C.....C.....T.....A.C..G.....T.....C.....T..G.....A..T
M.oenpelli .....A..T..T..T.....G..A.....G.....C.....T.....A.....A.....A.....A...
M.boeleni ..T..A.....T.....A.....G..ss..wr.m..w..C.....T...A.C..T..T.....A.....T.....A...
M.viridisS .....C..T.....A.....G.....C.....C.....T.....T.....G..T.....A.....T.....A...
M.viridisN ..T..C..T.....A.....G...T.....C.....C.....G..T.....AT.....G.....A...
L.olivaceus .....A..T.....T.....G.....A.....GG.....C.....C.....T.....T.....A.....T.....A...
L.mackloti .....A.....T.....A.....G.....C.....T.....CT.G.....G.....G.....
L.fuscus .....G.....T.....A....?..G..??
L.albertisii .....A.....A.....G...T.....C.....C.....A.....A.....T....G.....A..T
A.papuana .....A..T.....T.....G.....C.....T..G.....T..A.....G..T.....A.....T...T.....C...
B.boa ..T..A..T.....G.....G..A.....G.....C.....GT.A..A..A.....G.....A.....GT.G.....
A.maculosa .....T.....r..A.....G.....C.....C.....T....T.....C..A..T.....A.....A...T.....G..C..T
A.stimsoni .....C..T.....A.....G.....C.....C.....A.....G.....G.....G.....
A.childreni .....C..T..T.....A.....G.....C.....C.....A.....G.....A.....G.....A..A...
A.perthensis .....C.....G..A.....G.....A.....C..C.....T.....A.....G.....
A.melanoceph .....A..T..G..TA.....A.....A.G.....C.....C.....G.....A..A...
A.ramsayi .....A..T.....T.....A.....G.....C.....A.....T.....G.....A..A...
P.reticulat .....T.....A.....G.....C.....T.....A.....T.....A.....G..T.....A...
P.timoriens .....A..T..T..T.....A.....Gr.....T..A.....A.....G.....AT.....C.....A..T
P.sebae .....C.....A.....G.....A.....C..C.....T..T.....A...A.....T....G..A..C...
P.molurus .....A..T.....G.....A.....G...T.....C.....C.....T.A.....A...T.....A.C.....GT.....A..A...
P.curtus .....A..T.....G.....G.....C.....C..C..T.....A..T..T.....A.....A.C.....T.....A.....
P.regius .....G..T..A.....A..C....G.....A.....C.....T..T.....G.....G..A.CA....C.....T.....A.....
C.aspera .....A..T..A..A.....C.....A.....G...A.....T..C..T..T.....A..T..AA.T..A.....T..A..TA.AA.G.....T.....A...
600
M.spilota GCAATCAATGACCCCACCTTAACACGATTCTTCGCATTACACTTCATCCTACCATTCGCAATTATCTCTCTCTCCTCACTACACATTATTTTACTTCACGAAGAAGGCTCTAGCAACCCA
M.bredli .....GC.....C.....T.....C.....
M.amethist .....C.....C.....C..T.....T.....C.....CT.A.....C..CC....A.....
M.oenpelli .....A...C....C.....T.....T.....T.A.....T..C..CC....C..T.....T..C.....T...
M.boeleni .....T..T..AC.....T.....C.....T.....C..T..CT.A.....C..CC....C.....T..C.....
M.viridisS .....C.....A...C.....T..T...C.G.....C.....A.....T..TG....CC....C.....C.....T...
M.viridisN .....C.....A...C.C.....T...C....T.....T.....C.....A.....G....CC.T.....C.....
L.olivaceus .....C.....C.....C.....C.....T.A.....C..CC....C.....A.....
L.mackloti .....C.....C.....C.....T.A.....G.C..CC.C.....T.....G..G.....
L.fuscus ??
118


M.spilota GCAATCAATGACCCCACCTTAACACGATTCTTCGCATTACACTTCATCCTACCATTCGCAATTATCTCTCTCTCCTCACTACACATTATTTTACTTCACGAAGAAGGCTCTAGCAACCCA
L.albertisii
A.papuana
B.boa
A.maculosa
A.stimsoni
A.childreni
A.perthensis
A.melanoceph
A.ramsayi
P.reticulat
P.timoriens
P.sebae
P.molurus
P.curtus
P.regius
C.aspera


720
M.spilota TTAGGAACCAACCCAGACATCGACAAAATCCCATTCCACCCTTACCACACCCACAAAGACCTACTCCTACTAACACTTATAATCCTGTTCTTATTCATTGTCGTCTCATTCCTCCCAGAC
M.bredli
M.amethist
M.oenpelli
M.boeleni
M.viridisS
M.viridisN
L.olivaceus
L.mackloti
L.fuscus
L.albertisii
A.papuana
B.boa
A.maculosa
A.stimsoni
A.childreni
A.perthensis
A.melanoceph
A.ramsayi
P.reticulat
P.timoriens
P.sebae
P.molurus
P.curtus
P.regius


M.spilota TTAGGAACCAACCCAGACATCGACAAAATCCCATTCCACCCTTACCACACCCACAAAGACCTACTCCTACTAACACTTATAATCCTGTTCTTATTCATTGTCGTCTCATTCCTCCCAGAC
C.aspera .....T.....T..T.....T.....T..A.....AT.T.....TA.TT.AA..A.CT..A.....TA.TACAC..C.AC.AAC...A.....T....T...
840
M.spilota ATCTTTAATGACCCAGACAACTTCTCAAAAGCTAACCCTCTAGTAACACCACAGCACATTAAACCAGAGTGGTACTTCCTATTCGCCTATGGCATTCTACGATCCATCCCCAATAAATTA
M.bredli .....C.....CT.G.....A.....T.....C..
M.amethist .....C..C.....C..T.....A..A.....C.....C.....T.....C...C.C
M.oenpelli ..T..C..C.....C..T..CA.....A.....C.....A.....T.....C.....C.....A..C...C..
M.boeleni ..T.....C.....T.....T..C.....A.....C.....T.....A.....T.....C.....C...C..
M.viridisS ..T.....C..T..G.....AT.....A.....C.....A.....T.....C.....C...C..
M.viridisN .....C.....G..T.....C.....C.....C.....A.....C.....T..T.....C.....C...C..
L.olivaceus ..T..C..C.....T.....C.....CT.....A.....A..A.....C.....T.....C...C.T
L.mackloti .....C..C.....T.....C.....A.....C.....C.....T.....C...C.T
L.fuscus ??
L.albertisii .....T.....T.....C.....AT.....G.....A.....C.....A..A.....T..C.....C..
A.papuana ..T..C..C.....C.....A.....A.....C.....A..A.....G..C.....C...C..
B.boa .....C.....C.....C.....A.....C.....A..A.....C.....C.....T..A.....C..
A.maculosa .....C.....T..C.....A.....C.....A..T.....T.....T.....C.G
A.stimsoni .....C.....C.....A.....C.....A.....T.....C..T..C.....T.....C...C..
A.childreni .....C.....C.....C.....A.....A.TA.....T.....C.....C.....T.....C...C..
A.perthensis .....C..C.....T.....C.....A.....A.....G.....C..A.....T..A.....
A.melanoceph .....C..C.....C.....C.....C.....A..A.....C..A.....T.....C...C..
A.ramsayi .....C.....C.....T.....C.....C.....T.....C.....C.....C...C..
P.reticulat .....C..C.....T.....C.....C.....A.....A..A.....T..C..T.....CC..C.T
P.timoriens ..T..C..C.....T.....A.....G.....A.....A..A.....T.....C.....T..T.....C..
P.sebae ..T..C..C.....C..T..A.....C.....C.....T.....A..A.....C..A..C.....A.....A..C...C..
P.molurus .....C..C.....C..T..A.....T.....C..A.....G..C.....A..C...C..
P.curtus ..T.....T.....T.....C..T..C.....C..A.....A..A.....T.....T..C..A..C.....A..C...C..
P.regius .....C..C.....C.....AT....C..C..C..A.....A.....C.....A....GA.....GC.T
C.aspera ..TA.A.....A.....T..C.....A.....A..A..T...T....T.....C..A.....A.....A..C..GC.C
960
M.spilota GGAGGCGCACTAGCCCTAGTAATATCAATCCTAATTCTATTCACAATCCCATTCATACACACAGCCTATCTCCGCCCCATAACCTTCCGCCCCCTGTCACAACTCATATTTTGAACACTA
M.bredli .....A.....G.....G.....T.....T.....
M.amethist ..G.....G.....A....C.....T..C.....C.C.....A..T.....T.A.....
M.oenpelli .....T.....TA.....CGC.....C...T.....C..A.....T.....T.....A..T.....A.....C...G.....
M.boeleni .....T...T.....A.G..C..G..T..C.....G..T.C...T.....C.C.....T..T.....T..A..C.....A.....C....T.T..
M.viridisS .....T.....T.....A....C.....T..C.....T.C.....C.....T..T..A..A.....T.A.....T.G
M.viridisN .....T.....TA....C.....TT.C.C.....C.....C..A..T..A.....T.....T..A..A.....A.....T..
L.olivaceus .....T....T.....GC.....A....C.....C.C.....C.....A.C.....T.....A.....C.....T..
L.mackloti .....A...T.....A....CT....TTGT.C.....C.....C.C.....T.....A..A..C.....A.....C.....
L.fuscus ??
L.albertisii .....A.....T.....C.....G....C.....C....T.....AT.C.....A.....
120


M.spilota GGAGGCGCACTAGCCCTAGTAATATCAATCCTAATTCTATTCACAATCCCATTCATACACACAGCCTATCTCCGCCCCATAACCTTCCGCCCCCTGTCACAACTCATATTTTGAACACTA
A.papuana
B.boa
A.maculosa
A.stimsoni
A.childreni
A.perthensis
A.melanoceph
A.ramsayi
P.reticulat
P.timoriens
P.sebae
P.molurus
P.curtus
P.regius
C.aspera
1080
M.spilota ATCTCAACATTCGCCACCATTACATGAGCTGCCACAAAGCCAGTAGAACCCCCATTCATCATTATCAGCCAAGTAACCTCAACACTATACTTCACATTCTTCCTATCCATCCCTATTCTA
M.bredli
M.amethist
M.oenpelli
M.boeleni
M.viridisS
M.viridisN
L.olivaceus
L.mackloti
L.fuscus
L.albertisii
A.papuana
B.boa
A.maculosa
A.stimsoni
A.childreni
A.perthensis
A.melanoceph
A.ramsayi
P.reticulat
P.timoriens
P.sebae
P.molurus
P.curtus
P.regius
C.aspera

```
                                                        1114
M.spilota       GGGTGAATAGAAAACAAAATAATAAACATCTCCT
M.bredli        ..............................
M.amethist      ..A.....................T...C...
M.oenpelli      ..A.....T..................C.CA..
M.boeleni       ..A..G..................G....TC.A.
M.viridisS      ..A.....G.....T............C...
M.viridisN      .....G.......................???
L.olivaceus     ..A.........................C.C...
L.mackloti      ..A...G....G........T......GC.C...
L.fuscus        ??????????????????????????????????
L.albertisii    ..C....C......T............C.CT..
A.papuana       .....G.C...................C.C...
B.boa           ..C.......................A.CTC...
A.maculosa      .....GG.G..................T.......
A.stimsoni      ..A...G.....................T..A.
A.childreni     ..A...G.....................T..A.
A.perthensis    ..A.........................C.....
A.melanoceph    ..A...GC....................T..A.
A.ramsayi       ..A..........T.............T..A.
P.reticulat     ..A.....................C.TC..AA..
P.timoriens     ..C.....................C..C..AAT.
P.sebae         ..A...C.............C...TCA.C.A.
P.molurus       ..A.....T............C.....C.C...
P.curtus        ..C...........T.......C.....CTC...
P.regius        ......G.T??????????????????????????
C.aspera        ......C.T.....T.....T.C...ACA.AA..
```

Appendix 4.2

Summary of characters found by Kluge (1993) to define *Aspidites* that were examined in the present study.

Numbers in parentheses refer to Kluge's character state coding

1. The modal, or median, number of premaxillary teeth in an adult is none (0), on (1) or two (2) per ramus.
5. The ventral openings for the premaxillary channels are located posterior (0) or anterior (1) to the posterior margin of the premaxillary teeth.
31. The anterior head of the dorsomedian end of the postorbial under- or overlaps (0) or abuts (1) the postorbital process of the parietal.
41. Lateral crests on the supraoccipital are absent or weakly developed (0) or prominent (1).
45. The quadrate is tall (>21% of lower jaw length) (0) or short (<22% of lower jaw length) (1).
46. The ventral, horizontal, fluted wing of the vomer, posterior to the septomaxilla, is uniformly wide or varies in width (0) or is uniformly narrow (1).
48. Anterior to the vomeronasal fenestra, the anterodorsal margin of the septomaxilla is even and the vertical lamina is without a fenestra (0), or there exists a fenestra completely enclosed by bone or narrowly (1) or broadly (2) continuous with the dorsal margin.
50. The modal, or median, number of palatine teeth per ramus in an adult is four or less (0), five or six (1), or seven or eight (2).
55. The maxillary process is broad (0), narrow (1), or absent (2).
66. There is a slight (0) or marked (1) change in the height of the dentary teeth.
73. The neck is approximately as wide as (0) or slightly (1) or markedly (2) narrower than the head in adults.
85. The number of posterior sutures in the nasal scale is none (0), one (1), or two or more (2).
91. The first, anteriormost genial scale is short (0) or long (1) compared to other throat scales, particularly the posterior geneials.
95. Most subcaudal scales are single (0) or paired (1).
106. There are none (0), 3 (1), 4 (2), 5 (3), 6 (4), 7 (5), 8 (6) or 9 (7) thermoreceptive pits in the posterior infralabial series of scales.
111. Each posterior trunk intercostal artery usually supplies blood to one (0) or more (1) body segments.

Chapter 5.

Phylogenetic analysis of the green python, *Morelia viridis* reveals a cryptic species.

5.1 Introduction

The Green Python (*Morelia viridis*) is found throughout the island of New Guinea and its offshore islands (with the exception of the Bismarck Archipelago) and in a small area of rainforest in the Iron and McIlwraith Ranges of northern Queensland (Barker and Barker, 1994; McDowell, 1975; O'Shea, 1996) (Fig. 5.1). It has an altitudinal range from 0-2000 metres above sea level, inhabiting lowland and lower montane forests and populated areas (O'Shea, 1996). McDowell (1975) found no significant distinguishing morphological features between populations, with the possible exception that juveniles from the Sandaun Province, PNG populations are brick red rather than the usual yellow or orange (Parker, 1982). However, this may be a rare polymorphism that is yet to be detected in other regions. Anecdotal evidence suggests that colour variation and markings may be diagnostic for some island populations, eg. Aru Is. and Biak Is. (F. Yuwono, D. McCrae, *pers. comm*).

Evidence from a study presented in chapter 7 shows that there is distinct genetic structure within green pythons from eastern New Guinea. Analysis of partial *cytochrome b* (*cytb*) sequence indicates a genetic dichotomy between populations north and south of the New Guinea central mountain range. The age of this division correlates with the uplifting of the central mountain range in the Pliocene (see chapter 7).

As green pythons are also found in western New Guinea and the rainforest habitat of green pythons forms a continuous "ring" around the central mountain range of New Guinea (Johns, 1982; Pratt, 1982) it is possible that the north-south genetic pattern found in the east may extend for the entire length of the island. However, as the island narrows to a mountainous isthmus and then widens into the Vogelkop Peninsula at the western end, there

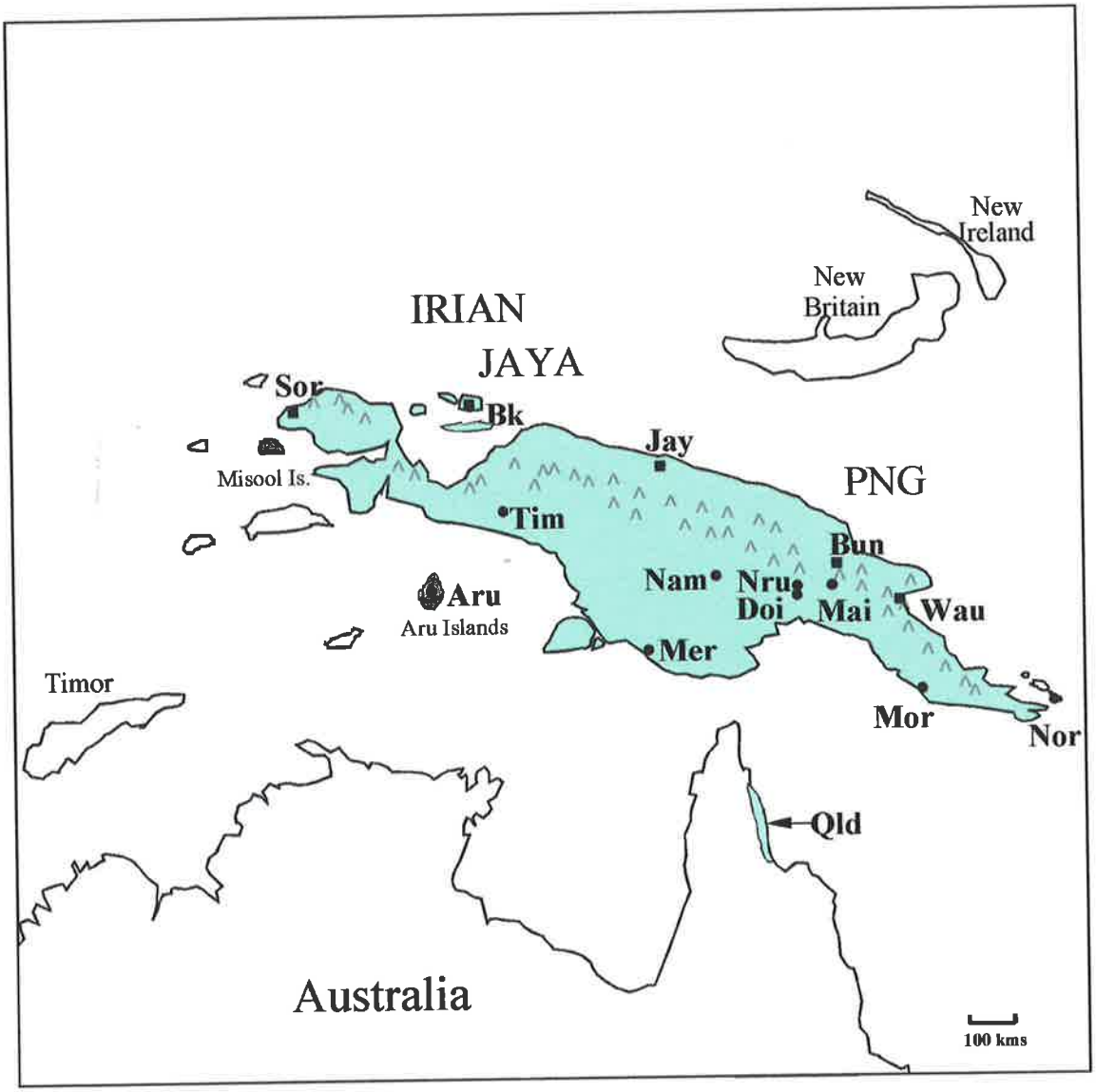


Fig 5.1. Map of northern Australia and New Guinea showing natural range of *Morelia viridis* (shaded area) and sample locations. The locality code is: Aru = Aru Islands, Bk = Biak Is., Bun = Bundi, Doi = Doido, Jay = Jayapura, Mai = Maimufu, Mer = Merauke, Mor = Port Moresby, Nam = Namasado, Nor = Normanby Is., Nru = Noru, Qld = Queensland, Tim = Timika, Wau = Wau. ● = southern populations, ■ = northern populations, ^ represents central mountain range.

may be a disjunction in suitable habitat sufficient to prevent gene flow between populations east and west of the narrow neck of land (Pratt, 1982). With climate and precipitation changes of the glacial/inter-glacial periods causing altitudinal shifts in vegetation patterns (Hope, 1996b; Pasveer and Aplin, 1998), opportunities existed for further separation and differentiation of populations of green pythons. East-west patterns of distribution have been found in some New Guinea species of mid-montane distributed tree kangaroos (Flannery, 1995; Groves, 1990) and birds (Pratt, 1982). If global patterns of climate, geography and vegetation changes caused these patterns of distribution, the distribution patterns of other mid-montane species such as the green python could have been similarly affected.

The distribution of green pythons in a small region of tropical rainforest in northern Queensland, Australia suggests that there was suitable habitat along the land bridge that was exposed between the two land masses during the times of lowered sea-level in the Tertiary and Quaternary periods. Because of the shallow depth of Torres Strait, the land bridge would have been in existence for up to half a million years (Galloway and Löffler, 1972) with the most recent marine incursion starting approximately 8,000 years ago. However, fluctuating climatic conditions would have influenced the vegetation types that inhabited the land bridge (Nix and Kalma, 1972; Read and Hope, 1996). The change from high rainfall/ high humidity to drier atmospheric conditions (Nix and Kalma, 1972) may account for a change in vegetation from the extensive rainforest that was present in the Miocene (Truswell, 1993) to the patches of remnant rainforest and sclerophyll forests of today. The extant rainforest remnants are essentially similar to lower montane forests of New Guinea but there is no northern Australian equivalent of the complex lowland humid rainforests (Nix and Kalma, 1972). Whilst some New Guinea rainforest species migrated into northern Australia and continued southwards to the Wet Tropics region of Atherton Tablelands (Winter, 1997), other New Guinea rainforest species, including *M. viridis,* extended their distributions no further south than the McIlwraith Range (~13°30'S).

The geological history of the Australia-New Guinea region is complex. New Guinea

is a composite of 22 separate terranes that successively accreted and collided with the Australian craton from the Oligocene to the Pliocene as tectonic plates moved Australia northwards and the Pacific island arcs south and westwards (Hill and Gleadow, 1989; Pigram and Davies, 1987). The western Irian Jayan terranes are different from the remainder of the island in that eight of the ten are of Gondwanan origin (Pigram and Davies, 1987). They were detached from the Australian craton by the Early Cretaceous and then had an independent history until at least until the Miocene.

With the evolutionary history of Australian and New Guinean pythons still undetermined, the possible ancestral population of green pythons could have been Australian, which entered New Guinea when the geological terranes from the north docked against the Australian craton (Pigram and Davies, 1987). Alternatively, the ancestral population could have been situated on one of the terranes that accreted to form New Guinea and then proceeded to move southwards into the Australian continent. Subsequent to either of these scenarios, northern and southern populations became isolated from each other as mountain orogeny proceeded (see chapter 7).

Aside from the natural history interest, genetic relationships within green pythons also have interest and applications in two other areas. Firstly, the striking appearance of green pythons makes them popular exhibits in zoos and in the pet trade. Much of the pet trade is supplied from Irian Jaya, although there is now considerable success with captive breeding in the US and Europe from Irian Jayan stock (D. Barker, *pers. comm*; D. MacCrae, *pers. comm.*). The Australian green python has had limited captive breeding success (Barker and Barker, 1994) and Australian law prohibits the unregulated capture of wild snakes. These factors combining with its very limited Australian range, have initiated the green python being given special status for conservation management (Banks, 1998). In a global effort to conserve biodiversity, zoological and fauna parks of the Australasian region established the Australasian Species Management Program (ASMP) to determine conservation selection criteria and appropriate management strategies for species at risk (Banks, 1998; Banks and

Meikle, 1994). These management programs are overseen by Taxon Advisory Groups (TAGs). Two of the main TAG goals for the green python are to clarify suggested differences between Australian and New Guinea snakes and to achieve captive breeding of Australian specimens. In earlier years, before quarantine restrictions closed off legal importation, green python exhibits for Australian zoos and fauna parks were imported from New Guinea as well as collected from Queensland. Therefore, to establish a breeding program in accordance with TAG guidelines, the origin of green pythons already in zoos and fauna parks needs to be determined.

Secondly, the rarity of the green python in Australian herpetologists' collections (and consequent high market price) and the strict regulations on importing exotic specimens into Australia have encouraged the illegal importation of green pythons into Australia (AQIS Press release, August 27, 1998). Currently, these illegal imports can easily be absorbed into the Australian pet trade undetected. The development of molecular makers that can be used to provenance or identify individual animals would ensure better quarantine management.

In the present study, I further investigated the genetic structure of green pythons by incorporating Australian and western New Guinean populations. I used nucleotide sequences of two mitochondrial genes and incorporated allozyme electrophoresis data (provided by S. Donnellan) to examine whether the "north/south" dichotomy extends along the length of the island of New Guinea and whether the Australian population could be distinguished from southern NG populations. I also considered the biogeography of the species and whether there is any east/west genetic partitioning within New Guinea.

5.2 Materials and Methods

5.2.1 Specimens examined

Fifty-four *Morelia viridis* were collected from 16 locations from across the species range in New Guinea and the Iron and McIlwraith Ranges of northern Queensland, Australia (Fig. 5.1). * = samples were used for allozyme electrophoresis. Tissue samples used in this

study were: **Australia, Mv1-18:** Iron Range, Queensland, ABTC 65592, ABTC 65605, ABTC 67593-4, ABTC 67596, ABTC 67627-34, QM J66805, NMVT888, Lockhart River Mission, Queensland, ABTC 51497-8*, McIlwraith Range, Queensland, QM CJS919; **Southern PNG,** Simbu Province, **Doi1-3:** AMS R115348-50*, **Nru1-2:** AMS R115355-6*, Southern Highlands Province, **Nam1-2:** AMS R122363-4*; Milne Bay Province, **Nor1:** AMS R129716*; Central District, **Mor1:** ABTC 68320; Eastern Highlands Province, **Mai1:** ABTC 67151; **Northern PNG,** Madang Province, **Bun1:** AMS R124531*; Morobe Province, **Wau1-2:** BPBM 11617*, BPBM 13798*; **Southern Irian Jaya, Aru1-4:** ABTC 66380-1, ABTC 68312-3; **Mer1-6:** ABTC 66384-5, ABTC 67170, ABTC 67172, ABTC 68315-6; **Tim1:** ABTC 66387; **Northern Irian Jaya, Bia1-4:** ABTC 66388, ABTC 67175-6, ABTC 67182; **Jay1-2:** ABTC 66383, ABTC 66386; **Sor1-6:** ABTC 67173-4, ABTC 67183, ABTC 68318-9, ABTC 66379. *Morelia spilota*, SAMA R26878, was used as the outgroup.

5.2.2 Mitochondrial DNA

DNA was extracted from frozen or alcohol preserved liver or blood, body scales or sloughed skin as per protocols in chapter 2.2. PCR primers L14841 and H15149 (see Table 2.1) and PCR protocols from chapter 2.7.1 and 2.7.3.1 were used to amplify a 350 bp *cytb* product. To amplify a partial *CR* product, whilst avoiding the amplification of the *CRL* fragment, nested PCR was done as per chapter 2.8. PCR primers L15926 and H690 (see Table 2.1) and PCR protocols from chapter and 2.7.2 and 2.7.3.2 were used for the *CR* product. Sequencing reactions in both directions of the *CR* product tended to stall at a C-rich region (described by Kumazawa *et al.*, 1998). Additional nested primers were used to amplify sequence across this region in each direction. These primers were Snake 1, Snake 6 and Snake 7 (see Table 2.1).

Sequences were aligned by eye and only 12 insertion/deletion (indel) events were inferred. Aligned sequences were phylogenetically analysed with PAUP*4.0b2a (Swofford, 1999). Aligned sequences are presented in Appendix 5.1. To determine whether the two data

sets should be combined, the ILD test (Farris *et al.*, 1995) (called the Partition Homogeneity Test) was performed in PAUP*4.0b2a. The potential for the *CR* primers to amplify mitochondrial genes rather than nuclear paralogues (Zhang and Hewitt, 1996) was tested using the protocol described in chapter 2.4.

Although the sample size for green pythons is small, I have used the mismatch distributions (Rogers, 1995; Rogers and Harpending, 1992) and raggedness indices (Harpending *et al.*, 1993) of the northern, southern and Australian populations calculated in DnaSP ver 3.00 (Rozas and Rozas, 1999) to infer population history. The shape of a mismatch distribution can give an indication if recent demographic history of the population sampled (Rogers and Harpending, 1992). If the population has undergone recent expansion the mismatch distribution has a distinct uni-modal, wave-like form. Populations that have been stable for long periods of time have a ragged, multi-peaked form. Also, the smaller the initial population size, the steeper will be the leading face of the wave. The distribution pattern of a population experiencing a bottleneck is similar to that produced by a sudden expansion model as the population recovers to a new equilibrium, but can be distinguished by elevated probabilities in the tail-region of the distribution. An L-shaped distribution will reflect the early stages of a reduction in population size, but the shape of the curve quickly flattens as the population converges to a new equilibrium (Rogers and Harpending, 1992). The observed mismatch estimators were compared to simulations of population growth estimates under a range of possible population histories were done by S. Chenoweth (Griffith University, Queensland) using the program Mismatch (Rogers, 1998) on a Unix computer. To calculate confidence regions, the Mismatch program simulates 1000 replicate data sets and estimates two parameters of the Mahalanobis distance (M, the vector of means, and C, the covariance matrix) using the two variables $\log_{10}\theta_0$ (MSE) and τ (Rogers, 1995). The Mahalanobis distance (D) is then calculated using these two parameters. The number of simulated data sets (n) for which the simulated D is at least as large as the observed D and the simulated MSE is at least as small as the observed MSE are counted. If $n/1000$ is less than or equal to 0.05, the

distribution is rejected as being outside the 95 percentile of confidence that the parameter estimates are correct.

5.2.3 Allozyme Electrophoresis

Allozyme electrophoresis was done by S. Donnellan for 12 specimens of *M. viridis* according to the methods of Richardson *et al.*, (1986) (Table 5.1). A phenogram was generated by the unweighted pair group method using arithmetic averages (UPGMA) from a matrix of percentage of fixed allele differences from 36 loci. Phylogenetic analysis was done using maximum parsimony criterion of optimality, with loci as characters and alleles as unordered character states. Polymorphisms were treated as uncertainties as per the recommendations of Kornet and Turner (1999).

5.3 Results

5.3.1 Mitochondrial DNA

Partial *cytb* and *CR* sequences from purified mitochondrial DNA were compared with sequences amplified from total cellular DNA for *Morelia viridis* AMS R115348. The two sequences for each gene were indistinguishable.

One *M. spilota* and thirty-six *M. viridis* samples were amplified for both partial *cytb* and partial *CR* genes to give fragments of approximately 350 bp and 850 bp respectively. In addition, specimens Aus15, Jay1, Mer1 and Nru1 were typed only for *cytb* and 14 specimens, Aus 3-13, Sor2 and Sor6 were typed only for *CR*. Among the 18 individuals sequenced either for *cytb* or *CR*, there were no additional haplotypes that were not already observed among the 36 individuals sequenced for both genes. From 40 individuals sequenced for *cytb* there were 14 haplotypes and from 50 individuals sequenced for *CR* there were 33 haplotypes. Unique haplotypes were represented in the combined data that were used for the final analyses. Sequence alignments are presented in Appendix 5.1.

Of 292 nucleotide sites of aligned *cytb* sequence, 64 were variable, and 42 were

Table 5.1 Allele frequencies, expressed as a percentage, in 7 OTUs of *Morelia* at 36 loci. Alleles are designated alphabetically, with 'a' being the most cathodally migrating allele. Where enzymes are encoded by more than one locus, the loci are designated numerically in order of increasing electrophoretic mobility. Where the allele frequencies are not given, the frequency is 100. The number of individuals sampled from each population (N) is given at the head of each column, except where except when fewer individuals were successfully typed, in which case N is indicated by the number in superscript beside the first allelic frequency entry for a locus. The following loci were invariant: *Acoh-1, Acp, Ak-1, Ak-2, Ca, Est, Fbp, Fumh, Gda, Iddh, Idh-2, Lap, Ldh-1, Ldh-2, Mdh-1, Mdh-2, Ndpk, PepD, Pgam, Pgdh* and *Pgk*. For *Aat-1, Lgl* and *Mpi*, OTUs Wau-Qld had the b, a and a alleles respectively and *M. spilota* had the a, b and b alleles respectively. OTU codes are Wau, Nor = Normanby, Nam = Namosado, Doi = Doido, Nru = Noru, Aus = Australia and spl = *M. spilota*.

Locus N	Wau 2	Nor 1	Nam 2	Doi 3	Nru 2	Aus 1	spl 1
Aat-2	a	c	c	c	c	-	b
Acoh-2	c(75) a(25)	b	b	b	b(75) a(25)	-	b
Ada	c	d	d	d	d	d	b(50) a(50)
Eno	b	b	b	b(17) a(83)	b(75) a(25)	b	b
Gpi	b	b	b	b	c(25) b(75)	b	a
Gr	b(50)[1] a(50)	b(50) a(50)	b	b(17) a(83)	a	b	a
Idh-1	c(25) b(75)	b(50) a(50)	a	a[1]	a[1]	a	a
PepA	b(75) a(25)	b	b	b	b	-	b
PepB1	c	b(50) a(50)	b(50) a(50)	b(83) a(17)	b(50)[1] a(50)	-	c
PepB2	b	a	a	a	a	a	-
Pgm-1	b(25) a(75)	a	b(25) a(75)	b(83) a(17)	b(25) a(75)	a	b
Pgm-2	b	a	a	a	a	-	b

parsimony informative. 786 nucleotide sites of aligned *CR* sequence (47 bp of $tRNA^{Pro}$ and 739 bp of *CR*) had 243 variable sites, with 87 parsimony informative. The Partition Homogeneity Test was not significant (P = 0.61), indicating that the estimate of sequence heterogeneity was not greater than would be expected due to chance alone and therefore, the two data sets were combined for analysis. The combined data sets of 1078 aligned sites (307 variable, 129 parsimony informative) were analysed by MP, NJ and ML. For maximum parsimony analysis, gaps were treated as a fifth character state. From a weighted heuristic search, a strict consensus tree was constructed from the 208 equally most parsimonious trees found. The model of nucleotide substitution found for the combined data set, using Modeltest3, was HKY85+Γ (Hasegawa *et al.*, 1985). Parameters estimated using successive approximations (as per chapter 2.12.3) for the HKY85+Γ model for the combined data set were: nucleotide frequencies A = 0.271227, C = 0.287991, G = 0.131905 and T = 0.308877, ts/tv ratio = 2.3341 and Γ = 0.2929. These parameters were then applied to a ML bootstrap analysis and corrected distances were used under the HKY85+Γ model for NJ analysis. The ML tree is shown in Fig. 5.2 with bootstrap proportions for MP, NJ and ML analyses indicated. Phylogenetic reconstructions for all three methods of analysis of mtDNA showed two monophyletic lineages supported by bootstrap pseudoreplicate proportions greater than 98%. One lineage, designated the southern lineage, comprised the southern New Guinea and Australian populations and the other, designated the northern lineage, comprised the northern populations. All three analyses also had the Australian specimens forming a monophyletic clade within the southern lineage with bootstrap pseudoreplicate support greater than 95%.

There was polyphyly of populations within the northern and southern lineages. Some of the western Sorong and Biak samples clustered with the eastern Wau and Bundi samples in the northern clade, and amongst the southern populations some samples from Merauke clustered with those from Aru Is., whilst other samples from Merauke grouped with individuals from Namasado. Relationships among haplotypes within the northern populations

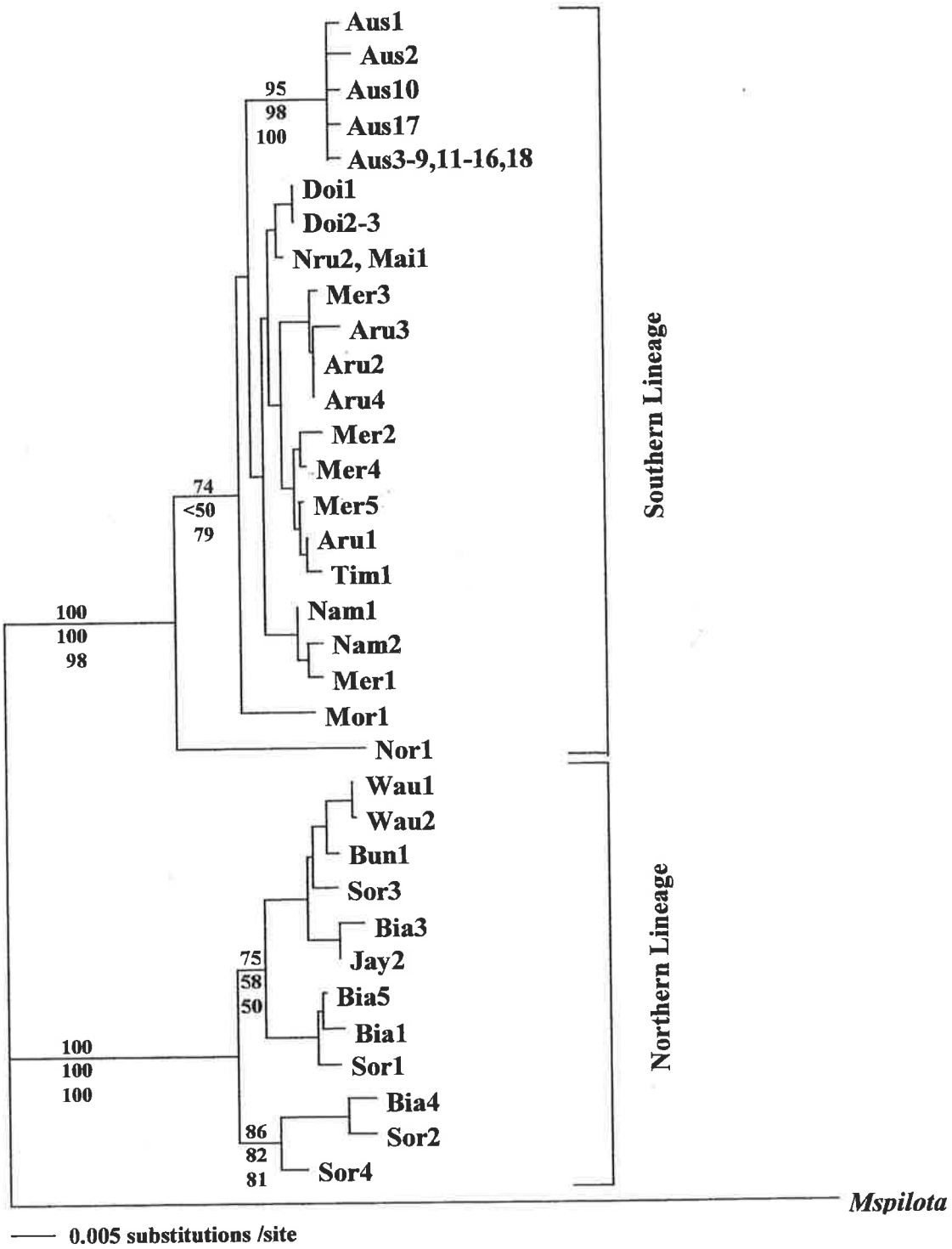


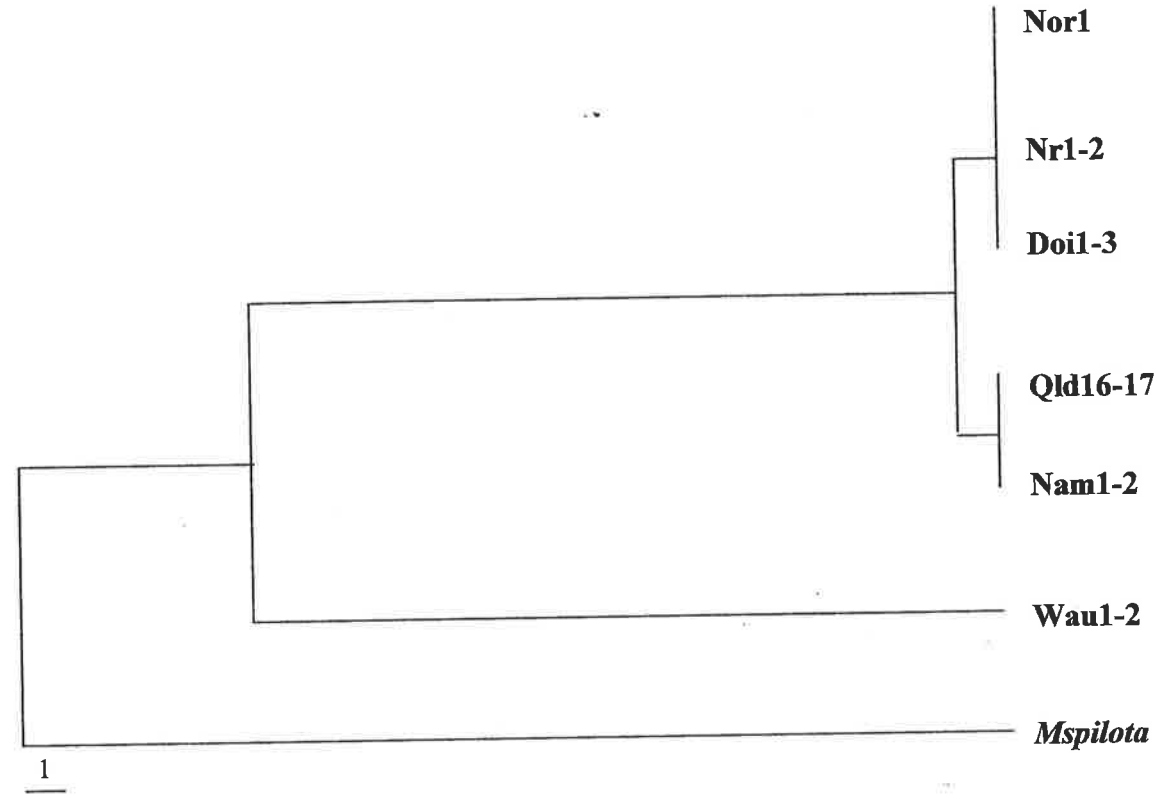
Fig. 5.2 Maximum likelihood phylogram of evolutionary relationships of Australian and New Guinea populations of *Morelia viridis*. Proportions of bootstrap pseudoreplicates are indicated for MP, ML and NJ analyses in descending order respectively.

were well supported, but much of the population structure amongst the southern populations was not supported by bootstrapping in MP and ML analyses.

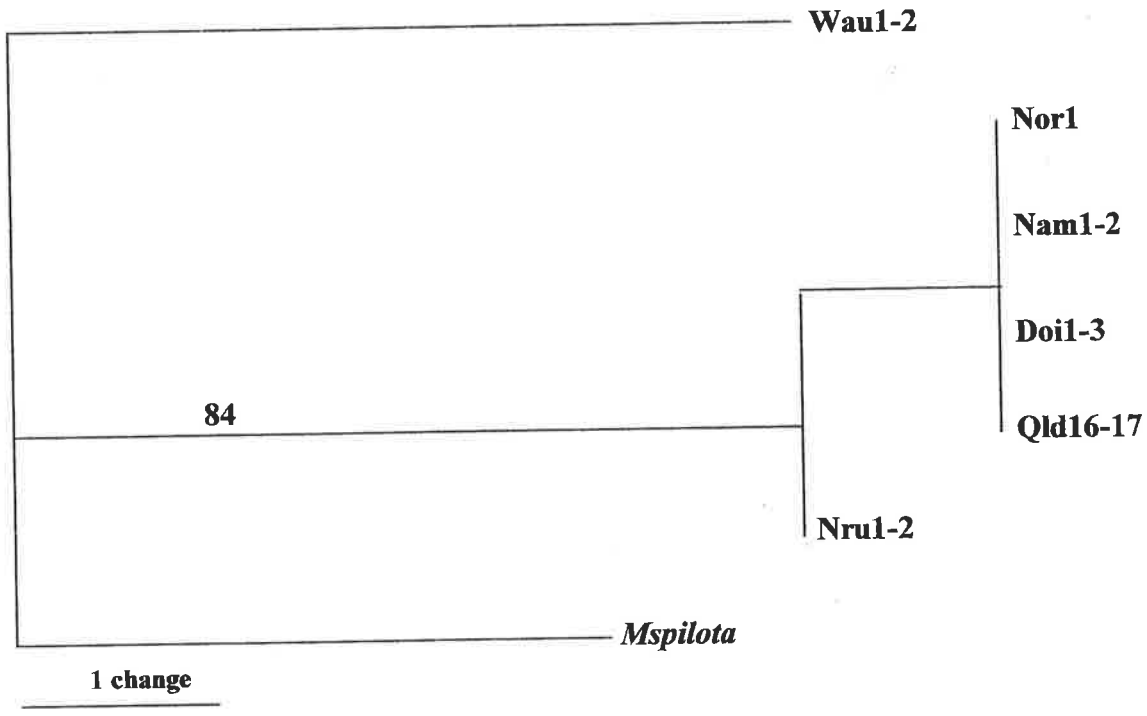
For the combined sequence data, there were four synapomorphies for the Australian clade - three transitions and an indel. One of the transitions was found in the *cytb*, and one was in $tRNA^{Pro}$. There was a C-T transition unique to the southern New Guinea individuals. There was also an A-G transition that was found in all the southern New Guinea animals except for the individual from Normanby Island (Nor1) which had an A at this site. Nor1 is the most divergent individual in the southern lineage, averaging more than 3% sequence divergence from all the other southern animals. Whilst this sequence has a character state that is synapomorphic for the southern lineage, it also shares 5 nucleotide substitutions that would otherwise be synapomorphies of the northern clade, one of which is an indel. There were 57 synapomorphies for the northern New Guinea lineage, (31 transitions, 23 transversions and 3 indels) 16 of these were in *cytb* and one in the $tRNA^{Pro}$. There was an average of 8.4% uncorrected sequence divergence between the northern and southern lineages for transitions and transversions and 1.5% divergence between Australia and southern New Guinea. Within the clades, mean sequence divergence was 0.2 % for Australia, 1.0% for southern New Guinea and 1.4 % for northern New Guinea.

5.3.2 Allozymes

Allele distributions at the 36 loci resolved are shown in Table 5.1. These data were converted into a matrix of percentages of loci showing fixed differences between OTUs. A fixed difference occurs at a locus when the two OTUs being compared share no alleles (Richardson *et al.*, 1986). UPGMA analysis of allozyme fixed differences showed the distinction between the northern and southern lineages (Fig. 5.3a). Phylogenetic analysis of the allozyme data using the maximum parsimony criterion of optimality showed the dichotomy of the northern and southern lineages with bootstrap support of 84% pseudoreplicates (Fig. 5.3b). Five loci showed fixed allelic differences between the northern



a.



b.

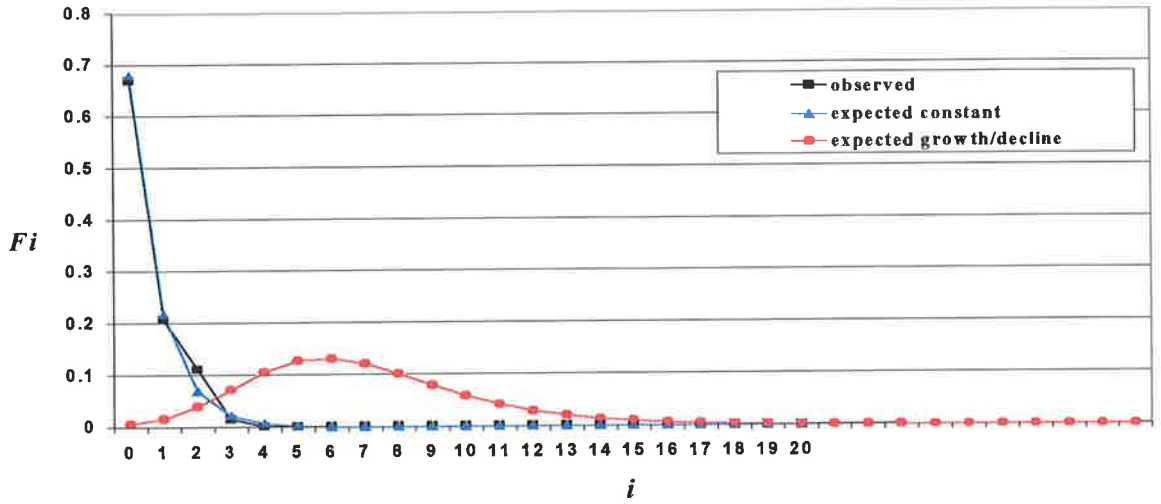
Fig. 5.3 Analysis of allozyme electrophoretic differences. a. A UPGMA tree of the percentage fixed allelic differences between OTUs of *Morelia viridis*. The scale bar represents 1% fixed allelic differences. b. One of twelve equally most-parsimonious trees constructed using branch and bound search method for maximum parsimony. Numbers represent proportion of bootstrap pseudoreplicates.

and southern lineages. For *Aat-2*, the northern lineage had the *a* allele, whereas southern New Guinea had *c* and the Australian population could not be typed for this locus. For *PepB1* and *Pgm-2*, the northern lineage had alleles *c* and *b* respectively, whereas within the southern lineage, the southern NG populations had alleles *ab* and *a* respectively, whilst the Australian population could not be typed for either locus. For *Ada* and *PepB2*, the northern lineage had alleles *c* and *b* respectively, whereas all other *M. viridis* had *d* and *a* respectively. For *Acoh-2* and *Idh-1*, the northern lineage had a unique allele *c* that was not fixed in the population and for *Gpi*, the Noru population had a unique allele *c* that was not fixed in the population. Five loci could not be typed for the Australian population and there was little differentiation within the southern lineage, with the Australian population sharing alleles with the southern New Guinean populations at the remaining 31 loci.

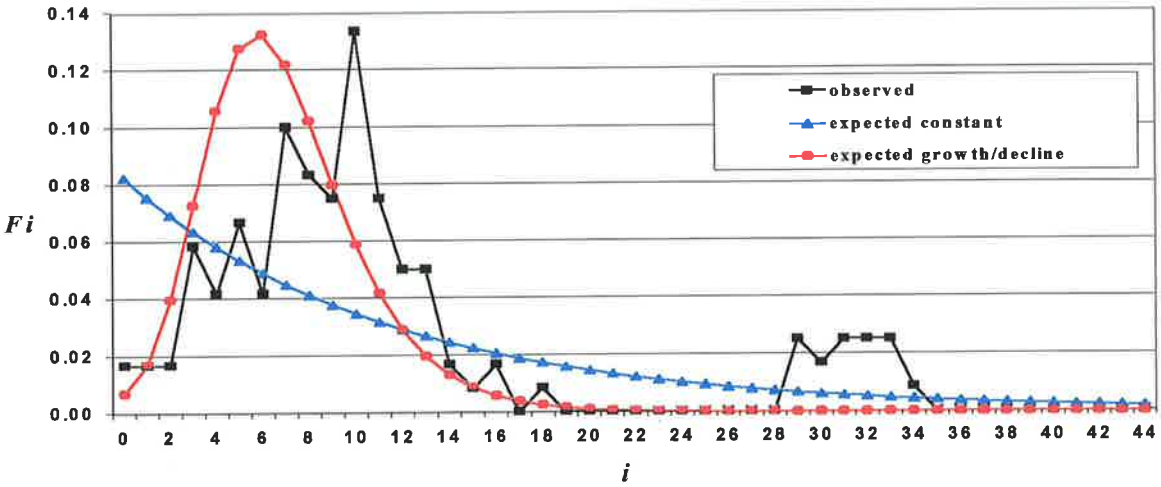
5.3.3 Mismatch distributions

Mismatch distributions, which measure the frequency distributions of the number of nucleotide sites at which all pair-wise combinations of sequences differ (Hartl and Clark, 1997), were determined between individuals for the Queensland, southern NG and northern NG populations using DnaSP (Rozas and Rozas, 1999) and distribution curves are shown in Fig. 5.4. The mismatch distribution for the Queensland population formed an L-shaped curve that showed a good fit to the expected distribution for a population in equilibrium. For the southern NG population, the curve was ragged with a steep leading edge. There was a peak in the tail of the curve that was not present when the Normanby Is. sample (NOR1) was removed, suggesting that the peak was representative of the high level of divergence of the Normanby Is. sample. The mismatch distribution for populations of the northern lineage had a ragged distribution with a much shallower leading edge. The distribution curves of both the southern and northern populations reflected the expected curves for population growth/decline.

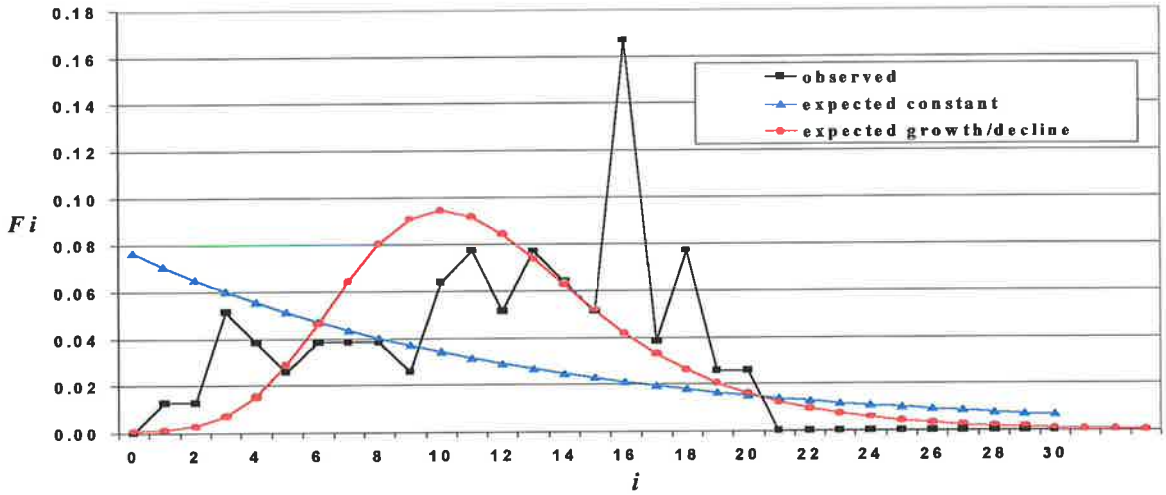
The confidence region of each observed mismatch distribution was assessed by



a.



b.



c.

Fig. 5.4 Graph showing the observed probability that two random neutral sequences will differ at exactly *i* nucleotide sites (*Fi*) against the number of nucleotide sites difference (*i*) for three populations of *M. viridis*. a. = Queensland, b. = southern New Guinea, and c. = northern New Guinea. The expected values are shown in red for a constant population size and blue for a growing or declining population size.

comparing it with 1000 simulated populations of constant size for each of the three *M. viridis* populations. To calculate the mismatch distributions, three parameters, θ_0, θ_1 and τ, were estimated using the method of moments (Rogers, 1995), where θ is the measure of population size and $\theta = 2Nu$ where N is the effective population size and u is the per-generation probability of a mutation anywhere in the sequence (*i.e.* the sequence mutation rate, as opposed to the nucleotide mutation rate. The mutation rate for a sequence m_T nucleotides is equal to $m_T\mu$, where μ is the mutation rate per nucleotide). Time and population size are measured in mutational units τ, such that $\tau = 2ut$, where t is time in generations and u is the per-generation probability of a mutation anywhere in the sequence. Raggedness (r) is measured by the sum of squared differences between successive entries of the empirical mismatch distribution. An r (obs) of ~ 0.012 suggests there has been a sudden expansion of the population, and an r (obs) of ~ 0.26 (> 0.03) is more consistent with the population size being constant for a long time (Harpending *et al.*, 1993).

The observed value of r (obs) for the Queensland population was 0.2331, τ was 0.168 and θ at time zero was 0.302. The measure of confidence of the mismatch distribution for the Queensland population was 548/1000 in the 1000 fold growth run and 712/1000 in the zero growth run, both of which are > 0.05, suggesting that we can accept with confidence that there was no evidence of a population expansion or decline. The r (obs) for Southern NG was 0.0172, τ was 3.394 and θ at time zero was 7.714. In the test of confidence for the southern NG populations in the 1000-fold growth run, the measure was 976/1000, which is not significant, suggesting evidence for a population expansion in southern PNG which was probably more severe than 1000 fold. However, the null hypothesis of no growth could not be rejected with a confidence region of 177/1000, which is also not significant. For Northern NG, the observed value of r (obs) was 0.0401, τ was 8.431 and θ at time zero was 3.557. Confidence regions for this population were estimated at 360/1000 for the 1000-fold simulations and 339/1000 for the zero growth simulations, neither of which are significant

and therefore the values estimated from the data cannot be rejected. This suggests that there was no evidence of population growth/decline and the population is likely to be in equilibrium. However, it must be noted that the confidence regions produced for expanding/shrinking populations are much narrower than for equilibrium populations (Rogers 1995). Populations in equilibrium are usually recognized by the roughness of the mismatch distributions (Rogers 1995).

5.4 Discussion

5.4.1 Patterns of diversity and biogeography

For green pythons examined in the present study, in the analyses of the mitochondrial and nuclear partitions, there were two reciprocally monophyletic clades separating individuals from northern New Guinea and southern New Guinea/Australia, with very strong support from high bootstrap proportions. The high sequence divergence and number of allozyme and nucleotide allelic fixed differences between the northern and southern lineages of *M. viridis* suggest that there has been no gene flow between populations north and south of the central mountain range for a substantial period of time. The presence of two such very distinct genetic lineages would beg the question as to whether the lineages actually represent two species of green pythons.

Traditionally, species were defined by the Biological Species concept (BSC) proposed by Dobzhansky (1937) and Mayr (1969) in which "species are groups of interbreeding natural populations that are reproductively isolated from other such groups". Some species don't interbreed because of natural geographical or chronological isolation, but they will interbreed and produce fertile offspring if placed in an artificial situation where males and females of each species come into contact with each other. Therefore, standing alone, the ability or not to interbreed may not be sufficient to define species. Also, Mayr outlined the difficulties of the BSC in determining the species status of allopatric populations. Some of the green python specimens used in this study represent island populations, invoking the difficulties of

allopatry if the BSC is used. Consequently an alternative approach to species definition is considered here.

The Evolutionary Species Concept proposed by Simpson (1951) states that a species is "a phyletic lineage (ancestral-descendent sequence of interbreeding populations) evolving independently of others, with its own separate and unitary evolutionary role and tendencies." This concept can be used to cover extinct and asexual groups. Species status is usually decided on patterns of phenotypic cohesion within a population as opposed to differences between populations. However, as this system is equivocal on the ranking of importance of each phenotypic trait it leads to inconsistencies (Avise and Ball, 1990).

The Phylogenetic Species Concept (PSC) of Cracraft (1983), focuses on evolutionary units that emphasise the results of evolution (*i.e.* the pattern of phylogenetic relatedness between organisms) rather than the process that brought about the separation. A phylogenetic species is defined as "the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent". The diagnostic character must have a genetic basis and must be defined within the community, not just in subsets of the community, ie. males, females, stages of life cycle, morphs, etc. A difficulty in applying this species concept lies in the somewhat arbitrary recognition of "the smallest diagnosable cluster of individual organisms".

In an attempt to combine the strengths of the BSC and the PSC, and to avoid their weaknesses, Avise and Ball (1990) devised the Concordance Principles (CP) concept. Under this species concept, species should be recognised by evidence of concordant phylogenetic partitions at multiple independent genetic attributes. In any individual, allelic phylogenies may differ from locus to locus due to stochastic lineage sorting of alleles and any number of individuals could group together differently depending on which locus is considered (Avise and Ball, 1990). However, if some evolutionary force such as reproductive isolation (intrinsic, *i.e.* biological, or extrinsic, *i.e.* geographical) pushes the gene genealogies towards a status of monophyly, congruent evolutionary patterns may occur across loci, causing alleles

to separate and cluster in a pattern reflective of their population history. Therefore, under natural conditions, phylogenetic partitioning could occur due to the separation of populations by chronological or geographical barriers as well as by intrinsic reproductive isolating barriers. Hence, there may be populations of organisms which under the BSC are classified as a single species that could be given separate taxonomic recognition under the CP.

Because some of the populations of green pythons are allopatric and therefore, species definition under the BSC could not be tested, I have adopted the Concordant Principles species concept for the present study. Under the CP (Avise and Ball, 1990), the pattern of relationships found for mitochondrial and nuclear genes would suggest the likely existence of two species of *Morelia viridis*, one present north of the central cordillera and the other present in southern New Guinea and Australia.

Because sampling of the eastern tip of New Guinea is very sparse, I am not able to define the geographic limits of the northern and southern species. The Huon Peninsula may be a candidate region of barrier to gene flow. Colgan *et al.* (1993) reported a possible zoogeographic barrier for lowland mammal species existing in the Huon Peninsula region and many lowland bird species which occur throughout the rest of New Guinea are not found in the lowlands of the Huon Gulf and Markham Valley. This is despite an apparent continuity of low elevation rainforest habitat through the Ramu-Markham divide (Pratt, 1982).

McDowell (1975) examined specimens from populations north and south of the central cordillera and found that scale counts were subject to geographic variation. However, there was no broad concordant geographic pattern apparent across a number of characters that would suggest morphological differentiation. Therefore, the northern and southern green python populations can be considered to be cryptic species.

Within each of the New Guinea lineages, there does not appear to be strong east/west genetic structure. In the northern lineage, genetic pattern does not correspond to geographic location, with haplotypes from Biak and Sorong present in all three minor clades of the phylogenetic tree (Fig. 5.2). There is also no distinct separation of eastern and western

populations in the northern lineage. One of these clades within the northern lineage contains representatives from each population sampled, covering two-thirds of the east-west range of the island.

In the mtDNA analyses, within the southern New Guinea clade, all the Australian individuals clustered together and this clade was well supported by high bootstrap and reliability values. There is a single character that would appear to be diagnostic of southern New Guinea populations, which in conjunction with four characters present in the Australian population, can be used to distinguish the two southern forms. The Australian population shows little haplotype diversity.

5.4.2 Population history

Mismatch distributions calculated from mtDNA alleles have been used to estimate times since population divergence and starting population size in primates (Goldberg and Ruvolo, 1997; Harpending, 1994; Rogers, 1995). Rogers (1995) admits that models used are "unrealistically simple" and the confidence intervals around the estimates are very large, thus making these divergence estimates for primates very rough. Also, initial population expansion tends to obscure the effects of later expansions for a very long time, therefore, the timing of sequential changes in population size would be difficult to determine (Rogers, 1995). As θ and τ have been calculated from mean mutation rates for primates, it may be inappropriate to use the values estimated from python sequences to extrapolate divergence times. However, the signals of population growth and decline can still be detected from the mtDNA sequences.

The small sample sizes in this study may have reduced the robustness of the statistical interpretation of the raggedness indices for each of the three populations of green pythons. The southern New Guinean population showed signs of population growth, but this evidence was not strong because a zero population growth scenario could not be excluded. However, estimates of past population history can be extrapolated from observations of the graphs of

raggedness index relative to expected outcomes (Fig. 5.4). The Australian population showed a very steep L-shaped curve that fits to the expected curve for a population that has had a dramatic reduction in size in relatively recent time (Rogers and Harpending, 1992). It is not yet reflecting a period of population growth from its newly established equilibrium. There is no evidence of population growth in the Australian curve. If green pythons evolved in New Guinea, it is possible that the Australian population has been through a bottleneck that would occur if only a small number of individuals made it across one or more of the land bridges of Torres Strait that formed each time an Ice Age caused a drop in sea level (ie. a founder effect). Because mismatch signals associated with later changes in population size are masked by earlier signals (Rogers, 1995), it is not possible to predict which era the reduction in population size may have occurred. However, if the migration across the land bridge was early, it may be possible to see evidence of population expansion in the mismatch signal as the Australian population established itself. Alternatively, the population of green pythons in Australia may have been larger but the population size was determined by the expansion and contraction of the rainforest brought about by climatic changes associated with the Ice Ages. If Cape York is reflective of Atherton Tableland, forests have continued to retract since the Ice Ages, thus constraining the potential for population expansion (Nix and Kalma, 1972; Winter, 1997). It is possible that the ancestral population of green pythons was influenced by a combination of the above scenarios. The low genetic diversity amongst the Queensland population supports the hypothesis of a bottleneck sometime in the past.

The observed curves for the northern and southern NG populations both show a degree of raggedness (Fig. 5.4) which tend to match the expected curves for growing or declining populations. The peak on the right of the southern curve, representing signal from the divergent Normanby Island individual, may indicate that the Normanby Island population has had a distinct population growth and expansion history to the southern mainland populations. The genetic divergence of the Normanby Island sample may reflect the isolation of an allopatric population, although geological evidence suggests that Normanby Island was

joined to mainland New Guinea during the Plio-Pleistocene (Keenan, 1994). The sequence divergence of the Normanby Island sample from the mainland southern populations is greater the divergence of the Australian population from the southern population. The present study only included a single individual from Normanby Island and therefore, divergence of this population is worthy of further investigation. The raggedness index is a non-directional indicator of change in population size, meaning that specific expansion or contraction cannot be determined from the curves (except for rapid decline). To further study the population trends of green pythons it will be necessary to have much larger sample sizes from within each of the northern and southern regions.

5.4.3 Management issues

It is clear from this study that populations of green pythons should be considered as at least three distinct management units. Whilst it has not been directly tested, it is possible that lack of breeding success in captive populations may be partially due to individuals from northern and southern lineages being paired. There are already many specimens in captivity in Australia that are descended from New Guinea lineages. As there has also been interbreeding with Australian individuals, the captive population is not representative of the Australian gene pool. Therefore, in order to conform to the ZooTag recommendations to only breed local fauna and to establish the uniqueness of the captive population, the management of any captive breeding program should maintain breeding stock taken from the rainforest of Queensland.

As poaching green pythons from the wild and illegal importation of specimens from New Guinea are major law enforcement issues in Australia, the determination of distinct mitochondrial lineages from Australia, southern and northern New Guinea provides a basis for the enforcement of protective legislation. However, as sample sizes from New Guinea were small, the haplotypic diversity present in each population needs further investigation before estimates of allele frequencies can be regarded as reliable.

Appendix 5.1

Nucleotide sequences of partial *control region* (786 bp) and *cytochrome b* (292 bp) genes for 35 *Morelia viridis* and one *M. spilota* specimens.

```
                                                                                                                                           130
control region
Aus1    TTCATCCCTG GCCCCCAAAA CCAGTATTTT AAATTAAACT ACTCTTTGAA TTCAACCCTC ACTTCCTCCT AAGCCATAGT CTGTAACTTA CAGACTATGG CTCATGCCTT AATATATAAA CCAAAAACCC
Aus2    .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus16   .......... .......... .......... .......... .......... .......... .......... .......... ......m... .......... .......... .......... ..........
Aus17   .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus18   .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi1    .......... .T........ .......... .......... .......... .-........ .......... .......... ......G... .......... .......... .......... ..........
Doi2    .......... .T........ .......... .......... .......... .-........ .......... .......... ......G... .......... .......... .......... ..........
Doi3    .......... .T........ .......... .......... .......... .-........ .......... .......... ......G... .......... .......... .......... ..........
Mor1    .......... .T........ .......... .......... .......... .yT......y .......... .......... .......... .......... .......... .......... ..........
Nam1    .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .........- .......... ..........
Nam2    .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .........- .......... ..........
Nru1    .......... .T........ .......... .......... .......... .C........ .......... .......... .......... .......... .......... .......... ..........
Mai1    .......... .T........ .......... .......... .......... .C........ .......... .......... .......... .......... .......... .......... ..........
Nor1    .......... .T........ .......... .......... ........G. .......... .......T.. .......... .......... .......... .......... .......... ..........
Mer2    .......T.. .T........ .......... .......... .......... .......... .......... .......... .......... .......... .........- .......... ..........
Mer3    .......... .T........ .......... .......... .......... CC........ .......... .......... .......... .......... .......... .......... ..........
Mer4    .......... .T........ .......... .......... .......... CC........ .......... .......... .......... .......... .......... .......... ..........
Mer5    .......... .T........ .......... .......... .......... CC........ .......... .......... .......... .......... .......... .......... ..........
Mer6    .......... .T........ .......... .......... .......... CC........ .......... .......... ......G... .......... .......... .......... ..........
Aru1    .......... .T........ .......... .......... .......... CC........ .......... .......... ...G...... .......... .G........ .......... ..........
Aru2    .......... .T........ .......... .......... .......... CC........ .......... .......... .......... .......... ...w...... .......... ..........
Aru3    .......... .T........ .......... .......... .......... CC........ ........y. .......... ......G... .......... .......... .......... ..........
Aru4    .......... .T........ .......... .......... .......... CC........ .......... .......... .......... .......... .......... .......... ..........
Tim1    .......... .T........ .......... .......... .......... CC........ .......... .......... ......G... .......... .......... .......... ..........
Wau1    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Wau2    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Jay2    .......... .T........ ....A..... .......... .y.......C .-........ .......T.C .-........ ......A... ...G...... ...G.A.... G......... ..........
Bun1    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Bia1    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Bia2    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... ...G.A.... G......... ..........
Bia3    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .......... .......... ..........
Bia4    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Sor1    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G......... ..........
Sor2    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .......... .......... ..........
Sor3    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .....A.... G........- ..........
Sor4    .......... .T........ ....A..... .......... .........C .-........ .......T.C .-........ ......A... ...G...... .......... .......... ..........
spilota .......... .T........ ....A..... .......... ........GC CA..C..... .........C ..-....... ....G.A... .......... T.G....... ......-..G ......T...

                                                                                                                                             260
Aus1    ATATAAT-CA CTGCACAATA AAACTCTCTC CTCGGCCCCC CCCCTACCCC CCCCCGGAAA A-CCATAGAA GAAGTCAGCA CATAAATAAA CCTATTAATC CCATTGCTTC CTCC-TATGTA TAATCTTAC
Aus2    .......-.. .......... .....y.y.. .......... .......m.. .......... .-........ .......... .......... .......... .......... ....-...... ..........
Aus16   .......-.. .......... ...y.y.y.. y.y....... .......m.. .......... .-........ .......... .......... .......... .......... ....-...... ..........
Aus17   .......-.. .......... .......... .......... .......... ....s..... .-........ .......... .......... .......... ......y... ....-...... ..........
```

```
                                                                                                                                   260
Aus1     ATATAAT-CA CTGCACAATA AAACTCTCTC CTCGGCCCCC CCCCTACCCC CCCCCGGAAA A-CCATAGAA GAAGTCAGCA CATAAATAAA CCTATTAATC CCATTGCTTC CTCC-TATGT ATAATCTTAC
Aus18    .......-.. .......... ...y.y.y.. y......... .......m.. .......... .-........ .......... .......... .......... .......... ....-..... ..........
Doi1     .......-.. ...A...... .......... .......... .......m.. .......... .-.....r.. r..k...... .......... ....C..... .......... ....-..... ..........
Doi2     .......-.. ...A...... .....y.y.. .......... .......m.. .......... .-........ .......... .......... ....C..... .......... ....-..... ..........
Doi3     .......-.. ...A...... .......y.. .......... .......m.. .......... .-........ .......... .......... ....C..... .......... ....-..... ..........
Mor1     .......-.. ........C. ...y.y.y.. ..y....... .......m.. .....C.... .-........ .......... .......... ....C..... .......... .C..-..... ..........
Nam1     .......-.. .......... .....y.y.. y.y....... .......m.. .......... .-........ .......... .......... ..C.C..... .......... ....-..... ..........
Nam2     .......-.. .......... .....y.y.. y.y....... .......m.. .......... .-........ .......... .......... ..C.C..... .T........ ....-..... ..........
Nru1     .......-.. .......... .......y.. ..y....... .......m.. .....C.... .-........ .......... .......... ....C..... .......... ....-..... ..........
Mai1     .......-.. .......... ...y.y.y.. y.y....... .......m.. .....C.... .-........ .......... .......... ....C..... .......... ....-..... ..........
Nor1     .......-.. ........C. ...T.y.y.y y.y....... .......m.. .....CC.C. .A...C.... ..-....... T......T.. ..C.C..... .......... .CT.C..... ..........
Mer2     .......-.. .......... .....y.y.. ..y....... .......m.. .......... .-........ .......... .......... ..C.C..... .T........ ...A-..... .......w..
Mer3     .......-.. .......... .......y.. y......... .......m.. ......A... .-........ .......... .......... .......... .......... ....-..... ..........
Mer4     .......-.. .......... .......y.. .......... .......m.. ......A... .-........ .......... .......... ....C..... .......... ....-..... ..........
Mer5     .......-.. .......... .....y.y.. .......... .......m.. ......A... .-........ .......... .......... ....C..... .......... ....-..... ..........
Mer6     .......-.. .......... .....y.y.. .......... .......m.. ......A... .-........ .......... .......... ....C..... .......... ....-..... ..........
Aru1     .......-.. .......... .......... ..y....... .......m.. .....-.... .A........ .......... .......... ....C..... .......... ....-..... ..........
Aru2     .......-.. .......... .......y.. .......... .......m.. .....-.... .A........ .......... .......... ....C..... .......... ....-..... ..........
Aru3     .......-.. .......... .....y.y.. y.y....... .......m.. .....-.... .A........ .......... .......... ....C..... .......... ....-..... ..........
Aru4     .......-.. .......... .....y.y.. y......... .......m.. .....-.... .A........ .......... .......... ....C..... .......... ....-..... ..........
Tim1     .......-.. .......... .....y.y.. ..y....... .......m.. .....-.... .A........ .......... .......... ....C..... .......... ....-..... ..........
Wau1     .......T.. .CA.....C. .......... .......... .......... .......... .A-....... .......... ..ATT.A.CT TACTGAT.A. ..C....... ....C..... ..........
Wau2     .......T.. .CA.....C. .......y.. y......... .......m.. .......... .A-....... .......... ..ATT.A.CT TACTGAT.A. ..C....... ....C..... ..........
Jay2     .......T.. .CA.....C. .....y.y.. .......... .......m.. .......... .A-....... .......... ..ATT.A.CT .ACTGAT.A. ..C....... ....C..... ..........
Bun1     .......T.. .CA.....C. ...y.y.y.. ..y....... .......m.. .....-.G.. .AA....... .......... ..ATT.A.CT TACTGAT.A. ..C....... ....C..... ..........
Bia1     .......T.. .CA.....C. .......... .......... .......m.. ....-..... .AA....... .......... ..ATT.A.CT TACTGAT.A. ..C..A.... ....C..... ..........
Bia2     .......T.. .CA.....C. .....y.y.. .......... .......m.. .......... .A-....... .......... ..ATT.A.CT .ACTGAT.A. ..C....... ....C..... ..........
Bia3     .......T.. .CA.....C. .....y.y.. y......... .......m.. ....-AACG. .AA....... ........T. ..ATT.A.CT .ACTGAT... ..C....... ....C..... ..........
Bia4     .......T.. .CA.....C. .......... .......... .......m.. ....-..... .AA....... .......... ..ATT.A.CT TACTGAT.A. ..C..A.... ....C..... ..........
Sor1     .......T.. .CA.....C. .......y.. .......... .......m.. .....-.... .AA....... .......... ..ATT.A.CT TACTGAT.A. ..C..A.... ....C..... ..........
Sor2     .......T.. .CA.....C. ...y.y.y.. y.y....... .......m.. ....-AACG. .AA....... ........T. ..ATT.A.CT .ACTGAT... ..C....... ....C..... ..........
Sor3     .......T.. .CA.....C. .......... .......... .......m.. .......... .A-....... .......... ..ATT.A.CT TACTGAT.A. ..C....... ....C..... ..........
Sor4     .......T.. .CA.....C. .....y.y.. y.y....... .......m.. .....-AC.. .AA....... .......... ..ATT.A.CT .ACTGAT... ..C....... ....C..... ..........
spilota  .......TT. .CA..A..C. ..Gy.y.y.y -.y....... .......... .....-A.GG .A-...T.GG ...A..G... ......ACC. TTAGAA..CT .TCAACAAA. y..TC..... ..........

                                                                                                                                   390
Aus1     ATTAATGGTT TGCCCCATGA ATATTAAGCA GGAATTTCCC TTCAAATATT TTAGCCTAAA TTAGCTTCTG TACAAAATAT CTAGTCCTCA TTTTCT-GGT CGTTCAATGC AATCGGGGTT AATAAATCTT
Aus2     .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Aus16    .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Aus17    .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Aus18    .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Doi1     .......... .......... .......... .......... .......... .......... ........C. .......... ....C..... ......-... .......... .......... ..........
Doi2     .......... .......... .......... .......... .......... .......... ........C. .......... ....C..... ......-... .......... .......... ..........
Doi3     .......... .......... .......... .......... .......... .......... ........C. .......... ....C..... ......-... .......... .......... ..........
Mor1     .......... .......... .....T.... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Nam1     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Nam2     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Nru1     .......... .......... .......... .......... .......... .......... ........C. .......... ....C..... ......-... .......... .......... ..........
Mai1     .......... .......... .......... .......... .......... .......... ........C. .......... ....C..... ......-... .......... .......... ..........
```

```
                                                                                                                                        390
Aus1     ATTAATGGTT TGCCCCATGA ATATTAAGCA GGAATTTCCC TTCAAATATT TTAGCCTAAA TTAGCTTCTG TACAAAATAT CTAGTCCTCA TTTTCT-GGT CGTTCAATGC AATCGGGGTT AATAAATCTT
Nor1     .......... .......... .......... .......... .......... .......... .......... .......... .C..C..... ......-... .......... ....A..... ..........
Mer2     .......... .......... .......... .......... .......... .......... ........S. .......... .......... ......-... .......... .......... ..........
Mer3     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Mer4     .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Mer5     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Mer6     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Aru1     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Aru2     .......... .......... .......... .......... .......... .......... .......... .......... .......... ...C..-... .......... .......... ..........
Aru3     .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Aru4     .......... .......... .......... .......... .......... .......... .......... .......... .......... ......-... .......... .......... ..........
Tim1     .......... .......... .......... .......... .......... .......... ........C. .......... .......... ......-... .......... .......... ..........
Wau1     .......... .......... .......... .......... .......... .......... A....C.T.. .........C ..T....... ......-... .......... .......TC. ..C.......
Wau2     .......... .......... .......... .......... .......... .......... A....C.T.. .........C ..T....... ......-... .......... .......TC. ..C.......
Jay2     .......... .......... .......... .......... .......... .......... A....C.T.. ....G....C ..T....... ......-... .......... .......TC. ..........
Bun1     .......... .......... .......... .......... .......... .......... A....C.T.. .........C ..T....... ......-... .......... .......TC. ..........
Bia1     .......... .......... .......... .......... .......... ....T..... ...A.C.T.. .........C ..T....... ......-... .......... .......TC. ..........
Bia2     .......... .......... .......... .......... .......... .......... A....C.T.. ....G....C ..T....... ......-... .......... .......TC. ..........
Bia3     .......... .......... .......... .......... .......... ....T..... A..A.C.T.. .........C .CT....... ......-... .......... .......TC. ..........
Bia4     .......... .......... .......... .......... .......... ....T..... ...A.C.T.. .........C ..T....... ......-... .......... .......TC. ..........
Sor1     .......... .......... .......... .......... .......... ....T..... ...A.C.T.. .........C ..T....... ......-... .......... .......TC. ..........
Sor2     .......... .......... .......... .......... .......... ....T..... A..A.C.T.. .........C .CT....... ......-... .......... .......TC. ..........
Sor3     .......... .......... .......... .......... .......... ...A...... A....C.T.. .........C ..T....... ......-... .......... .......TC. ..........
Sor4     .......... .......... .......... .......... ..T....... ....T..... A..A.C.T.. .........C ..T....... ......-... .......... .......TC. ..........
spilota  .......... ....T..... .....T.... .......... ..TT.T.... ....T..... A.G..C.T.. .......... TCT....... .A....T... .....T.... .GCAC.A... ..CT......

                                                                                                                                        520
Aus1     ACTAACCATG GATATCCTTG ATCAAGTGGT GTCCCTTAAT TTAGTA-CTT CCC-GTGAAA TCTTCTATCC TTCCATAGAA TGCTAACCAT TCGACTTCTC ACGTCCATAT TATGCCAATC CTCCCTTCTA
Aus2     .......... .......... .......... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... .......... ..........
Aus16    .......... .......... .....T.... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... .......... ..........
Aus17    .......... .......... .......... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... .......... ..........
Aus18    .......... .......... .......... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... .......... ..........
Doi1     .......... .......... ....G..... ...T...... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Doi2     .......... .......... ....G..... ...T...... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Doi3     .......... .......... ....G..... ...T...... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mor1     .......... .......... ....G..... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Nam1     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Nam2     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Nru1     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mai1     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Nor1     .A........ .......... .......... ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ......C...
Mer2     .......... .......... ....G..... .......... C.....-... ..-G...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mer3     .......... .......... ....G..... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mer4     .......... .......... ....G..... ...T...... C..T..-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mer5     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Mer6     .......... .......... ....G..... .......... ......-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Aru1     .......... ......G... ....G..... .......... C.....-... ...-...... ..C......S .......... .......... .......... .......... ....T..... ..........
Aru2     .......... .......... ....G..... ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Aru3     .......... .......... ....G..... ...T...... C..T..-... ..--.GT... ..C....... .......... .......... .......... .......... ....T..... ..........
```

```
                                                                                                                                     520
Aus1     ACTAACCATG GATATCCTTG ATCAAGTGGT GTCCCTTAAT TTAGTA-CTT CCC-GTGAAA TCTTCTATCC TTCCATAGAA TGCTAACCAT TCGACTTCTC ACGTCCATAT TATGCCAATC CTCCCTTCTA
Aru4     .......... .......... ....G..... ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... ....T..... ..........
Tim1     .......... .......... ....G..... .......... C.....-... ...-...... ..C....... .......... .......... .......... .......... ...TT..... ..........
Wau1     .......... .......... .......T.. ...T...... C.....A... ...C...... ..C....... .......... .......... ....-..... .......... A......... .......TA.
Wau2     .......... .......... .......T.. ...T...... m.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Jay2     .......... .......... .......T.. ...T...... C..T..-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Bun1     .......... .......... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Bia1     .......... .......... .......T.. ...T....G. C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Bia2     .......... ......G... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Bia3     .......... .......... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Bia4     .......... .......... .......T.. ...T....G. C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Sor1     .......... .......... .......T.. ...T....G. C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Sor2     .......... .......... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Sor3     .......... .......... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
Sor4     .......... .......... .......T.. ...T...... C.....-... ...-...... ..C....... .......... .......... .......... .......... A......... .......TA.
spilota  .T........ ....T....A .C.T.AG... ...T....G. C...CG-... ...-...... ..C....... ....G..T.. .......... .......... .......... C....T.... ...TT.A..T

                                                                                                                                     650
Aus1     GCTTTTTCCA AGGCCGCTGG TTACACTCTC AAGATCATCT CAATGGTCCG GAACCACCCC TCCATCCTAG CTTTTTCCAA GACCTTTGGT CGCACCCTTT ATATGGTACA TCTTGTCTCA TGTTCTGATC
Aus2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus16    .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus17    .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus18    .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi3     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mor1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nam1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nam2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nru1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mai1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nor1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer3     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer4     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer5     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer6     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru2     .......... .......... .......... .......... .......... .......... .......... .......... A......... .......... .......... .......... ..........
Aru3     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru4     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Tim1     .......... .......... .......... ..A....... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Wau1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Wau2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Jay2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bun1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Bia1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .T........ ..........
Bia2     .......... .......... .......... ........?. .......... .......... ......TAG. .......... .......... .......... .......... .T........ ..........
Bia3     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .T........ ..........
```

```
                                                                                                                                          650
Aus1     GCTTTTTCCA AGGCCGCTGG TTACACTCTC AAGATCATCT CAATGGTCCG GAACCACCCC TCCATCCTAG CTTTTTCCAA GACCTTTGGT CGCACCCTTT ATATGGTACA TCTTGTCTCA TGTTCTGATC
Bia4     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .T........ ..........
Sor1     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Sor2     .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .T........ ..........
Sor3     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Sor4     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .T........ ..........
spilota  ...C...... .......... .......... .......... .G........ .......... ....A..... .......... .......... ......C... .......... .A.CAC.... ..........

                                                                                                                                          780
Aus1     ACCTATGCTA GTCCACCCCT GGTT-CCCTT TTTTTCTCGG TACCTTTCAT CTGACTACCA TATATGCACA CACACAGTAA GGCTCTTTAG TCCGGGTGGA ACATGATGTT ACTTCCTCAT TTACCCCAAT
Aus2     .......... .......... ....-..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus16    .......... .......... ....-..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus17    .......... .......... ....-..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus18    .......... .......... ....-..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .-........ ....G.....
Doi2     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi3     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mor1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... ......C... .......... .......... .......... ..........
Nam1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nam2     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nru1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mai1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nor1     .......... .......... ....T..... ........T. .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer2     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer3     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer4     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer5     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer6     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru2     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru3     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru4     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Tim1     .......... .......A.. ....T..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Wau1     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-G. .......... ....T..... TTC...C... .A....m...
Wau2     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-G. .......... ....T..... TTC...C... .......C..
Jay2     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... .......C..
Bun1     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-G. .......... ....T..... TTC...C... ..........
Bia1     .......T.. ....G..A.. ....T..... ........T. .........C .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... .......C..
Bia2     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... .......C..
Bia3     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... ..........
Bia4     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... ......-...
Sor1     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... ..........
Sor2     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... ..........
Sor3     .......T.. ....G..A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTG...C... .......C..
Sor4     ..G....T.. .......A.. ....T..... ........T. .......... .......... .......... .......... ....TAC-.. .......... ....T..... TTC...C... ..........
spilota  .T.....TCT AC.....AT. ....G.T..C .....-..T. .......... .....C.... .......... ........T. ...A-..... .......T.. G..C.C..C. T.C....... .......T.G
```

150

cytochrome b

910

```
Aus1     GGATAC TTCG GCTCAATACT ACTAACATGC CTAGCCCTAC AAGTATTAAC CGGCTTCTTC CTAGCCGTTC ACTACACAGC AAACATTAAT CTAGCATTCT CATCCATCAT CCACATCTCC CGAGATGTTC
Aus2     ...... .... .......... .......... .......... ..A....... .......... .......... .........s .......... .......... .......... .......... ..........
Aus16    ...... .... .......... .......... .......... .......... .......... .......... .........s .......... .......... .......... .......... ..........
Aus17    ...... .... .......... C......... .......... .......... .......... .......... .........s .......... .......... .......... .......... ..........
Aus18    ...... ???? ???....... .......... .......... .......... .......... .......... .........G .......... .......... .......... .......... ..........
Doi1     ...... .... .y........ .T........ .......... .......... .......... .......... .........s .......... .......... .......... .......... ..........
Doi2     ...... .... .......... .T........ .......... .......... .......... .......... .........s .......... .......... .......... .......... ..........
Doi3     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mor1     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... ........T. .......... .......... ..........
Nam1     ...... .... .......... .T........ .......... .......... .......... .......... .........s .........C .......... .......... .......... ..........
Nam2     ..... .... .......... .T........ .......... .......... .......... .......... .........s .........C .......... .......... .......... ..........
Nru1     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mai1     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Nor1     .....T .... .......... .T........ .......... .......... .......... T.......A. .........s .......G.. .......... .......... .......... .....C....
Mer2     ...... .... .......... .T........ .......... .......... .......... .......... .......... .........C .......... .......... .......... ..........
Mer3     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer4     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer5     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Mer6     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru1     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru2     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru3     ...... .yy. .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aru4     ...... .... .s........ .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Tim1     ...... .... .......... .T........ .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Wau1     ...... .... ........T. .T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Wau2     ...... .... ........T. .T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Jay2     ...... .... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Bun1     ...... .... .y......T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Bia1     ...... .... ........T.. T......... T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. .......A.T ........C.
Bia2     ...... .... ........T. .T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Bia3     ....... ... ........T.. T.......T T......... .....C.... .......... T.......C. ....T..... ......C..C .......... .......T.. T......A.T ........C.
Bia4     ...... .... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. y......A.T ........C.
Sor1     ...... .... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. .......A.T ........C.
Sor2     ....... ... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. y......A.T ........C.
Sor3     ...... .... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. ...T...A.T ........C.
Sor4     ....... ... ........T.. T.......T T......... .....C.... .......... T.......C. .......... ......C..C .......... .......T.. .......A.T ........C.
spilota  ....... ... .......... .T........ ....T..... ....CC.... .......... T....T..C. .......... .........C ..G....... .......... ......TA.. .....C..C.
```

1040

```
Aus1     CATACGGTTG AATAATACAA AACCTACACG CCATCGGAGC ATCCATATTC TTCATTTGCA TCTACATCCA TATTGCACGA GGGTTATACT ATGGATCCTA CCTCAACAAA GAAACCTGAA TATCCGGTAT
Aus2     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus16    .......... .......... .......... .......... .......... .......... .......... .......... .....y.... .......... .......... .......... ..........
Aus17    .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Aus18    .......... .w........ .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Doi1     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... ..........
Doi2     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... ..........
Doi3     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... ..........
Mor1     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..AC..... .C........ .......... ..........
Nam1     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... ..........
```

```
                                                                                                                                     1040
Aus1    CATACGGTTG AATAATACAA AACCTACACG CCATCGGAGC ATCCATATTC TTCATTTGCA TCTACATCCA TATTGCACGA GGGTTATACT ATGGATCCTA CCTCAACAAA GAAACCTGAA TATCCGGTAT
Nam2    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ..........
Nru1    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ..........
Mai1    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ..........
Nor1    .......... .......... .....G.... .......... .......... .....C.... .......... .......... ..AC...... .......... .......... .......... ........G.
Mer2    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ..........
Mer3    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ..........
Mer4    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ........G.
Mer5    .......... .......... .......... .......... .......... .......... .T........ .......... ..A....... .......... .......... .......... ..........
Mer6    .......... .......... .......... .......... .......... .......... .T........ .......... ..A....... .......... .......... .......... ..........
Aru1    .......... .......... .......... .......... .......... .......... .T........ .......... ..A....... .......... .......... .......... ..........
Aru2    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ........G.
Aru3    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ........G.
Aru4    .......... .......... .......... .......... .......... .......... .......... .......... ..A....... .......... .......... .......... ........G.
Tim1    .......... .......... .......... .......... .......... .......... .T........ .......... ..A....... .......... .......... .......... ..........
Wau1    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Wau2    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Jay2    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Bun1    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Bia1    .......C.. .......... .......... .......... .......... .......... .......... C..C...... ..A....... .C........ ...A...... ..G..T.... ..........
Bia2    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Bia3    .......C.. .......... .......... .......... .......... .......... .......... C......... ..AC...... .C.....T.. .......... ..G..T.... ..........
Bia4    .......C.. .......... .......... .......... .......... .......... .......... C..y...... ..Ay...... .C.....y.. .......... ..G..T.... ..........
Sor1    .G.....C.. .......... .......... .......... .......... .......... .......... C..C...... ..A....... .C........ ...A...... ..G..T.... ..........
Sor2    .......C.. .......... .......... .......... .......... .......... .......... C..y...... ..Ay...... .C.....y.. .......... ..G..T.... ..........
Sor3    .......C.. .......... .......... .......... .......... .......... .T........ C..C...... ..AC...... .C........ .......... ..G..T.... ..........
Sor4    .......C.. .......... .......... .......... .......... .......... .T........ C......... ..AC...... .C........ .......... ..G..T.... ..........
spilota .......C.. .......... .......... .......... ...T..G... .......... .......... .........C ..A....... .C........ T......... .......... .......C..

                                          1078
Aus1    TACCCTACTC ATCACACTAA TAGCAACCGC CTTCTTCG
Aus2    .......... .......... .r???????? ????????
Aus16   .......... .......... .........y ........
Aus17   .......... .......... .G........ ..y.....
Aus18   .......... .......... ......???? ????????
Doi1    .......... .......... .......... ........
Doi2    .......... .......... .......... ........
Doi3    .......... .......... .......... ........
Mor1    .......... .......... .......... ........
Nam1    .......... .......... .......... ........
Nam2    .......... .......... ...?.....T ........
Nru1    .......... .......... .......... ........
Mai1    .......... .......... .......... ........
Nor1    ......G... .......... .......... ........
Mer2    .......... .......... .......... ........
Mer3    ......T... .......... .......... ........
Mer4    ......T... .......... .......... ........
Mer5    ......T... .......... .......... ........
Mer6    ......T... .......... .......... ........
```

```
                                                      1078
Aus1     TACCCTACTC ATCACACTAA TAGCAACCGC CTTCTTCG
Aru1     ......T... .......... .......... ........
Aru2     ......T... .......... .......... ........
Aru3     ......T... .......... .......... ........
Aru4     ......T... .......... .......... ........
Tim1     ......T... .......... .......... ........
Wau1     C......... ......T... .G........ ......?.
Wau2     C......... ......T... .G........ ......T.
Jay2     C......... ......T... .......... ......T.
Bun1     C......... ......T... .......... .....???
Bia1     .......... ......T... .......... ......T.
Bia2     C......... ......T... .......... ......T.
Bia3     C......... ......T... .......... ......T.
Bia4     y......... ......T... .......... ......T.
Sor1     .......... ......T... .......... ......T.
Sor2     y......... ......T... .......... ......T.
Sor3     C......... ..T...T... .......... ......T.
Sor4     C......... ......T... .......... ......T.
spilota  .......... .......... .......... T.......
```

Chapter 6.

Phylogenetic relationships of *Liasis* pythons based on mitochondrial DNA.

6.1 Introduction

Since their description in the mid-1800's, pythons of the genus *Liasis* have had a dynamic taxonomic history (Kluge, 1993; McDowell, 1975; Underwood and Stimson, 1990). McDowell (1975) divided the many species of *Liasis* into two groups. One group, called the *L. olivaceus* Group, comprised the olive python, *L. olivaceus,* the Papuan python *L. papuana* and the water pythons *L. fuscus/L. mackloti.*

The olive python is endemic to Australia and is found in two disjunct geographic regions (Fig. 6.1). *L. olivaceus olivaceus* Gray 1842 is distributed from the Kimberley district of Western Australia to the western slopes of the Great Dividing Range in Queensland. *L. o. barroni* Smith 1981 is found only in the Pilbara region of Western Australia. The olive python is unpatterned, reddish-brown in colour and reaches a length of 4-6 metres (Shine, 1991). The two subspecies are distinguished by differences in mid-body and ventral scale counts (Smith, 1981a).

The Papuan python, whose distribution is limited to New Guinea, is comparable in size and colour to *L. olivaceus* and has a similarly high number of ventral scales. The two taxa were considered to be conspecific by Stuhl (1932). In more recent studies using morphological characters to determine the evolutionary relationships of pythons, *Liasis papuana* was placed as a sister taxon to *Liasis olivaceus /L. fuscus/L. mackloti* by McDowell (1975) and to *L. olivaceus* only by Underwood & Stimson (1990). Kluge (1993) considered it to be sufficiently different to warrant erection of a new monotypic genus, *Apodora.*

Water pythons are found in the Lesser Sunda islands of Indonesia, the trans-Fly River

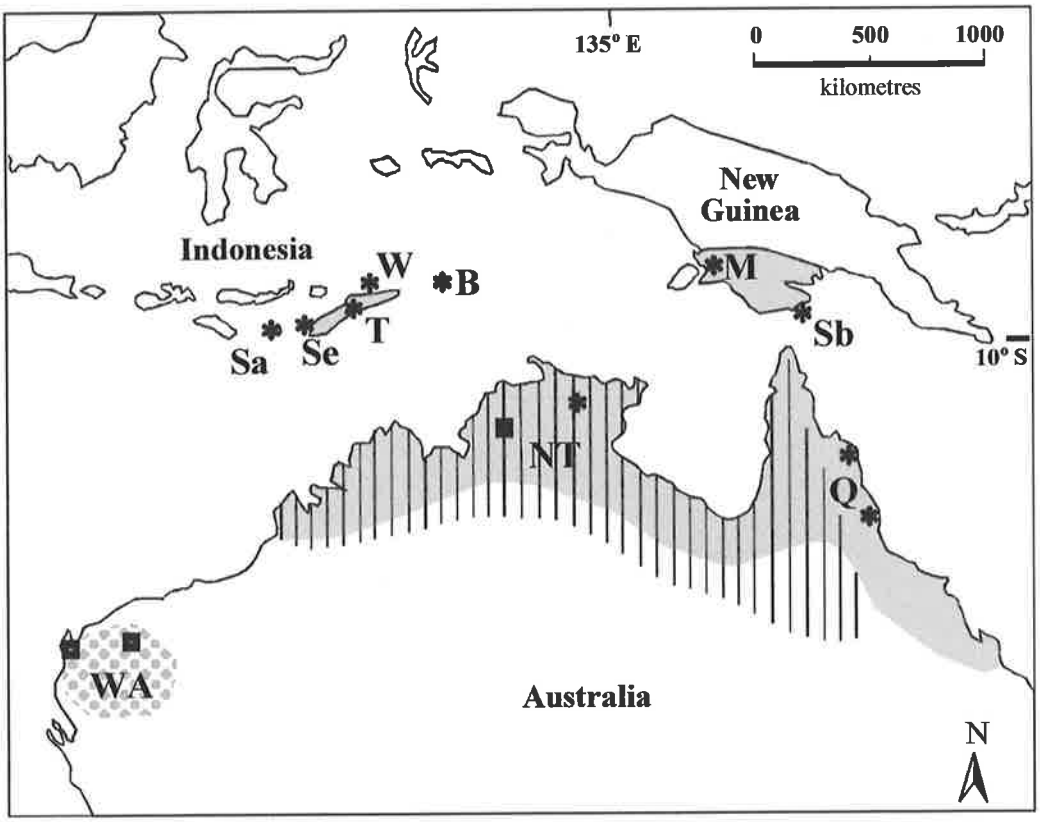


Fig. 6.1. Map showing populations of *Liasis mackloti* (✱) and *L. olivaceus* (■) sampled. Letters indicate populations as follows: **B**= Babar Island, **M** = Merauke, **NT** = Northern Territory, **Q** = Queensland, **Sa** = Sawu Island, **Sb** = Saibai Island, **Se** = Semau Island, **T** = Timor, **W** = Wetar, **WA** = Western Australia. Shaded area indicates the natural range of *L. fuscus/mackloti*, hatched area indicates the natural range of *L. olivaceus olivaceus* and the spotted area represents the range of *L. o. barroni*.

region of southern New Guinea and across northern Australia from the Kimberley district of Western Australia, through the tropical region of Northern Territory and Queensland as far south as Bowen (Fig. 6.1) (McDowell, 1975). They are also found on some islands in Torres Strait. Water pythons are predominantly shades of brown to glossy black dorsally with some patterning on the Indonesian specimens (Brongersma, 1956) and no patterning on the Australian and New Guinean specimens (Barker and Barker, 1994; O'Shea, 1996). They have an average length of 1.5-2.0 metres (Cogger, 1996; Shine and Slip, 1990) and are usually found in wetter areas such as freshwater swamps and lagoons.

Water pythons have been recognised as one or two species by different reviewers. *Liasis mackloti* Duméril & Bibron 1844, was originally described from Timor. The Australian water python *L. fuscus* Peters 1874, was described from Port Clinton in Queensland and this name was also applied to specimens from New Guinea (e.g. de Rooij, 1917). Boulenger (1893) presented a key to separate *L. fuscus* and *L. mackloti* based on the presence/absence of a groove or pit in the rostral scale and ventral and sub-caudal scale counts. This key was determined from published records of specimens from Timor and Semau and observations of specimens from Queensland and New Guinea available in the Natural History Museum, London (McDowell, 1975). McDowell (1975) used Boulenger's key to examine the status of *L. fuscus/L. mackloti* by comparing two specimens from Wetar and four from New Guinea, one from Northern Territory and three from Queensland. He could not differentiate populations according to Boulenger's key, which may be due to the limited samples available from Indonesia. McDowell could distinguish specimens from Australia-New Guinea and Wetar Island based on colour variation, but attributed the difference to geographical variation within a species.

Water pythons from Wetar and Sawu Island had been classified previously as the sub-species *L. m. dunni* by Stuhl (1932) and *L .m. savuensis* by Brongersma (1956) respectively. Barker & Barker (1994) recognised *L. fuscus* from Australia and New Guinea and *L. mackloti* from Indonesia. Kluge (1993) conservatively recognised only a single species, *L. mackloti*

but indicated the need for further investigation of geographic variation.

In summary, the taxonomic history of the *Liasis* Group suggests that there are two distinct types of olive python - *L. olivaceus barroni* and *L. o. olivaceus,* four distinct types of water python - *L. fuscus* (from Australia and New Guinea), *L. mackloti*, *L. m. savuensis* and *L. m. dunni* (from Indonesia) and the Papuan python, *Apodora papuana* as the sister lineage to all *Liasis* species.

To date there has been no comprehensive molecular study of the systematics of McDowell's *L. olivaceus* Group. Because Kluge (1993) placed *A. papuana* as the sister lineage to *L. fuscus/L. mackloti* and *L. olivaceus,* I included *A. papuana* as an ingroup taxon in order to examine the relationships of the genus *Liasis* with respect to *Apodora*. To examine the contentious taxonomy of the group, for completeness, I included representatives of the two olive python subspecies and specimens of water pythons from across the range in Australia, New Guinea and Indonesia. Because of a relatively rapid rate of nucleotide evolution, mitochondrial DNA markers are powerful tools in evolutionary biology and have been used widely to elucidate phylogeographic structure and species boundaries (Moritz, 1994; Moritz *et al.*, 1987). However, because I recovered the phylogenetic history of a mitochondrial gene, *i.e.* a "gene" tree, which may or may not be representative of the "species" tree (Avise, 1994), I do not attempt to use the results of the present study to alter taxonomic classifications. Rather, any phylogenetic structure that may be present in the mitochondrial gene tree is used to provide a focus for further investigations and biogeographic speculation.

6.2 Materials and Methods

6.2.1 Specimens examined

Samples were available for thirty-five specimens of *Liasis fuscus/L. mackloti* from five Indonesian islands, New Guinea and Australia (Northern Territory and Queensland, including Saibai Island). Three specimens of *L. olivaceus barroni*, four specimens of *L. o.*

olivaceus, and single specimens of *Apodora papuana* and the outgroup taxa *Morelia spilota* and *M. viridis*. Individuals included in this study are given below.

Liasis fuscus/L. mackloti*: Indonesia** Babar Island, **wB 1** ABTC 66396; Sawu **wSa 1-5** ABTC 67166, ABTC 68266-69; Semau **wSe 1-3** ABTC 66392-3, ABTC 67165; Timor **wT 1-4** ABTC 66390-1, ABTC 66394-5; Wetar **wW 1-2** ABTC 67167/86; **New Guinea** Merauke **wM 1-2** ABTC 68258-9; **Australia Northern Territory**, Arnhem Highway **wNT 1** ABTC 13467, Darwin **wNT 2-4** ABTC 68260, SAMA R21422, SAMA R27497, Fogg Dam, **wNT 5-9** ABTC 13445, ABTC 13460-2, NTM R10694, Humpty Doo **wNT 10** ABTC 68261; **Queensland wQ 1-6** Cairns ABTC 68262, ABTC 51982-3, Tully Gorge ABTC 66026, ABTC 51981, Yorkey's Knob ABTC 68263, **Saibai Island wSb 1-2** ABTC 67637/9. ***Liasis olivaceus*:** Northern Territory **oNT 1-4** ABTC 6503, ABTC 13427, ABTC 13463, ABTC 66237; Western Australia **oWA 1-3** ABTC 68265, ABTC 66137, ABTC 66214. ***Apodora papuana ABTC 68240, ***Morelia spilota*** SAMA R26878 and ***M. viridis*** AMS R115348.

6.2.2 Mitochondrial DNA

DNA was extracted as per chapter 2.2 and mitochondrial DNA (mtDNA) was enriched from frozen liver for *L. mackloti* SAMA R21422 as per chapter 2.3.2. Nested PCR as per chapter 2.8 was used to amplify a partial *transfer RNA^{Thr}/control region* (*CR*) product using first round primers L15926 and H690 and second round primers L15926 and snake17. *CR* products were amplified using PCR2 (chapter 2.7.2) and program 2 (chapter 2.7.3.2). Both strands of the PCR products were sequenced with the primers used for PCR or with nested primers Snake 1, Snake 6 and Snake 7 as per chapter 2.10 and 2.7.2.3. Sequences are included as Appendix 6.1.

The potential for the *CR* primers to amplify mitochondrial genes rather than nuclear paralogues was tested using the protocol of chapter 2.4 on enriched mtDNA and total cellular DNA of *Liasis mackloti* SAMA R21422.

Sequences were aligned by eye and with only six insertion/deletion (indel) events

inferred within *L. fuscus/L. mackloti* and 12 insertion/deletion events between *L. fuscus/L. mackloti* and *L. olivaceus*, it was not necessary to use a secondary structure model to optimise the alignment. Aligned sequences were phylogenetically analysed with PAUP*4.0b2a (written by D. Swofford, 1999). The model used for maximum likelihood analyses, selected using Modeltest (Posada and Crandall, 1998), was the HKY85+Γ model (Hasegawa *et al.*, 1985). The HKY85+Γ evolutionary model was also used in the Neighbor-Joining (NJ) analysis. The robustness of the trees was evaluated by bootstrap analysis using 2000 fast heuristic pseudoreplicates for MP and NJ analyses and 200 for ML analysis. Support for the internal nodes of branches of MP trees was determined from decay indices (Bremer, 1994) calculated with the program Autodecay ver. 2.9.9 (written by T. Eriksson, 1997). The Wilcoxon signed-ranks test was applied to examine statistical significance of the most parsimonious tree relative to alternative hypotheses (Templeton, 1983). To find the most parsimonious tree(s) under the alternate hypothesis, a heuristic search was performed under the constraint of the hypothesis to be tested. All the most parsimonious trees generated under the hypothesis constraint were tested against all the most parsimonious trees generated under no constraint using the Templeton (1983) test.

6.3 Results

Partial *CR* sequences from enriched mitochondrial DNA and from total cellular DNA for *Liasis mackloti* SAMA R21422 were indistinguishable, providing no evidence that the primers amplify paralogous sequences.

For most of the samples tested, there are 795 aligned nucleotide sites (48 bp of $tRNA^{Pro}$ and 747 bp of *CR*), 208 sites were variable, with 119 parsimony informative. Due to sequencing difficulties, full length sequence could not be obtained for three samples. For samples wQ2, wNT9 and wNT10 there are 617, 671 and 684 bp respectively. The ts/tv ratio as estimated by ML was 2.56 and the gamma shape parameter was 0.241410. The nucleotide sequences had estimated nucleotide frequencies of A=0.282054, C=0.281717, G=0.122758

and T=0.313472. Amongst 7 individuals of *L. olivaceus* and 35 individuals of *L. fuscus/L. mackloti* sequenced there were 5 and 18 haplotypes respectively. In olive and water pythons, there were two and seven haplotypes respectively shared by more than one individual. Shared haplotypes were: (oNT3,4) and (oWA1,3) for olive pythons and (wM1,2, wSB1,2), (wNT2,6), (wNT5,8), (wQ1, 4-6), (wSa1-4), (wSe1,2) and (wSe3, wT1-4) for water pythons.

For all the *Liasis* sequences, there was a mean HKY85 corrected haplotype divergence of 8% between *L. olivaceus* and *L. fuscus/L. mackloti*. There were 46 synapomorphies for the *L. olivaceus* haplotypes, including 5 indels. One of these indels comprised four nucleotides in *L. o. olivaceus* and five nucleotides in *L. o. barroni* that were not present in the *L. fuscus/L. mackloti* haplotypes. Within the *L. olivaceus* clade there were 13 synapomorphies for *L. o. barroni* and five for *L. o. olivaceus*.

It is important to note that informative sites coded as alignment gaps could not be included in the maximum likelihood analyses, which is reflected in lower ML bootstrap support for most clades. A ts/tv ratio of 2.5 was used to down-weight transitions relative to transversions for weighted parsimony analysis. From both unweighted and weighted heuristic searches under the maximum parsimony criterion of optimality found 30 equally most parsimonious trees (unweighted tree length of 441steps). A maximum likelihood tree generated under the HKY85+Γ model is presented in Fig. 6.2. Maximum likelihood and distance analyses had similar tree topologies to the parsimony analyses with the exception that the distance tree had *Apodora* as a sister taxon to *M. spilota*. Bootstrap proportions for ML, MP and NJ analyses are indicated on Fig. 6.2. *Apodora papuana* was the sister taxon to *Liasis* and there was reciprocal monophyly of *L. fuscus/L. mackloti* and *L. olivaceus* with strong bootstrap support (73-100%). Within *L. olivaceus* there was reciprocal monophyly of *L. o. barroni* and *L. o. olivaceus,* although there was not strong bootstrap support from distance or likelihood analyses. However, a decay index of 14 supported the branch uniting the two subspecies.

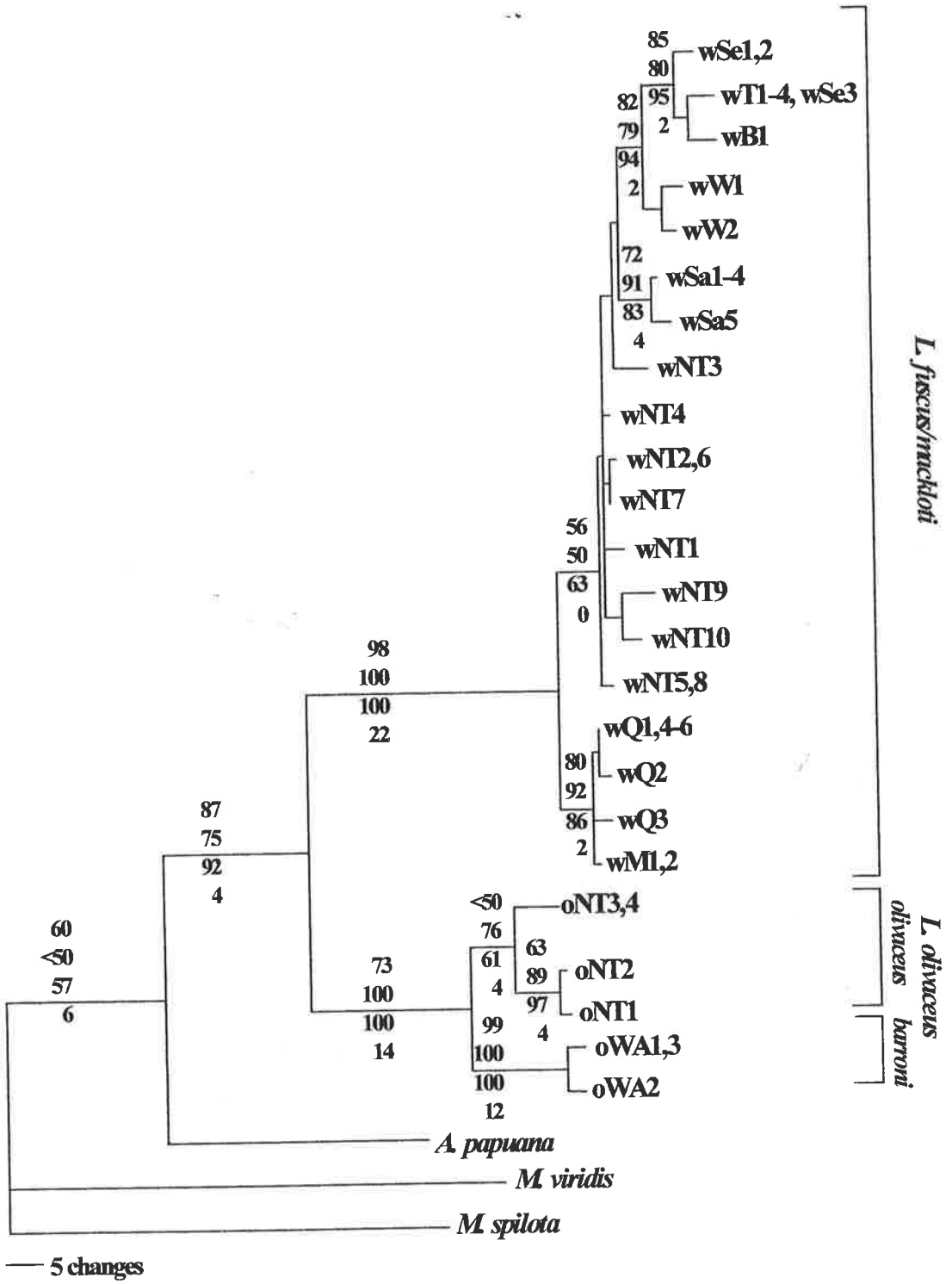


Fig. 6.2. Phylogram of maximum likelihood tree showing the genetic relationships of olive and water pythons for partial sequence of the mitochondrial *control region*. Numbers in descending order represent bootstrap proportions of 200 and 2000 pseudoreplicates for maximum likelihood, maximum parsimony and distance analyses respectively and decay indices.

The relationships between haplotypes of populations within the water python clade (*L. fuscus/L. mackloti*) were not fully resolved by any analysis, although there was well-supported monophyly (>72% bootstrap proportions) for most populations, except NT and Qld. There was strong support (>80% from bootstrap analysis; decay index of 2) for the monophyly of the New Guinea/Queensland/Saibai populations (hereafter called the "eastern" clade) with the HKY85 NJ tree and strict consensus trees of the MP analyses showing this clade as the sister lineage to the Northern Territory/Indonesian clade (hereafter called the "western" clade). There were five nucleotide substitutions and an indel unique to the "eastern" clade. There was also strong support (>72% bootstrap pseudoreplicates) for monophyly of each of the Sawu, the Babar/Semau/Timor and Wetar clades. There were four transitions unique to the Babar/Semau/Timor cluster and two transitions and two indels defining the Sawu cluster. There were two unique haplotypes from Wetar that had two synapomorphies separating them from the remainder of the *Liasis* populations. Within the Babar/Semau/Timor clade, four Timor individuals and one Semau individual (wSe3) shared a haplotype, which differed from a Babar haplotype (wB1) by a single nucleotide. However, there were nine nucleotides difference between these haplotypes and the other Semau haplotype (wSe1,2).

There was no structure amongst the Northern Territory haplotypes, except in the HKY85 analysis where all individuals but one (wNT3) formed a clade (58% bootstrap pseudoreplicates). In 11 of the 30 equally most-parsimonious trees and the ML tree, this Northern Territory haplotype was more closely associated with the Sawu haplotypes than with the remaining Northern Territory haplotypes.

Within the "western" clade, intra-clade HKY85-corrected divergence was 0.0-0.7% and inter-clade divergence was approximately 2.5%. Within the Babar/Semau/Timor clade, inter-island divergence was approximately 1.2%. The divergence between the Semau haplotypes, wSe1,2 and wSe3, was 1.13%, whereas there was 0.6% divergence between the two haplotypes,wSa5 and wSa1-4, observed from Sawu.

Templeton's (1983) test was applied to the data set to examine the statistical significance of the shortest trees in MP analysis relative to two alternate hypotheses. The alternate hypotheses tested were a) monophyly of the Australian haplotypes, *i.e.* no distinction between the "eastern" and "western" clades and b) monophyly of the Northern Territory haplotypes.

For hypothesis a) the constraint tree was (*L. fuscus/L. mackloti* from Semau, Timor, Babar, Sawu, Wetar (Queensland, Merauke, Saibai, Northern Territory), *L. olivaceus*, *A. papuana*, *M. viridis*, *M. spilota*). There were 55 equally most-parsimonious trees generated in an unweighted heuristic search under this constraint. The constraint tree for hypothesis b) was (*L. fuscus/L.mackloti* from Semau, Timor, Babar, Sawu, Wetar, Queensland, Merauke, Saibai,(Northern Territory), *L. olivaceus*, *A. papuana*, *M. viridis*, *M. spilota*). There were 148 equally most-parsimonious trees generated in the unweighted heuristic search under this constraint. Each of the 30 equally most-parsimonious trees generated in a heuristic search with no constraints was tested against all the most-parsimonious constraint trees. None of the trees from either alternate hypothesis was significantly different from the overall shortest trees (Table 6.1). Therefore, the hypotheses of Australian or Northern Territory monophyly could not be rejected.

6.4 Discussion

6.4.1 Taxonomic implications

The phylogenetic analyses in the present study provide genetic support for the taxonomic groupings summarised in the introduction. *Apodora papuana* was the sister lineage to *L. olivaceus* and *L. fuscus/L. mackloti*, supporting the separate status of the genus *Liasis*. As well as very strong support for the monophyly of *CR* haplotypes from the island populations of Sawu (*L. m. savuensis*) and Wetar (*L. m. dunni*), there was also strong support for monophyly of haplotypes from the island cluster of Timor, Semau and Babar. Reciprocal

Table 6.1. Results from Wilcoxon signed–ranks tests for monophyly of the Australian and Northern Territory *Liasis fuscus/L. mackloti*. The number of equally most parsimonious trees is shown in brackets after each hypothesis. None of the tests were significant.

Alternative hypothesis	No. sites compared	Test score	P-value
Australian monophyly (55)	8	13.5	0.4795
Tree length = 443	10	22.0	0.5271
	12	32.5	0.5637
Northern Territory monophyly (148)	4	2.5	0.3173
Tree length = 443	5	5.0	0.4795
	6	7.0	0.4142
	8	13.5	0.4795
	10	22.0	0.5271

monophyly of Timor and Semau was not supported in the present analysis and both populations share a haplotype (wSe3). However, the amount of divergence between the two haplotypes from Semau (wSe1,2 and wSe3) was 60% greater than the maximum intra-clade divergence for other islands, suggesting either that there has been dispersal, naturally or in association with human movement around the region, or that the samples were mislabeled. Timor and Semau were last connected by dry land during the last glacial maximum 18000 yr BP (Heaney, 1991; Kitchener and Suyanto, 1996) providing a means of easy dispersal. These alternatives can only be distinguished with further sampling.

The genetic relationships of the Indonesian and Northern Territory water pythons remain unclear. The present data are insufficiently informative to distinguish the Indonesian and Northern Territory water pythons as separate lineages, reflected in the inability of Templeton's (1993) test to reject the monophyly of Australian water pythons.

Whilst a study of *Mhc* polymorphisms comparing two specimens from Fogg Dam in the Northern Territory with two specimens from eastern Queensland did not show any differentiation between the two populations (T. Madsen *pers. comm.*), there is evidence from the present mitochondrial study for the presence of two distinct lineages of water pythons in Australia. The monophyly of the "eastern" clade of water pythons would support the

application of the name *L. fuscus* to this lineage. However, because I present a gene tree that may or may not reflect the species tree, allusions to taxonomic classifications need to be taken cautiously and warrant further investigation. Increased sampling of the region between the Northern Territory and Queensland, west of the Great Dividing Range around the Gulf of Carpentaria and investigating nuclear loci, would be useful to elucidate the possible distinction between *L. fuscus* and *L. mackloti.*

6.4.2 Biogeographic implications

Mitochondrial haplotypes from New Guinea and eastern Australia are more closely related than are the haplotypes of eastern Australia and the Northern Territory. Likewise, the Northern Territory haplotypes are more closely related to the Indonesian haplotypes than they are to those of eastern Australia. This haplotypic distribution can be explained by a vicariant event that separated eastern Australian and Northern Territory populations whilst maintaining a link between New Guinean and eastern Australian populations.

I propose that the divergence of the "eastern" from the "western" populations of water pythons within Australia may be the first documented evidence in a land-dwelling organism of the Plio-Pleistocene Torres Strait land bridge causing an east-west directed barrier to gene flow, with the Northern Territory aligning with Indonesia to the exclusion of Queensland and New Guinea. Land barriers such as peninsulae have previously been implicated as the cause of genetic sub-structuring in a regional suite of both inshore marine and coastal organisms (Avise, 1992). Six out of ten coastal species or species–complexes (including birds, reptiles, fish, mollusks and arthropods) surveyed from the Florida peninsula, showed substantial genetic subdivision between the Atlantic Ocean and the Gulf of Mexico (Avise, 1992). Likewise, the Torres Strait land bridge has been implicated as a barrier to gene flow in a teleost fish, the barramundi (Chenoweth *et al.*, 1998; Keenan, 1994), and there are a number of inshore marine organisms such as oysters (Johnson and Joll, 1993), and turtles (Norman *et al.*, 1994) that show strong genetic divergences between populations from the

Australian perimeters of the Indian *versus* the Pacific Oceans.

Geological evidence indicates that Australia and New Guinea have been periodically joined by land for long periods of time (up to half a million years) throughout the Plio-Pleistocene (Galloway and Löffler, 1972). During the most recent glacial maximum at 18,000 years BP, an extensive land bridge was present from the eastern edge of what is now Cape York to the edge of the Arafura Sill, approximately 500km north-east of Darwin (Lambeck and Nakada, 1990) (Fig. 6.3). Potentially, a land bridge between Australia and New Guinea would allow gene flow between populations from the Northern Territory and New Guinea. However, for water pythons, it seems that gene flow between Northern Territory and New Guinea/eastern Australia populations has not occurred in the recent past. The most recent opening of Torres Strait began at the peak of the warmest period 7000 years BP. During the period 50,000 – 10,000 years BP, the sea level sometimes dropped below –50 metres, whence a large lake (Lake Carpentaria) formed in a depression in the present Gulf of Carpentaria. (Jones and Torgersen, 1988; Torgersen *et al.*, 1983; Torgersen *et al.*, 1985) (Fig. 6.3). The lake was connected to the sea by a large westerly flowing river that reached the coast at the edge of the Arafura Sill (Jones and Torgersen, 1988) and until ~35,000 years BP was the endpoint for the Fly River (Torgersen *et al.*, 1988). These large rivers may have caused a barrier to west-coastal taxa moving between Australia and New Guinea. There was a rapid rise in sea level between 18,000 and 6,000 years BP that would have had significant effects on the marine/land interface. The marine incursion would have occurred at a rate of 125m per year (Galloway and Löffler, 1972) requiring coastal populations to find new suitable habitat. The climate postulated for 17,000-8,000 years BP would have sustained an open woodland/savannah/open forest vegetation across the emergent Torres Strait land bridge (Nix and Kalma, 1972) with a swampy flood plain brought on by flooding from intense wet season in the region of Lake Carpentaria (Torgersen *et al.*, 1988). As sea levels rose and marine inundation began to open Torres Strait from the west, the changing climate caused the vegetation of the land bridge to develop an almost continuous, narrow band of closed forest

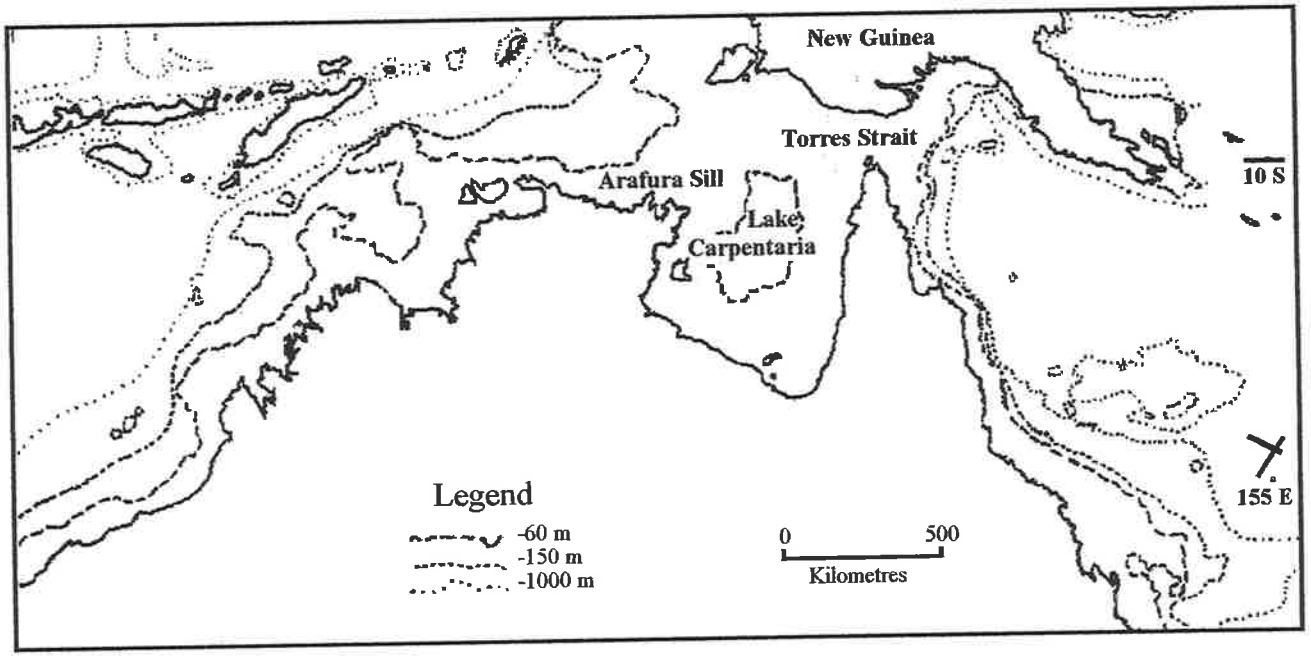


Fig. 6.3. Map of northern Australia showing three depth contours that illustrate the placement of Lake Carpentaria and the extent of the land bridge between Australia and New Guinea. Adapted from Keenan, 1994.

between New Guinea and the east coast of Australia, whilst maintaining woodland/open forest in the west. I hypothesise that the distribution of water pythons on the land bridge between Australia and New Guinea was limited by their requirement for watery habitats in drainage areas along the coast. They would have been isolated by the dry open forest of the interior of the land bridge, leading to the divergence between the "eastern" and "western" clades within Australia.

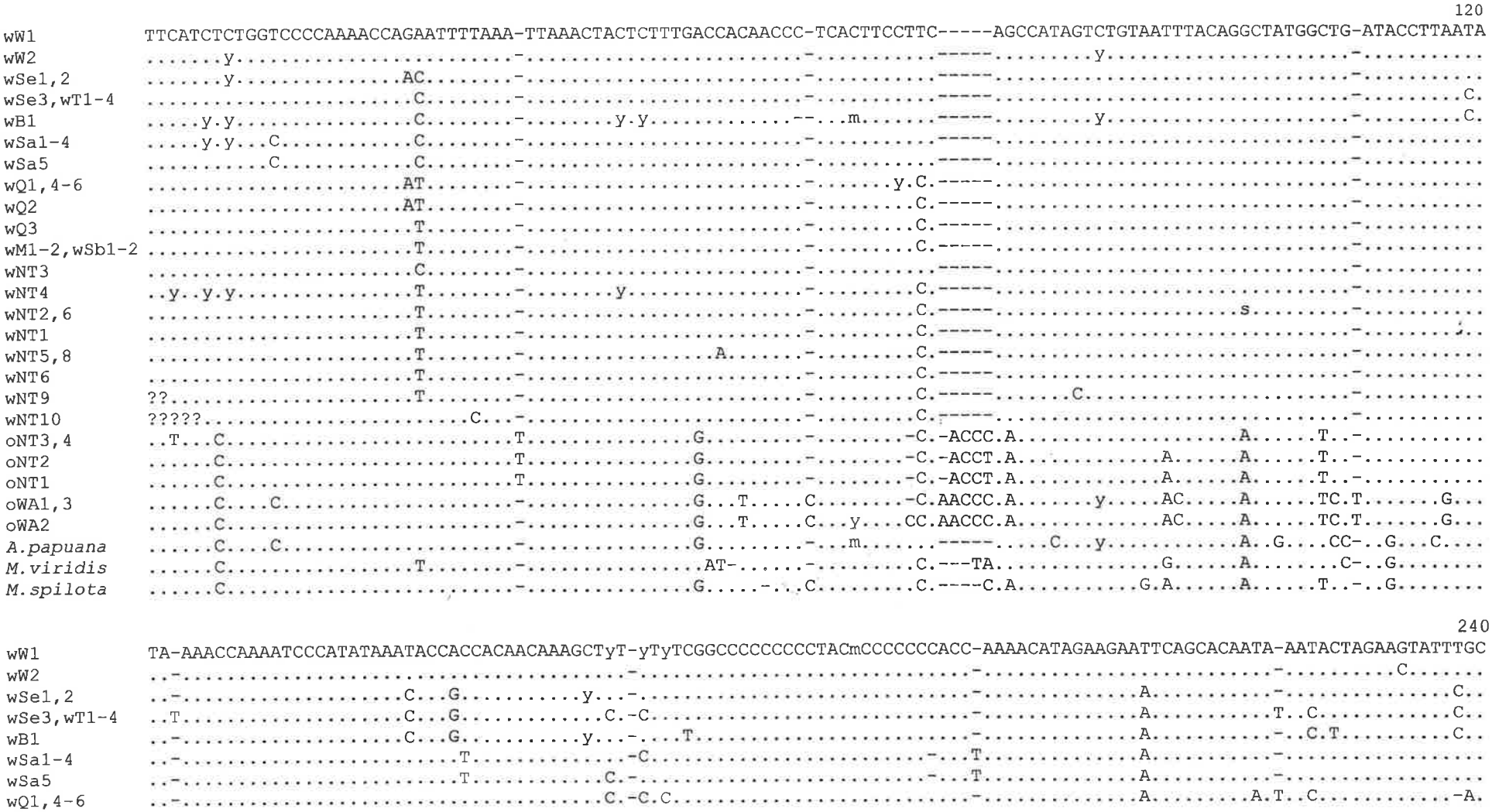
A second vicariant event to affect the distribution of the mitochondrial haplotype distributions of water pythons could have been the widening of the sea barrier between the Lesser Sunda Islands of Indonesia and the emergent Australian landmass. During the height of the last glacial maximum, the Lesser Sunda Islands may have been separated from the Australian landmass by as little as 72-99 km (Goodwin, 1979; Hooijer, 1975) possibly close enough for individuals to raft between the two land masses. This could account for the close genetic relationship of Northern Territory populations with those from the Indonesian islands.

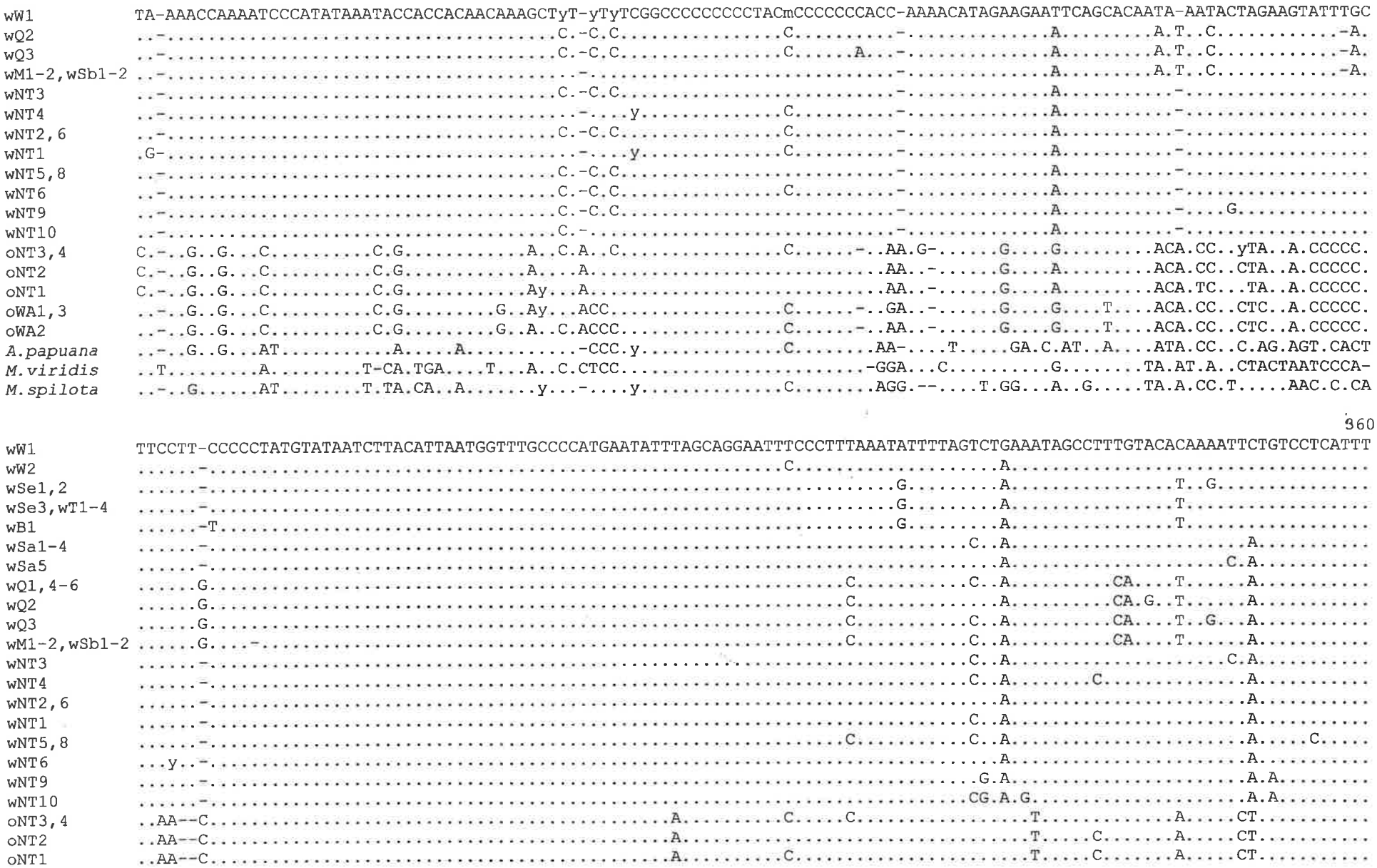
In the absence of suitable tools for establishing a molecular clock, I do not attempt to estimate the timing of divergence between the "eastern" and "western" lineages. It would first be necessary to calibrate of the rate of molecular evolution for a gene suited to detecting more recent divergences. Calibration of the rate of molecular evolution is also dependent on the availability of well-dated fossils related to the taxa under consideration.

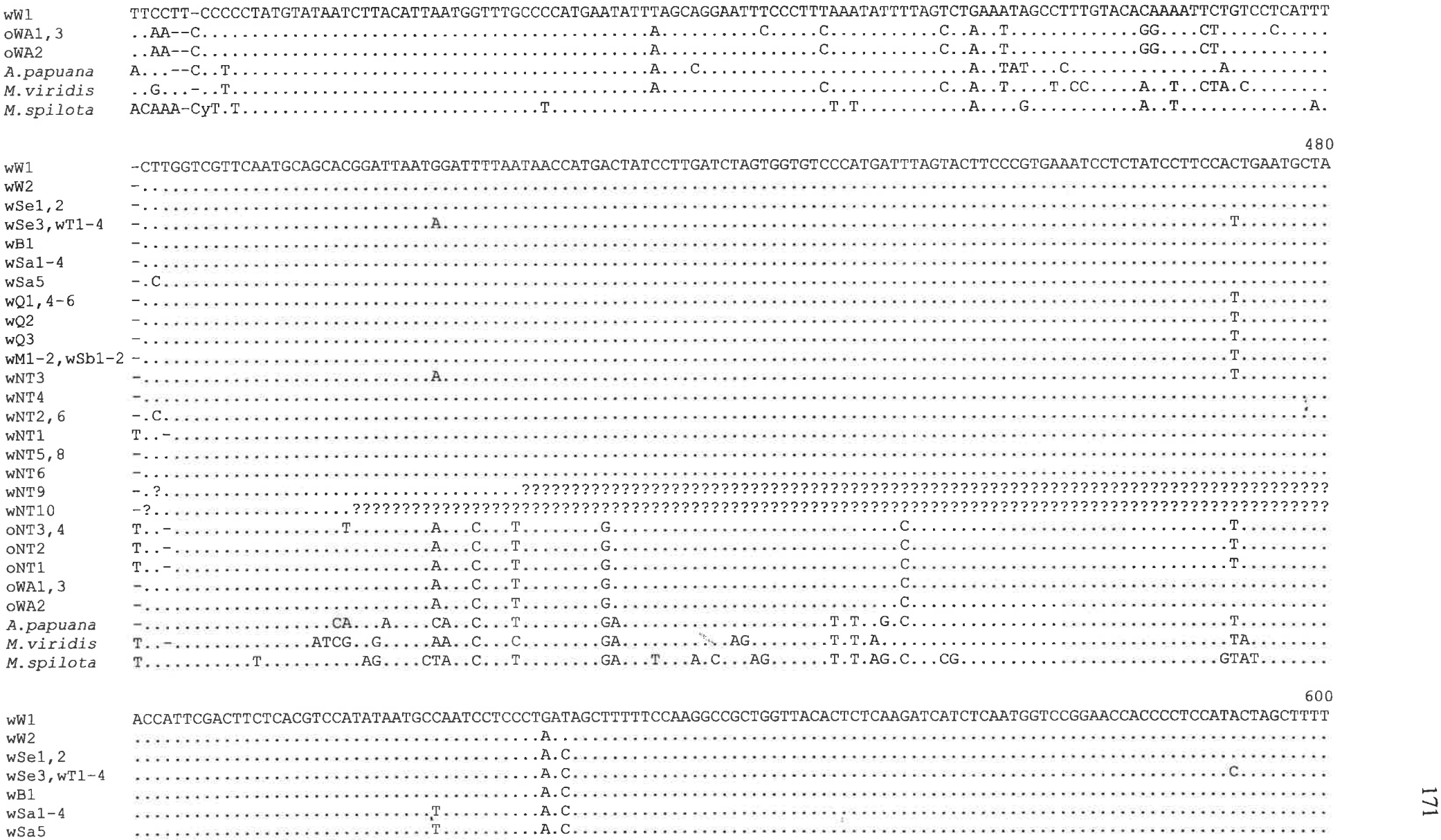
The role of the Pleistocene Torres Strait land bridge as an east-west barrier to gene flow can be further investigated by examining a suite of semi-aquatic, freshwater-dependent taxa from northern Australia for congruence in phylogeographic patterns.

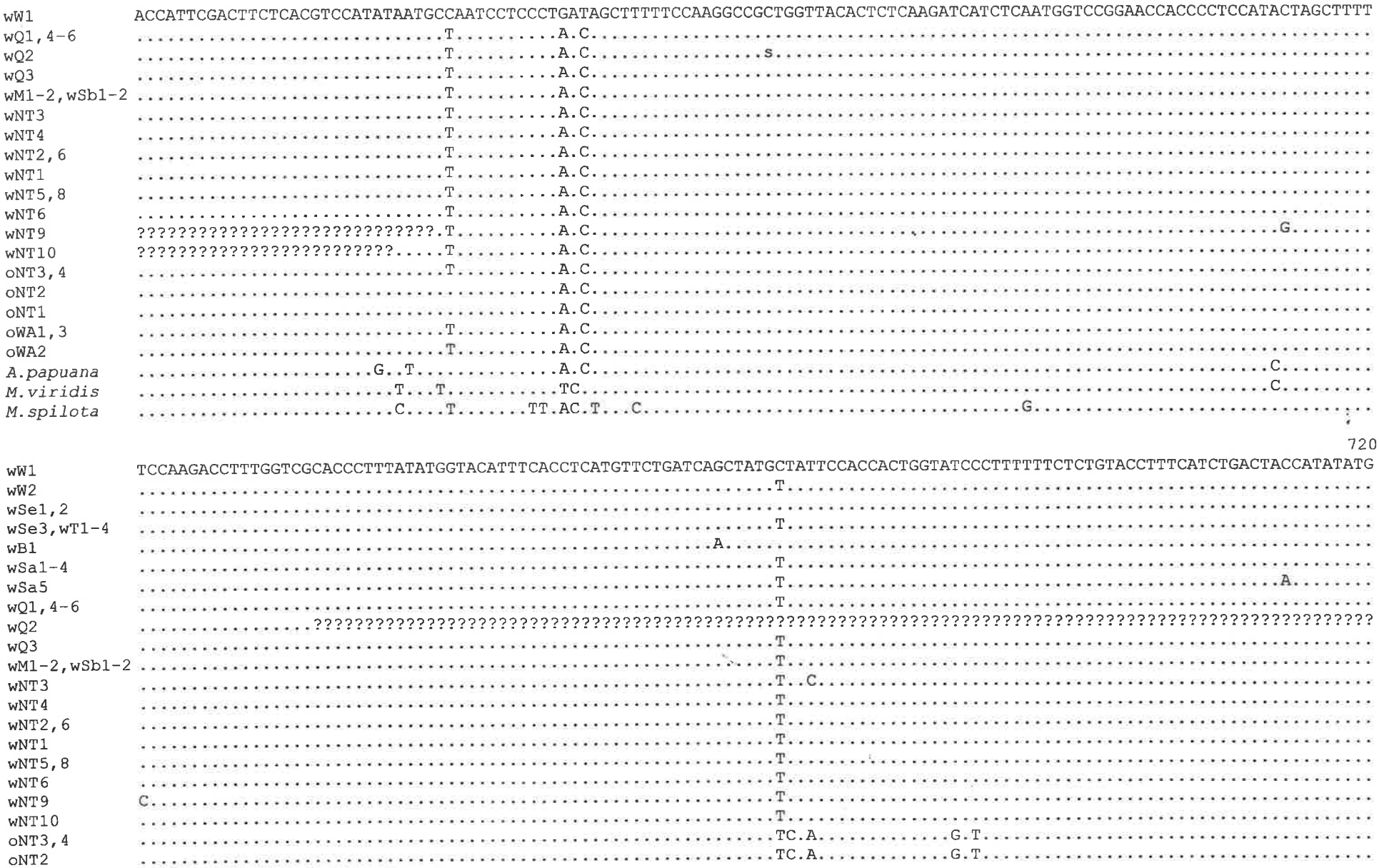
Appendix 6.1.

Alignment of nucleotide sequences for *Liasis mackloti/fuscus* (w) and *L. olivaceus* (o) samples used in this study for partial *control region*. Due to sequencing difficulties, sequences are shorter for wQ2 (617 bp), wNT9 (684 bp) and wNT10 (671 bp).




wW1 TA-AAACCAAAATCCCATATAAATACCACCACAACAAAGCTyT-yTyTCGGCCCCCCCCCTACmCCCCCCCACC-AAAACATAGAAGAATTCAGCACAATA-AATACTAGAAGTATTTGC
wQ2
wQ3
wM1-2,wSb1-2
wNT3
wNT4
wNT2,6
wNT1
wNT5,8
wNT6
wNT9
wNT10
oNT3,4
oNT2
oNT1
oWA1,3
oWA2
A.papuana
M.viridis
M.spilota
360
wW1 TTCCTT-CCCCCTATGTATAATCTTACATTAATGGTTTGCCCCATGAATATTTAGCAGGAATTTCCCTTTAAATATTTTAGTCTGAAATAGCCTTTGTACACAAAATTCTGTCCTCATTT
wW2
wSe1,2
wSe3,wT1-4
wB1
wSa1-4
wSa5
wQ1,4-6
wQ2
wQ3
wM1-2,wSb1-2
wNT3
wNT4
wNT2,6
wNT1
wNT5,8
wNT6
wNT9
wNT10
oNT3,4
oNT2
oNT1


wW1 TTCCTT-CCCCCTATGTATAATCTTACATTAATGGTTTGCCCCATGAATATTTAGCAGGAATTTCCCTTTAAATATTTTAGTCTGAAATAGCCTTTGTACACAAAATTCTGTCCTCATTT
oWA1,3 ..AA--C.....A.....C.....C.....C..A..T.....GG....CT.....C.....
oWA2 ..AA--C.....A.....C.....C..A..T.....GG....CT.....
A.papuana A...--C..T.....A...C.....A..TAT...C.....A.....
M.viridis ..G...-..T.....A.....C.....C..A..T....T.CC.....A..T..CTA.C.....
M.spilota ACAAA-CyT.T.....T.....T.T.....A....G.....A..T.....A.
480
wW1 -CTTGGTCGTTCAATGCAGCACGGATTAATGGATTTTAATAACCATGACTATCCTTGATCTAGTGGTGTCCCATGATTTAGTACTTCCCGTGAAATCCTCTATCCTTCCACTGAATGCTA
wW2 -.....
wSe1,2 -.....
wSe3,wT1-4 -.....A.....T.....
wB1 -.....
wSa1-4 -.....
wSa5 -.C.....
wQ1,4-6 -.....T.....
wQ2 -.....T.....
wQ3 -.....T.....
wM1-2,wSb1-2 -.....T.....
wNT3 -.....A.....T.....
wNT4 -.....
wNT2,6 -.C.....
wNT1 T..-.....
wNT5,8 -.....
wNT6 -.....
wNT9 -.?.....??
wNT10 -?.....??
oNT3,4 T..-.....T.....A...C...T.....G.....C.....T.....
oNT2 T..-.....A...C...T.....G.....C.....T.....
oNT1 T..-.....A...C...T.....G.....C.....T.....
oWA1,3 -.....A...C...T.....G.....C.....
oWA2 -.....A...C...T.....G.....C.....
A.papuana -.....CA...A....CA..C...T.....GA.....T.T..G.C.....T.....
M.viridis T..-.....ATCG..G.....AA..C...C.....GA.....AG.....T.T.A.....TA.....
M.spilota T.....T.....AG....CTA..C...T.....GA...T...A.C...AG.....T.T.AG.C...CG.....GTAT.....
600
wW1 ACCATTCGACTTCTCACGTCCATATAATGCCAATCCTCCCTGATAGCTTTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTT
wW2 .....A.....
wSe1,2 .....A.C.....
wSe3,wT1-4 .....A.C.....C.....
wB1 .....A.C.....
wSa1-4 .....T.....A.C.....
wSa5 .....T.....A.C.....


wW1 ACCATTCGACTTCTCACGTCCATATAATGCCAATCCTCCCTGATAGCTTTTTCCAAGGCCGCTGGTTACACTCTCAAGATCATCTCAATGGTCCGGAACCACCCCTCCATACTAGCTTTT
wQ1,4-6
wQ2
wQ3
wM1-2,wSb1-2
wNT3
wNT4
wNT2,6
wNT1
wNT5,8
wNT6
wNT9
wNT10
oNT3,4
oNT2
oNT1
oWA1,3
oWA2
A.papuana
M.viridis
M.spilota
720
wW1 TCCAAGACCTTTGGTCGCACCCTTTATATGGTACATTTCACCTCATGTTCTGATCAGCTATGCTATTCCACCACTGGTATCCCTTTTTTCTCTGTACCTTTCATCTGACTACCATATATG
wW2
wSe1,2
wSe3,wT1-4
wB1
wSa1-4
wSa5
wQ1,4-6
wQ2
wQ3
wM1-2,wSb1-2
wNT3
wNT4
wNT2,6
wNT1
wNT5,8
wNT6
wNT9
wNT10
oNT3,4
oNT2

```
wW1            TCCAAGACCTTTGGTCGCACCCTTTATATGGTACATTTCACCTCATGTTCTGATCAGCTATGCTATTCCACCACTGGTATCCCTTTTTTCTCTGTACCTTTCATCTGACTACCATATATG
oNT1           ..............................................................TC.A.............G.T......................................
oWA1,3         ..............................................................TC.A.............G.T......................................
oWA2           ..............................................................TC.A.............G.T......................................
A.papuana      ...................................C..........................TC.A.............G.T..C...................................
M.viridis      ....................................C.TGT...............C........G............T.............G...........................
M.spilota      ......................C.............A...................T.....TCTAC......T....TG.T..C....T...................C..........

                                                                                      795
wW1            CACACACACAGTTAGGCTTCCAGTCCAGGTGGAGCACTATGCTTTCTCTTACACTTCCCCTATAGGGATACATCT
wW2            ...........................................................................
wSe1,2         .................................................C.........................
wSe3,wT1-4     ...........................................................................
wB1            ...................................G.......................................
wSa1-4         ............A..............................................................
wSa5           ............A..............................................................
wQ1,4-6        ............A..............................................................
wQ2            ???????????????????????????????????????????????????????????????????????????
wQ3            ............A.......................................G............??????????
wM1-2,wSb1-2   ............A..............................................................
wNT3           ..................................................G........................
wNT4           ...........................................................................
wNT2,6         ...........................................................................
wNT1           ...........................................................................
wNT5,8         ...........................................................................
wNT6           ...........................................................................
wNT9           ...........................................................................
wNT10          ...........................................................................
oNT3,4         ............A......TT.........................CT........T.......T..........
oNT2           ............A......TT.......................C.CTC-G.....T.......T..........
oNT1           ............A......TT.......................C.CTC-G.....T.......T..........
oWA1,3         ............A......TT.......................C..TC-.....CT.......T..........
oWA2           ............A......TT.......................C..TC-....CCT.......T..........
A.papuana      ............A......TT.........A.............C.C.C.CAC..A...ATGG.AT.C.TC...
M.viridis      ............A.....CTT.....G......A..TG...T.A-T..C.CATT.A..G.A.--T......????
M.spilota      .................A.TT.....G...T......GC.....C...C.CATT.A......--...........
```

173

Chapter 7

Concordant molecular phylogeographic structuring in six New Guinean vertebrates

7.1 Introduction

The power of comparative phylogeography lies in its ability to detect recent, broad-ranging influences on population structuring (Avise, 1992; Riddle, 1996; Taberlet, 1998). By considering a range of species that are diverse in habitat preferences and life histories, the finding of phylogeographic congruence implies a common event that has overridden the unique aspects of the ecology of individual taxa. A common, wide-ranging influence that has affected the distributions of the species tested may have far-reaching implications in conservation efforts for other co-distributed fauna and flora.

Phenotypic divergences between conspecific populations or sibling species are often predicted to be primarily influenced by the most recent Pleistocene Ice Ages because the taxa involved are thought to be very young and there are no other obvious barriers to dispersal. Indeed, this situation has often been behind theories to explain the generation of genetic diversity in lowland rainforest communities, *e.g.* the Pleistocene Refuge model (Haffer, 1969). Recent comparative phylogeographic studies among vertebrates from tropical Australia and the Americas have aimed to test the effect of the expansion/contraction of rainforests due to the latest Pleistocene climate oscillations on the historical aspects of species distributions and the distribution of intraspecific genetic diversity (da Silva and Patton, 1998; Patton and da Silva, 1998; Schneider *et al.*, 1998). These studies found that a large proportion of intra-specific genetic divergences with strong geographic partitioning predate the predicted influence of the Pleistocene Ice Ages.

The island of New Guinea presents another opportunity to test the Pleistocene Refuge model of speciation for tropical rainforests. New Guinea is the world's largest and most

biologically diverse tropical island, and has a continuous "ring" of lowland rainforest and lower montane or hill forest (Johns, 1982; Pratt, 1982) skirting the central mountain range or cordillera. While the width of these forests varies considerably around the island, both forest types are thought to be continuously distributed around the island (Pratt, 1982). In any biogeographic analysis of the New Guinean biota throughout the Tertiary and Quaternary, the influence of two factors cannot be ignored: the substantial changes in the geomorphology of the area due to complex tectonism over the entire period (Dow, 1977; Pigram and Davies, 1987), and the dramatic climate oscillations of the late Tertiary and the Quaternary (Axelrod and Raven, 1982; Haig and Medd, 1996; Read and Hope, 1996). During this period, tectonic movements due to continental and island arc collisions have seen the formation and infilling of several major sedimentary basins to produce new lowland habitats and the rapid uplift of the central cordillera (Dow, 1977; Hill and Gleadow, 1989; Hill *et al.*, 1993). Vegetation and faunal evidence suggests that climate oscillations caused major shifts in the altitudinal range of many plant and animal species, with lowland rainforest becoming "patchier" and the upper altitudinal limit of lower montane forest becoming depressed, the effect varying according to the aspect and region (Hope, 1996a; Pasveer and Aplin, 1998). Unfortunately, palaeo-environmental reconstructions are too few at the moment to derive a wide-ranging temporal and geographic perspective across New Guinea. However, differences in the magnitude of the climate oscillations could have caused a differential retreat of lowland/lower montane rainforest to higher elevations, with steep valley/ridge systems along the heavily dissected central cordillera leading to fragmentation of the rainforest. Rainforest refugia formed in this way could have produced divergences amongst isolates without necessarily producing phylogeographic concordance among different taxa.

The ranges of the north coast and the Vogelkop, in northwestern New Guinea, could also have acted as rainforest refugia, promoting diversification of isolated populations to be reconnected later by a continuous lowland rainforest cover. Their involvement as "refuges" could have begun when they were part of the 'Outer Melanesian Arc" before they achieved

accretion with the Australian craton to form the emergent landmass of present day New Guinea. It is also prudent to consider the role that the adjacent landmass of Australia, repeatedly connected by land-bridges to southern New Guinea in the Plio-Pleistocene, could have played as another lowland rainforest refuge area.

Given the complexity of the Neogene geomorphological and climatic history of New Guinea, it would be surprising if wide scale patterns of phylogeographic concordance would be observed on the island. Indeed, in general, continent-wide phylogeographic concordance is far less frequent in comparison with the strong patterns of concordance observed for smaller geographic regions (reviewed by Taberlet *et al.*, 1998). Nevertheless, if any factor influencing past rainforest distributions had a dominant effect, then patterns of concordance might be expected. Evidence from the distributions of lowland rainforest and lower montane forest birds (Pratt, 1982) and marsupials (Colgan *et al.*, 1993) supports the latter contention. Many apparent sibling species of birds and several species or subspecies pairs of marsupials show concordant patterns of distribution within each of these two groups. However, the pattern of concordance at the eastern end of the island differs between birds and marsupials. Furthermore, there is no estimate available of when the diversification patterns arose for any of these taxa.

I tested the generality of the pattern of taxic diversification in New Guinean lowland rainforest vertebrates observed for birds (Pratt, 1982) and marsupials (Colgan *et al.*, 1993) by using a molecular phylogeographic approach on a selection of snakes and a rodent. I studied the intra-specific phylogeography of one colubrid snake (*Boiga irregularis*), two congeneric boine snake species (*Candoia aspera* and *C. carinata*), two pythonine snake species (*Leiopython albertisii* and *Morelia viridis*), and a murid rodent (*Melomys rufescens*). These taxa have a wide range of differing ecologies and life histories and importantly they are co-distributed throughout lowland rainforest and lower-mid montane forest (Flannery, 1995; O'Shea, 1996). I used a mitochondrial molecular clock to estimate the timing of initial events promoting divergences, providing new information to assist in the analysis of historical

influences on the distribution of genetic diversity for an important component of the world's biodiversity.

7.2Materials and Methods

7.2.1 Materials

Samples were available from a wide geographic region of eastern New Guinea (Fig. 7.1). Collection locality and museum voucher details of the specimens sequenced in this study are listed below. Some *Candoia* sp. sequences were supplied by C. Austin and the *Melomys cervinipes* and *Morelia viridis* sequences were supplied by J. Birrell. These are marked by ^{A} and ^{B} respectively. $^{\#}$ = samples used for allozyme electrophoresis only.

Ten ***Boiga irregularis*** AMS R124799 Usino, Madang Prov. (MP); AMS R124360 Nokopo, Morobe Prov. (MoP); BPBM 13758 Gusap, MoP; AMS R135588 Mt Sue, Sandaun Prov. (SP); AMS R130424 Wilbeite, SP; AMS R115331 Yuro, Simbu Prov. (SiP); AMS R115357 Noru, SiP; AMS R122371 Bobole, Southern Highlands Prov. (SHP); AMS R122373/368 Namosado, SHP. Two ***B. cynodon*** Flores Is, Indonesia WAM R104558, Sumba Is, Indonesia WAM R101936. Eleven ***Candoia aspera*** AMS R124100$^{\#}$/1^{A}, Madang, MP; AMS R1335506$^{\#}$/7^{A}/8^{A}/9$^{\#}$ Mt Sue SP; AMS R122352^{A}/3^{A} Namosado SHP; AMS R115354$^{\#}$ Noru, SiP; AMS R115337$^{\#}$/8* Yuro, SiP. Seven ***C. carinata*** AMS R135578 Mt Sue SP; AMS R124689 Nobonob MP; AMS R124623 Karkar Is, MP; AMS R122347$^{\#}$/8$^{\#}$ Fogamaiyu SHP; AMS R122345/9 Waro SHP. Six ***Leiopython albertisii*** AMS R115346 Doido SiP; AMS R124686/798 Madang area, MP; AMS R124481 Kabum MP, BPBM 11612/3 Wau, Morobe. Eleven ***Morelia viridis*** AMS R122363^{B}/4^{B} Namosado SHP; AMS R115348^{B} /49^{B}/50^{B} Doido, SiP; AMS R115355^{B}/6^{B} Noru, SiP, AMS R129716^{B} Normanby Is., MBP; AMS R124531^{B} Bundi MP; BPBM 13080^{B}/1^{B} Wau, Morobe. One ***Melomys cervinipes*** QM M10914^{B} Kroombit Tops, Qld. Nine ***M. rufescens*** AMS M15955 Telefomin, SP; AMS M15956 Wigote, SP; AMS M21678 Madang, MP; AMS M21677 Karkar Island MP; AMS M16248 Namosado, SHP; AMS M19859 Waro, SHP; AMS M15162 Yuro, SiP;

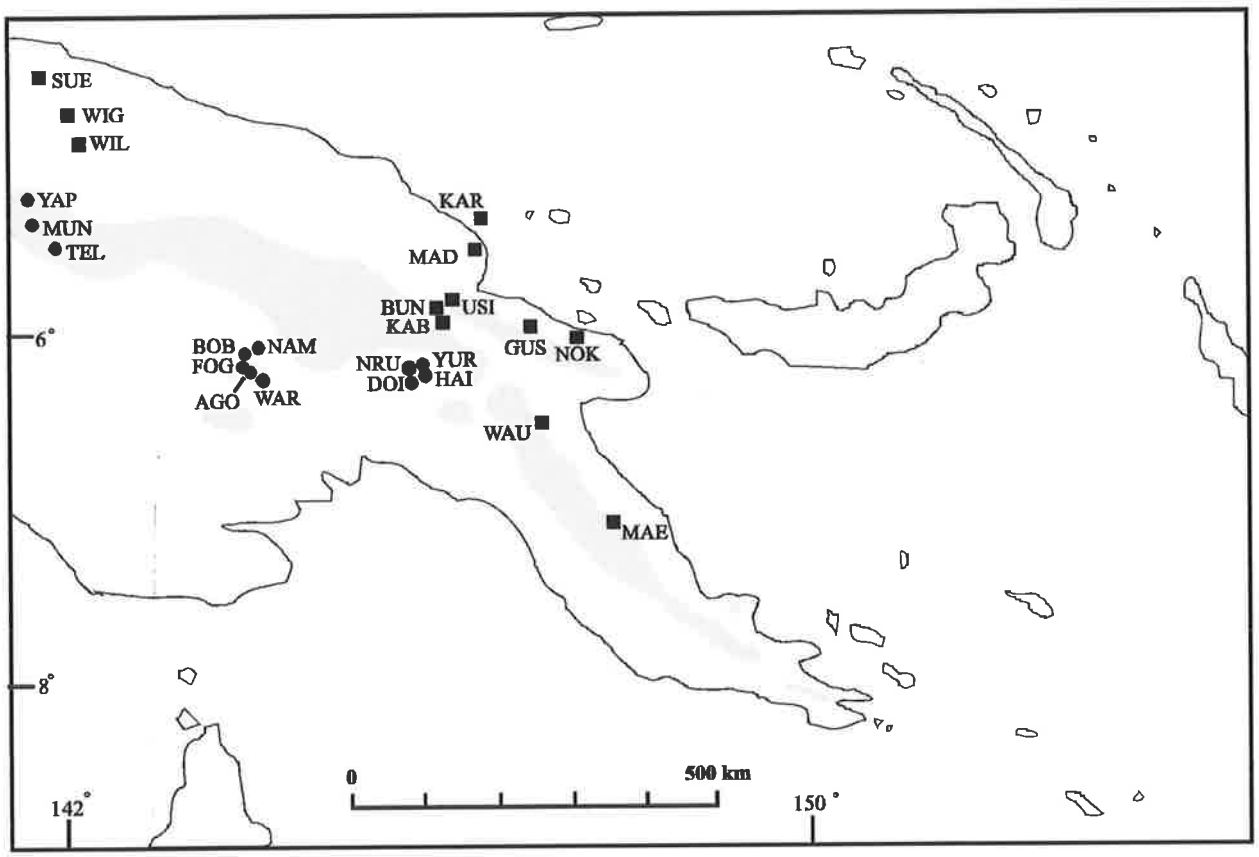


Fig. 7.1. Map of eastern New Guinea showing sample locations. The locality code is: AGO=Agofia, BOB=Bobole, BUN=Bundi, DOI=Doido, FOG=Fogamaiyu, GUS=Gusap, HIA=Haia, KAB=Kabum, KAR=Kar Kar Is., MAD=Madang, MAE=Mt. Albert Edward, MUN=Munbil, NAM=Namasado, NOK=Nokopo, NRU=Noru, SUE=Mt. Sue, TEL=Telefomin, USI=Usino, YAP=Yapsiei, YUR=Yuro. Shading represents the 2,000 metre contour of the central mountain range. Squares represent 'northern' populations, circles represent 'southern' populations.

AMS M14732 Noru, SiP; AMS M12651 near Kosipe, Mount Albert Edward, Central Prov. One ***Python reticulatus*** SAMA R28534 no location data.

7.2.2 Mitochondrial genes

DNA was extracted by methods described in chapter 2.2. A ~360 bp product of the partial *cytochrome b* (*cytb*) gene was amplified by PCR (protocols in chapter 2.7.1 and 2.7.3.1) and directly sequenced with primers L14841 and H15149. Aligned sequences generated in the present study are listed in Appendix 7.1. The boid and pythonine trees were reciprocally rooted, with each species or genus respectively acting as the outgroup for the other. Other outgroups and species used for molecular clock calibration (GenBank accession numbers where appropriate) included *Boiga cynodon*, *Coluber constrictor* (U49300), *Elaphe flavolineata* (U49301), *Melomys cervinipes*, *Mus* (Z96069), *Rattus norvegicus* (J01436) and *Python reticulatus*. Sequences were aligned by eye, as there was no length variation. Analysis of the aligned sequences with Modeltest3.0 (Posada and Crandall, 1998) suggested the Kimura 2-parameter model of evolution (Hasegawa *et al.*, 1985) for maximum likelihood and distance analyses. Phylogenetic trees were produced from the aligned sequences using all nucleotide positions in PAUP*4.0b2a, by unweighted maximum parsimony (MP) and maximum likelihood (ML) and the Neighbour-Joining algorithm (NJ) using the Kimura 2-parameter model of evolution.

7.2.3 Molecular clock analysis

For the purposes of molecular clock dating, sequence divergence was estimated according to the genetic distances calculated under the Kimura 2-parameter model (Hasegawa *et al.*, 1985). *Cytb* data were not available for the colubrid snake taxa used for fossil-based calibration, instead I used the mitochondrial *ND4* data of Kraus & Brown (Kraus and Brown, 1998). Because *cytb* has a fast rate of base substitution leading to saturation at third base positions, I estimated divergence using transversion substitutions only. Rate homogeneity

within each taxon (including the species used to calibrate our molecular dating approach) was tested by Tajima's (Tajima, 1993) test at a 95% confidence level. Time of separation of major clades within *Leiopython albertisii, Morelia viridis* and *Melomys rufescens* was estimated by the method described in chapter 2.13 equation 1. The times of separation of major clades within each *Candoia* species, for which there were no suitable fossils, were estimated as described in chapter 2.13 equation 3. For comparison of evolutionary rates with other vertebrates, the evolutionary rate (% substitutions/million years) was estimated using chapter 2.13 equation 2.

7.2.4 Allozyme Electrophoresis

Allozyme electrophoresis was done by S. Donnellan for four taxa examined in this study according to Richardson *et al.* (1986). The enzymes stained were: AAT, ACOH, ACP, ADA, AK, ALAT, ALDH, CA, ENO, EST, FBP, FUMH, GDA, GPDH, GPI, GSR, GUK, IDDH, IDH, LDH, LGL, MDH, MPI, NDPK, PEP-A, PEP-B, PEP-C, PEP-D, PGAM, PGDH, PGK, PGM, PNP, SOD, and TPI. (Allozyme frequencies are listed in Appendix 7.2).

Phenograms were generated by the unweighted pair group method using arithmetic averages (UPGMA) from matrices of percentage of fixed allele differences from 28 loci for boines, 36 loci for pythonines, and 24 loci for rodents. Allozyme frequencies for *Melomys rufescens* were obtained from Flannery *et al.* (1994), excluding the Central Province population because of missing data, and for *Boiga* from L.Rawlings (unpublished honours thesis).

7.3 Results

7.3.1 Phylogeographic patterns

Phylogenetic relationships of a portion of the mitochondrial *cytb* gene were recovered with unweighted MP and NJ from Kimura 2-parameter distances. To display the evolutionary distances among samples, I present the NJ trees in which branch lengths are proportional to evolutionary distance (Fig. 7.2i and 7.3i). Maximum parsimony trees showed similar overall

topologies and are not presented. Allozyme UPGMA phenograms are shown in Fig. 7.2ii. For *Boiga irregularis*, the *Candoia* species and *Morelia viridis*, phylogenies displayed reciprocal monophyly of clades either side of the central cordillera, hereafter referred to as the "north/south clades" (Fig. 7.2i *a-b* and 7.3i *a*). There was strong bootstrap support (>70%) for reciprocal monophyly in almost every case. For *Melomys* (Fig. 7.3i *b*), the NJ tree has the southern populations monophyletic to the exclusion of the northern populations which form two groups, albeit with poor bootstrap support for paraphyly of the northern populations. However, the northern and southern populations are reciprocally monophyletic in the MP analysis and they group reciprocally in the allozyme analysis (see below). For *Leiopython* (Fig. 7.3i *a*), there is a dichotomy between the Wau population (north of the central cordillera) and Doido (south of the central cordillera), with both populations monophyletic to the exclusion of a second group from the northern side of the central cordillera (96% bootstrap support). For each of the six species, the north/south clades represent immediate sister lineages in comparison with their closest congeners or confamilial relatives (see Chapter 5, L.H. Rawlings unpublished honours thesis; Austin, 2000; Torrance, 1997).

Divergence at nuclear loci, assessed by allozyme electrophoresis, was used to give another measure of genetic divergence and to test whether the observed pattern is due to population history or gene specific phenomena such as gene trees. In the absence of any evidence about constancy of rates of allozyme evolution, I do not draw conclusions about phylogenetic relationships from these trees. With one exception, each of the UPGMA trees showed a distinct dichotomy correlated with the north/south clades seen in the mitochondrial trees. For *Candoia carinata* (Fig. 7.2ii *b*), a multifurcation of northern and southern populations is present, but both southern populations form a subgroup within the multifurcation. The lack of resolution in this phenogram arises because one third of the loci in some populations could not be typed due to poor tissue quality. The magnitude of percentage of loci showing fixed allelic differences largely reflects the extent of sequence divergence in *cytb*. Indeed, the means of Kimura 2-parameter distances across all nucleotide

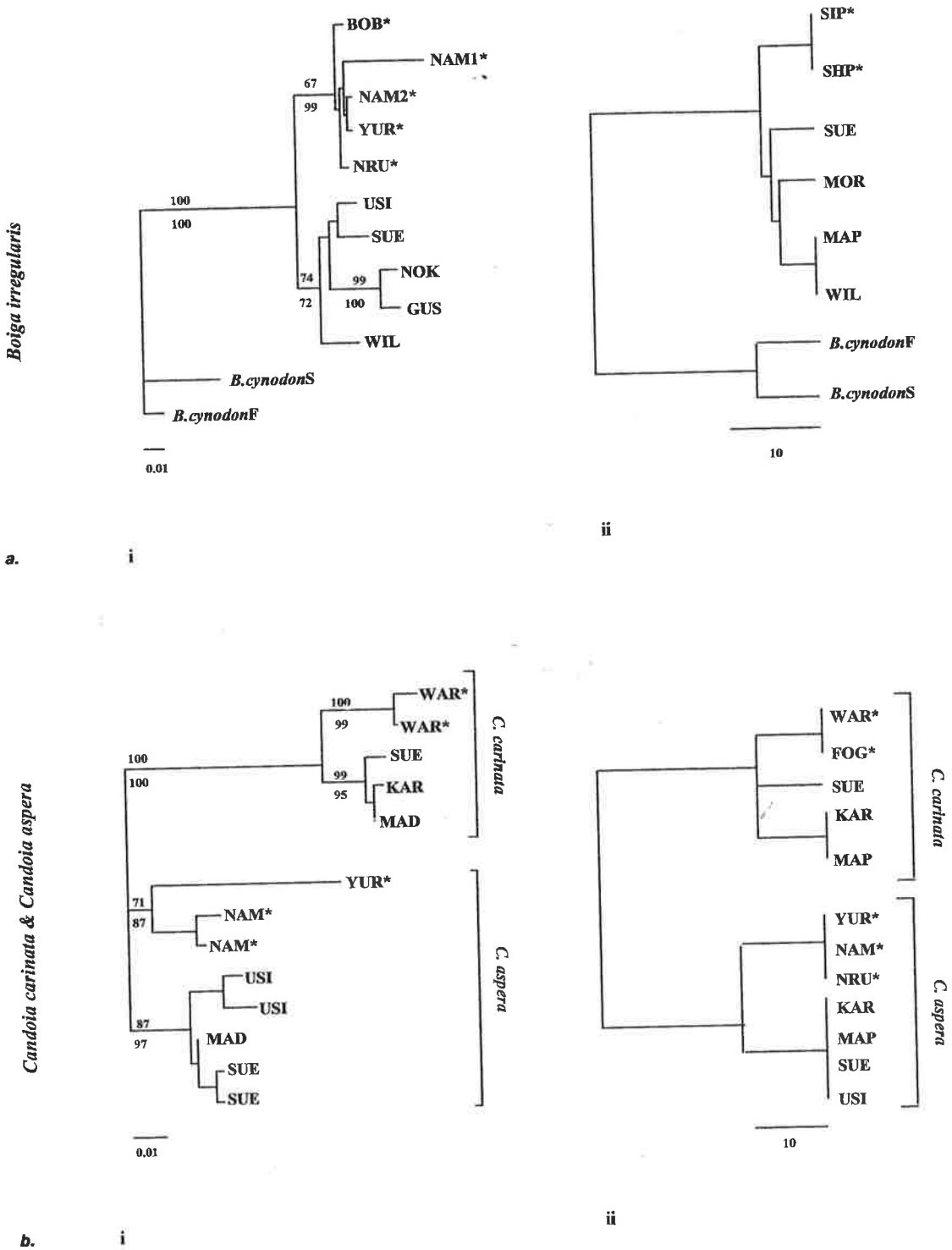


Fig. 7.2. Mitochondrial gene tree and allozyme tree of **a)** *Boiga irregularis* and **b)** two *Candoia* species. i.) Neighbour Joining (NJ) trees of Kimura 2-parameter distances among mitochondrial *cytochrome b* haplotypes. Numbers above and below branches represent MP and NJ bootstrap proportions from 1000 and 2000 pseudoreplicates respectively and the asterisk denotes southern New Guinean populations. ii.) UPGMA trees of percentage of loci showing fixed allelic differences. The scale represents 10% fixed allelic differences. Terminal taxon names correspond to locations on Fig. 7.1 or where populations were pooled for analysis **MAP** = Madang Province (Kabum [KAB], Madang [MAD], Usino [USI]), **MOR** = Morobe Province (Gusap [GUS], Nokopo [NOK]), **SHP** = Southern Highlands Province (Agofia [AGO], Bobole [BOB], Namasado [NAM], Waro [WAR]), **SIP** = Simbu Province (Doido [DOI], Haia [HAI], Noru [NRU], Yuro [YUR]), For B. cynodon **F** = Flores Is, Indonesia and **S** = Sumba Is., Indonesia.

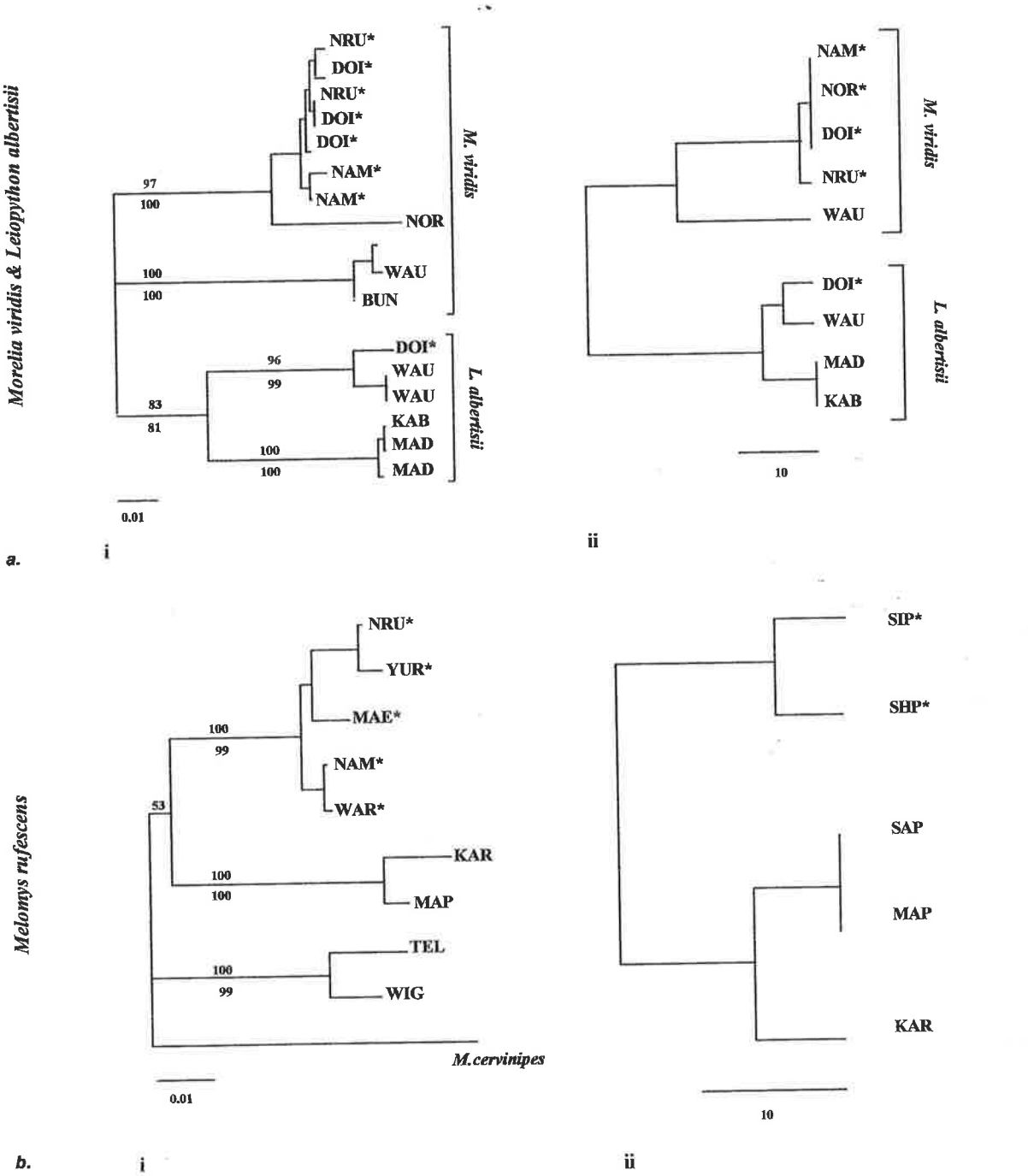


Fig. 7.3. Mitochondrial gene tree and allozyme tree of *a.*) two pythonine snake species (*Leiopython albertisii* and *Morelia viridis*) and *b.*) *Melomys rufescens*. i.) NJ trees of Kimura 2-parameter distances among mitochondrial *cytochrome b* haplotypes. (ii.) UPGMA trees of percentage of loci showing fixed allelic differences. The scale represents 10% fixed allelic differences. Terminal taxon names correspond to locations on Fig. 7.1. or where populations were pooled for analysis **MAP** = Madang Province (Kabum [KAB], Madang [MAD], Usino [USI]), **SAP** = Sandaun Province (Munbil [MUN], Telefomin [TEL], Wigote [WIG], Yapsiei [YAP]), **SHP** = Southern Highlands Province (Agofia [AGO], Bobole [BOB], Namasado [NAM], Waro [WAR]), **SIP** = Simbu Province (Doido [DOI], Haia [HAI], Noru [NRU], Yuro [YUR]).

sites between north/south clades are highly correlated with the means of percentage of loci showing fixed allelic differences ($r^2 = 0.89$, $F_{1,4} = 34.85$, $p < 0.0035$).

The concordance of the geographical arrangement of genetic profiles between the DNA and allozyme analyses across all six species shows that population sub-structuring is not random and implies that some wide-reaching phenomenon has impacted on their history. Furthermore, the relative amount of mitochondrial sequence divergence and the number of fixed allelic differences between the north/south clades imply that their separation is not due to the latest of the Plio-Pleistocene climatic oscillations.

7.3.2 Timing of Diversification

I used a mitochondrial molecular clock to estimate the timing of north/south clade divergence. Initially, I tested for the presence of a molecular clock with the Tajima test (Tajima, 1993). As none of the tests were significant, I was unable to reject the clock hypothesis in any comparison. To account for the possibility that each lineage may have a different rate of molecular evolution (Lopez *et al.*, 1997), divergence times were estimated separately for each lineage. As there are no fossils of any of the test species presently known, I used the divergence from the nearest sister lineage with an adequately described and dated fossil taxon as the calibration point for each lineage. The *Python* to *Leiopython/Morelia* divergence dates to at least 25 MYA (Rage, 1987; Scanlon, 1996), and I used the pythonine calibration for the boine lineage as suitable fossils are not known for the boines, and pythonines and boines are immediate sister lineages (Kluge, 1991). For *Boiga*, the evolutionary rate was calculated from the mitochondrial *ND4* sequence divergence of two colubroid species, *Coluber constrictor* and *Elaphe flavolineata*, for which the earliest fossils are dated to the early Miocene at 22.5 MYA (Rage, 1987).

There is much controversy over the age of the *Mus/Rattus* divergence. Fossil dates for *Mus* and *Rattus* are estimated at approximately 12 MYA (Catzeflis *et al.*, 1992). Much earlier

divergence times have been estimated using sequence divergences from other mammalian lineages and their relevant fossils (e.g. Janke *et al.*, 1994), however, I have chosen to be conservative and use the rodent fossil date. Whilst nucleotide substitution in *cytb* is not always clock-like at deeper divergences (Janke *et al.*, 1994), the Tajima test on our data did not reject rate homogeneity for the *Melomys/Mus/Rattus* lineages. The corrected sequence divergence between *Mus* and *Rattus* in the *cytb* gene for all substitutions and for transversions only is 20% and 7% respectively. These data are consistent with estimates of sequence divergence between *Mus* and *Rattus* determined from many nuclear genes (Bulmer *et al.*, 1991; O'hUigin and Li, 1992). The rate of sequence evolution between *Mus* and *Rattus cytb* genes for all substitutions is 0.87% per MY and for transversions is 0.3% per MY. The rate for transversions only at third codon positions is 0.9% per MY. This rate is faster than the artiodactyl rate for *cytb* of 0.5% per MY (Irwin *et al.*, 1991). Honeycutt *et al.* (1995) also found that rodents have a faster rate for *cytb* and the mitochondrial *cytochrome oxidase II* gene in comparison with artiodactyls.

For the snakes, the rate of base substitution for pythons and colubrids was 0.21% and 0.41% per MY respectively for all substitutions and 0.06% and 0.16% per MY for transversions only. The only other directly fossil-calibrated estimate of the rate of *cytb* sequence divergence for all substitutions for reptiles is ~0.4% per MY for marine turtles (Bowen *et al.*, 1993). These estimates are consistent with other data supporting a slower rate of molecular evolution in heterotherms than in the homoeothermic placental mammals (Caccone *et al.*, 1997; Kocher *et al.*, 1989; Martin *et al.*, 1992; Rand, 1993; Rand, 1994; but see Seddon *et al.*, 1998).

The minimum estimated times of divergence between the north/south clades are 5.5 - 3.1 MYA for the snake species and ~ 1.7 MYA for *Melomys rufescens* (Table 7.1). It is possible that the ages of the fossil divergences used in the calibrations have been under-estimated. This is a significant source of error in fossil-based estimates of lineage age (Springer, 1995). Importantly, in terms of testing for a late Pleistocene effect on divergence,

underestimating lineage age would cause a higher estimate of the rate of base substitution, leading to an earlier estimate of the divergence date. Similarly, if the earlier molecular clock based dates for the *Mus/Rattus* divergence are used (i.e. 22 – 35 MA - Janke *et al.*, 1994; O'hUigin and Li, 1992; Sarich, 1985; Wilson *et al.*, 1977), the divergence of the north/south clades of *Melomys* is pushed back to 3.0 - 4.7 MYA.

7.4 Discussion

7.4.1 Biogeographic Interpretations

There are at least four major factors other than latest Pleistocene climate oscillations that could explain recent population separation and hence genetic substructuring in New Guinean lowland fauna. First, there has been a tremendous uplift of the central cordillera that began approximately 5.8 MYA in the late Miocene (Haig and Medd, 1996). The orogeny of the cordillera has been episodic rather than a continuously gradual event, with an initial major burst of uplift at 5.8-5.3 MYA and less intense episodes of folding and thrusting spanning 5.3-4.7 MYA (Haig and Medd, 1996). Evidence of the effect that the orogeny of the central cordillera has had on vicariance of lowland faunas is most dramatically demonstrated in a range of aquatic fauna. The New Guinea crocodile *Crocodylus novaeguineae* (Hall, 1989), the soft-shelled tortoise, *Pelochelys* (Webb, 1995) and rainbowfishes, *Melanotaenia affinis* and *M. goldei* (Allen and Cross, 1982), show putative paired sister taxon distributions across the central cordillera. Furthermore, the rainbowfish faunas of the northern and southern watersheds share just a single species out of a total of 39 species (Allen and Cross, 1982). Few opportunities exist for lowland rainforest or lower montane forest fauna to disperse across the cordillera as there are only a few mountain passes at about 1200-2000m elevation (Pratt, 1982). The assumption that lowland rainforest is continuously distributed around the island (Pratt, 1982; the present study) should also be re-examined.

Second, vicariance could be due to isolation of founders on the accreting elements of the "Outer Melanesian Arc". Polhemus (1998) points out that the "Outer Melanesian Arc" is

Table 7.1 Corrected sequence divergence and fossil calibration data used to estimate minimum divergence times for northern and southern clades in six New Guinean vertebrates. *a* - *e* are the values used to estimate evolutionary rates and divergence time for the major clades in each taxon.

		L. albertisii	*M. viridis*	*B irregularis*	*C. carinata*	*C. aspera*	*M. rufescens*
a	Av distance N vs S (A)	0.0069	0.0039	0.0131	0.0070	0.0067	0.0101
	Standard Error	0.0000	0.0002	0.0016	0.0000	0.0004	0.0006
b	Outgroup dist. from target taxon	0.0359	0.0316				0.0712
c	Min. fossil divergence time (MY)	25	25	16 - 22.5			12
d	ND4 Colubrid dist. Tv only			0.0712			
e	Python evol. rate (% bp/MY)				0.0006	0.0006	
e	Colubrid evol. rate x 16MY			0.0022			
e	Colubrid evol. rate x 22.5MY			0.0016			
	Min N/S diverg. time (MY) #	**4.8**	**3.1**				**1.7**
	Min N/S diverg. time (MY) +			**2.97 - 4.09**	**5.5**	**5.3**	

Python evol. rate (% bp/MY)	**Ts + Tv**	**.2% / MY**
Python evol. rate (% bp/MY) ^	**Tv only**	**.06% / MY**
Colubrid evol. rate x16MY(% bp/MY) ^	**Tv only**	**.22% / MY**
Colubrid evol. rate x22.5MY (% bp/MY) ^	**Tv only**	**.16% / MY**

\# (a/b)*c

\+ (a/2)/e

^ (d/2c)

in fact a series of at least four elements with dissimilar ages and origins. By the late Miocene (7-5 MYA), all of these elements except the eastern end of the Finisterre terrane had docked with the northern margin of the Australian craton. By 2 MYA the eastern end of the Finisterre terrane had also completely docked (Pigram and Davies, 1987). If vicariance on any of these elements was responsible for producing the diversification observed among the taxa that I have sampled, then it is difficult to reconcile the timing of the diversification from our molecular data (up to 5 MYA) with the proximity of the these "Outer Melanesian Arc" elements to the remainder of present day New Guinea at that time. Furthermore, if each taxon was present on these elements prior to their docking with the rest of the New Guinean landmass, it is difficult to conceive that diversification could have continued in the presence of a higher likelihood of dispersal as the elements approached more closely to the gradually growing New Guinean landmass.

Third, vicariance could be due to isolation of founders on various elements of the "Inner Melanesian Arc" and northern Australia. The present study does not rule out vicariance from northern Australia as an explanation for some species but it certainly does not explain the patterns for all of the species. There is no evidence that four of the taxa, the two *Candoia* species, *Leiopython albertisii* and *Melomys rufescens*, occurred in northern Australia and there are presently no paleo-environmental data available for this region that are pertinent to the time period involving the taxa studied here. Future phylogeographic studies of the several vertebrate species with distributions in northern Australia and throughout lowland New Guinea *e.g.* the marsupials *Echymipera rufescens* and *Spilocuscus maculatus*, the rodents *Hydromys chrysogaster* and *Uromys caudimaculatus* and the reptiles *Boiga irregularis*, *Morelia amethistina* and *M. viridis*, could be informative by incorporating analysis of both the Australian and New Guinean parts of their ranges.

Fourth, the earlier episodes of the Plio-Pleistocene climatic oscillations could have produced range fragmentation and subsequent divergence *in situ* within New Guinea. Whilst

the molecular clock estimates of intra-specific divergences for the snakes and rodents correlate with the 5-1 MYA period of orogeny of the central cordillera, they also overlap the Plio-Pleistocene climate oscillations starting at approximately 2.6 MYA (Webb and Bartlein, 1992). Palaeoclimatic records for southern Australia indicate climatic oscillations for the Pliocene comparable in magnitude to those of the Pleistocene (Truswell, 1993) but data for the Pliocene and earliest Pleistocene are not yet available for the New Guinean region which would have had both a more maritime and temperate to tropical climate. Climatic changes in the Late Pleistocene in New Guinea lowered both the upper and lower altitudinal boundaries of montane forests and caused expansion of grasslands at the expense of rainforest (Hope and Golson, 1995). The effectiveness of climatic oscillations in producing fragmentation of lowland rainforest and hill forest would have been exaggerated by the presence of the increasingly uplifted central cordillera. Indeed, Pratt (1982) predicts that the most important effect would have been to greatly restrict the distribution of hill forest.

The paraphyly of northern lineages of *Leiopython* and *Melomys rufescens* with their respective southern counterparts may indicate a dispersal pattern such as the "leading edge hypothesis" advocated by Hewitt (1996). Under this model, northern populations would have persisted and have been the source of long distance dispersants into the southern lowlands where the taxon was absent or had been extirpated.

Congruence of the phylogeographic structure between species strongly supports the hypothesis that the uplifting of the central cordillera has made a major contribution to distribution and divergence of lowland vertebrates in eastern New Guinea. As the central mountain range extends the length of the island, I predict that populations in western New Guinea would have been similarly affected. I also predict that a wider range of fauna and flora will show a similar distribution pattern, which may provide a biogeographic principal for defining conservation management units in the entire lowland New Guinean biota. Indeed allozyme electrophoretic studies of skinks and frogs also suggest distributions of sister taxa separated by the central cordillera (Donnellan and Aplin, 1989; Donnellan and Aplin, in

press). Mapping the precise areas of contact or overlap for each of the divergent taxon pairs will no doubt increase the level of understanding of the impact of orogeny and climate on the New Guinean biota. It may also help determine whether these divergent taxon pairs represent separate biological species through an examination of gene flow in areas where they may come into contact.

Appendix 7.1.

Aligned nucleotide sequences for partial *cytochrome b* gene of six New Guinea taxa.

Boiga irregularis= #, *B. cynodon*=*

```
                                                                                                                                      120
#AMS 122373   AAACTTCGGA TCCATACTAC TAACCTGCTT AATTACACAA ACTATGACTG GTTTCTTCCT GGCTATTCAC TACACAGCAA ACATTAACCT AGCTTTTTCA TCCATTATCC ACATCACCCG
#AMS 122368   .......... .......... ........?. ...A...... ..C....... .......... .......... C......... .......... .......... ......G... ..........
#AMS 122371   .......... ..k....... .......... .......... .......... .......... .......... .......... ......G... .......... .......... ..........
#AMS 115331   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..y....... .......... ..........
#AMS 115357   .......... .......... .......... .......... .......... .......... .......... ......C... .......... .......... .......... ..........
#AMS 124799   .......... .....G.... .......... ...C...... ..C..A..C. .w..T..... .......... .......... ....C..... ...C...... .......... ......y...
#AMS 124360   .......... .....G.... .......... ...C...... ..C..A..C. .G..T..... A......... .......... .......... .......... .......... .......T..
#BPBM 13758   ......y... .....G.... ...T...T.. ...C...... .....A..C. .G..T..... A......... .......... .......... .......... .......... .......T..
#AMS 135588   .......... .....G.... .......... ...C...... ..C..A..C. .A..T..... .......... .......... .......... .......... .......... ..........
#AMS 130424   ?????????? ..T....... .......... ..?....... ..C.....C. ....T..... .......... .......... .......... .......... .......... ..........
*WAM 101936   ......T..G ..T...T... ....T..... ...A.T.... ...G.A..A. .A........ .......... .......... .T.....T.. C..C...... ........T. ..........
*WAM 104558   ......T..G ..T....... .......... ...A.T.... .....A..A. .A........ r......... .......... .T.....T.. T..C...... .....CG.T. ..........

                                                                                                                                      240
#AMS 122373   AGATGTTCCC TACGGmTGAA TAATACAAAA TCTTCATGCA ATTGGCGCAT CAATGTTTTT TATCTGCATC TATATTCATA TTGCACGAGG ACTTTACTAC GGGTCTTACC TAAATAAAAA
#AMS 122368   .......... .T...C..G. .......... .......... .......... .......... C......... .....A.... .G.yr..... ...C...... .......... ..........
#AMS 122371   .......... .....C.... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
#AMS 115331   G......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
#AMS 115357   .......... .......... .......... .......... .......... .......y.. .......... .......... ...s...... .......... .......... ..........
#AMS 124799   ......C... .....m.... .......... .......... .....A.... ....A..... .......... .....C..C. .......... .......... ..A....... ..........
#AMS 124360   ....A..... .......... .......... .A........ .....A.... ....A..... .......... .....C..C. .......... .A........ ..A....... ..........
#BPBM 13758   ....A..... .....C.... .......... .A........ .....A.... ....A..... .......... .....C..C. .......... .......... ..A....... ..........
#AMS 135588   ......C... .....C.... .......... C......... .....A.... .T..A..... .......... ..C..C..C. .......... .......... ..A....... ..........
#AMS 130424   ......C... .....A.... .......... C......... .....A.... .T..A..... .......... .....C..C. .......... .......... ..A....... ..........
*WAM 101936   .........A ..TC.C.... .C........ CA.A...... .......... .C..A..... ...T...... ..C.CC.... .......... .....w.... .....C...T ..........
*WAM 104558   .........A ..T....... .C........ CA.A...... .......... .C..A..... ...y...... ..C..C..C. .......... .......... .....C...T ..........

                                                                             309
#AMS 122373   TGTCTGATTA TCAGGAACAA TTCTACTATT TATCCTCATA GCAACAGCCT TCTTTGGTTA TGTACTACC
#AMS 122368   .......... .......... ....G..... .????????? ?????????? ?????????? ?????????
#AMS 122371   .......... .......... .......... .......... .......... .......... ...C.....
#AMS 115331   .......... .......... .......... .......... .......... .......... ...m.....
#AMS 115357   .......... .......... .y........ .......... .......... .......... ...C.....
#AMS 124799   .......... .......... .......... .......... .......... .......A.. C..C.....
#AMS 124360   C......C.. .......... ....T..... ......T... .......... .......... C..C.....
#BPBM 13758   C......C.. .......... .......... ......T... .......... .......... C..C.....
#AMS 135588   .......... .......... .......... ......T... .......... .......G.. C..C.....
#AMS 130424   .......C.. .......... .......... .......... .......... ..G??????? ?????????
*WAM 101936   ...A...C.. .......... CC...T.... ...T..T... .......... C...C..... ...C.....
*WAM 104558   ...A...C.. .......... .C..TT.... .......... .......... .......... ...C.....
```

Candoia aspera =# (S= R1335507/8) and *C. carinata*=*

```
                                                                                                                                    120
*AMS R122349  TTTGGGTCAA TGCTGCTGAC CTGCTCAGCA TTACAAGTAA TAACCGGGTT TTTCTTATCA ATACACTACA CAGCAAATAT CAACCTAGCA TTTTCATCCA TTATCCACAT TACACGAGAC
*AMS R122346  .......... .......... .......A.. .......... .......A.. .......... .......... .......... .......... .......... .......... ..........
*AMS R124689  .....A.... .A........ ...T...... .......... .......A.. .......... .......... ....G..... .......... ..C....... .......... ..........
*AMS R124623  .....A.... .A........ ...T...... .......... .......A.. .......... .......... ....G..... .......... ..C....... .......... ..........
*AMS R135578  .....A.... .A........ ...T...... .......... .......A.. .......... .......... ....G..... .......... ..C....... .......... ..........
#AMS R115337  ..C..A.... .A..A..A.. ...T...... C......... ....A..... C......... .....T..T. .......... TG........ ..C.....T. .......... ..........
#AMS R124101  ..C..A.... .AT.A..A.. ...T...... .......... ....A..A.. C......... .....T.... .......... TG........ ..C.....T. .......T.. ..........
#SUE S        ..C..A.... .AT.A..A.. ...T...... .......... ....A..A.. C......... .....T.... .......... TG........ ..C.....T. .......T.. ..........
#AMS R122353  ..C..A.... .A..A..A.. ...T...... .......... ....A..A.. C......... .....T.... .......... TG....G... ..C.....T. .......... ..........
#AMS R122352  ..C..A.... .A..A..A.. ...T...... .......... ....A..A.. C......... .....T.... .......... TG....G... ..C.....T. .......... ..........

                                                                                                                                    240
*AMS R122349  GTCCCACACG GATGAATAAT ACAAAATCTG CACGCAATCG GGGCCTCCAT ATTTTTTATC TGTATTTACA TACACATCGC ACGAGGGTTA TACTACGGAT CCTACTTAAA TAAAGAGACA
*AMS R122346  .......... .......... .......... .......... .......... ...y...... .....y.... .......... .......... .......... .......... .w........
*AMS R124689  .....C.... .C........ ......C... ........T. .......... G......... .......... .......... ......A... .......... .......... ..........
*AMS R124623  .....C.... .C........ ......C... ........T. .......... G......... .......... .......... ......A... .......... .......... ..........
*AMS R135578  .....C.... .C........ ......C... ........T. .......... G......... .......... .......... .......... .......... .......... ..........
#AMS R115337  ........T. .......... .........A .....T.... .A..A..... ......C... .....C..T. C......T.. ......A... ........C. .A........ C...A.....
#AMS R124101  ........T. .......... ......C..A .....C.... .A..A..T.. .......... .....C.... C......T.. ......AC.. ........C. .G...C.... C.........
#SUE S        ..T.....T. .......... ......C..A .....C.... .A..A..T.. .......... .....C.... C......T.. ......AC.. ........C. .G........ C.........
#AMS R122353  ........T. .......... .........A .....C.... .A..A..T.. G.....C... .....C.... C......T.. ........... .T.....C. .A........ C...A.A...
#AMS R122352  ........T. .......... .........A .....C.... .A..A..T.. G.....C... .....C.... C......T.. ......A... ..T.....C. .A........ C...A.A...

                                                                                309
*AMS R122349  TGACTATCTG GCACCACACT ACTAATCATA CTAATAGCGA CCGCATTCTT CGGATACkTA CTACCGTGA
*AMS R122346  .......T.. .......m.. .......... .......... ....m..... .......G.m ..m.??...
*AMS R124689  .......... .......... .......... .......... .......... .......G.. .....A...
*AMS R124623  .......... .......... .......... m......... ........y. .......G.. T....A...
*AMS R135578  .......... .......... .......... .......... .......... .......G.. .....A...
#AMS R115337  ...T....G. ....G..... .......... ..T.....A. .T........ T......G.. .....A.AG
#AMS R124101  ........A. .......... .......... ..T.....A. .......... ......TG.. .....A...
#SUE S        ........A. .......... .......... ..T.....A. .......... ......TG.. .....A...
#AMS R122353  .......... ....G..... .......... ..T.....A. .......... .......G.. .....A...
#AMS R122352  .......... ....G..... .......... ..T.....A. .......... .......G.. .....A...
```

Morelia viridis=*, (^=AMS R115348/49/50; +=AMS R115355/6; <=BPBM 11617/BPBM 13798, AMS R124481) and *Leiopython albertisii*= # (†=BPBM 11612/3)

```
                                                                                                                          120
*DOI ^         TTCGGCTCAA TACTATTAAC ATGCCTAGCC CTACAAGTAT TAACCGGCTT CTTCCTAGCC GTTCACTACA CAGCAAACAT TAATCTAGCA TTCTCATCCA TCATCCACAT CTCCCGAGAT
*NRU +         ..s.sT.... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
*AMS R122364   .......... .......... .......... .......... .......... .......... .......... ...s...... ...C...... .......... .......... ..........
*AMS R122363   .......... .......... .......... .......... .......... .......... .......... ...s...... ...C...... .......... .......... ..........
*AMS R129716   .......... .......... .......... .......... .......... ....T..... ..A....... ...s...... .G........ .......... .......... .........C
*WAU <         .......... ..T....... ...TT..... .........C .......... ....T..... ..C....... .......... C..C...... .......... .T.....T.. .A.T......
AMS R124531    .....y.... ..T....... ...TT..... .........C .......... ....T..... ..C....... .......... C..C...... .......... .T.....T.. .A.T......
#MADKAB        ..T....... .......... .......k.T ......A... .......... ......G... ..C....... .......... C..C...... .......... .......... .A.......C
#AMS R115346   ..T..T.... .......... ...TT..... .........C .......... .......... ..C....... .......... C..C...... ..T....... ....T..... TA........
#WAU †         ..T..T.... .......... ....T....T .....G...C .......... .......... ..C....... .......... C..C...... ..T....... .......... TA........

                                                                                                                          240
*DOI ^         GTTCCATACG GTTGAATAAT ACAAAACCTA CACGCCATCG GAGCATCCAT ATTCTTCATT TGCATCTACA TCCATATTGC ACGAGGATTA TACTATGGAT CCTACCTCAA CAAAGAAACC
*NRU +         .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
*AMS R122364   .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
*AMS R122363   .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
*AMS R129716   .......... .......... .........G .......... .......... .........C .......... .........s .......C.. .......... .......... ..........
*WAU <         ..C....... .C........ .......... .......... .......... .......... .....T.... ....C..C.. .......C.. .....C.... .......... ......G..T
*AMS R124531   ..C....... .m........ .......... .......... .......... .......... .....T.... ....C..C.. .......C.. .....C.... .......... ......G..T
#MADKAB        ..C....TT. .......... .......... .......... .......... .........C .......... .T..C..C.. .......C.G .....C.... .A.....A.. ..........
#AMS R115346   ..C....T.. .C........ .......... .......... .......... .......... .......... .T..C..C.. .......C.C .....C.... .A.....A.. ..........
#WAU †         ..C....T.. .C........ .......... .......... .......... .......... .......... .T..C..C.. .......C.C .....C.... .A.....A.. ..........

                                                       309
*DOI ^         TGAATATCCG GTATTACCCT ACTCATCACA CTAATAGCAA CCGCCTTCT
*NRU +         .......... .......... .......... .......... .........
*AMS R122364   .......... .......... .......... .......?.. ...T.....
*AMS R122363   .......... .......... .......... .......... .........
*AMS R129716   .......... ..G....... G......... .......... .........
*WAU <         .......... ....C..... .......... T....G.... .........
*AMS R124531   .......... ....C..... .......... T......... .....w...
#MADKAB        ..G....... .C........ .......... .....G.... ....A....
#AMS R115346   .....G.... .C........ .......... ..G....... .........
#WAU †         .....G.... .C........ .......... ..G....... .........
```

Melomys rufescens (▲=AMS M19859/AMS M16248; *=*M. cervinipes*)

```
                                                                                                                                120
WARNAM▲       TTTGGCTCTy TTTTAGGACT ATGCCTTACA ATTCAAATCA TCACAGGACT ATTyTTAGCC ATACATTATA CATCAGATAC AATAACAGCA TTTTCATCAG TTACACACAT TTGCCGAGAC
AMS M14732    .........C ........m. .......... .......... .......... .......... ........C. .......... .......... ..y....... .......... ..........
AMS M15162    .........C .......... .......... .......... .......r.. ...C...... ........C. .......... .......... ..C....... .......... ..........
AMS M12651    .......y.T .....s.... .......... .......... .......... .......... .......... .......... .......... .........r ..G....... ..........
AMS M15955    ..C.....CC .......... G.......T. ........T. ........T. .........T .........T .......... .......... .......... .C....T... ...T......
AMS M15956    ..C.....CC .......... ........T. ........T. .T........ .........T .......... .......... .......... .......... .C....T... ...T......
AMS M21677    .........C .......... ......C.T. .......... .......... .......... .....C.... .......... .......... .......... .C..T..... ..........
AMS M21678    .......... ........m. ......C.T. .......... .......... .......... .....C.... .......... ...C...... .......... .C..T..... ........r.
QM  M10914*   .........C .C........ T.......T. .......... .T........ ...CC....T .......... .......C.. ..C....... .......... .C........ ..........
Mus           .....G...C .GC.....A. C.....A.T. G.C......C .......C.. ...C.....T .....C..C. .......C.. .........C ..C....... .A....T... C........T
Rattus        ..C..T...C .AC.....G. ......C.T. G.A......C .......CT. ...CC....A .....C..C. .G..T..... C......... ..C....... .C..C..... C.........

                                                                                                                                240
WARNAM▲       GTAAATTACG GATGACTAAT TCGATATATA CACGCAAACG GAGCATCAAT ATTCTTCATC TGCTTATTCC TCCATGTAGG ACGAGGAATA TATTATGGAT CCTACACCTT CATAGAAACA
AMS M14732    .......... .......G.. ..r....... .......... .......... ...T...... .......... .......... .......... .......... .......... ..........
AMS M15162    .......... .......G.. .......... .......... .......... ...T...... .......... .......... .......... .......... .......... ..........
AMS M12651    .......... .......G.. .......... .......... .......... ...T...... .......... .......... .......... .......... .......... ..........
AMS M15955    .......... .......... C......... .......... ........C. ...T...... ........T. .......... .......... .....C.... .T..T..A.. ..........
AMS M15956    .....C.... .......... C......... .......... ........C. ...T...... .......... .......... .......... .....C.... .T..T..A.. ..........
AMS M21677    ..T....... .....T.... C......... .....T.... .......G.. G..T...... .......... .......... .......... ..C..C.... .T........ ...G......
AMS M21678    ..T....... .......... C......... .....T.... .......... ...T...... .......... .......... .......... ..C..C.... .T........ T..?......
QM  M10914*   .....C.... ....GT.... .......... .......... .......... ...T...... ...C...... ....C..... C.....C... .......... .T..T..... T.........
Mus           .....C..T. .....T.... C......T.. .....C.... ....C..C.. .......T.. ...C.....T .A..C..... ...G..G..G ..C..C.... .T..T..A.. ..........
Rattus        .....C.... .C........ C.....CC.. ..A..C.... .C..C..... ...T...... ...C...... .......G.. .......C.. ..C....... .......T.. .C.......C

                                                                     306
WARNAM▲       TGAAACATTG GAGTCATCCT ACTATTTGCC GTAATAGCAA CCGCATTTAT AGGCTACGTC CTCCCA
AMS M14732    .....T.... .......... ....C..... .......... .......... .......... ......
AMS M15162    .......... .r........ ....C..... .......... .......... .......... ......
AMS M12651    .....T.... .......... .......... .......... .......... .......... ......
AMS M15955    .....T.... .......Ty. .......... ......k... .......C.. .......... ..T...
AMS M15956    .....T.... .......T.. .........T .......... .......C.. .......... ..T...
AMS M21677    .....T..C. ....T...T. .......A.. .......... .......C.. .......... ..T...
AMS M21678    .....T..C. ....T...T. .......A.. .......... .......C.. .......... ..T...
QM  M10914*   .....T.... .......... .T........ .......... .T.y...C.. .........T ......
Mus           .....T.... ....T..T.. .T.......A ........C. .T........ ...T..T..T ..T...
Rattus        .......... .GA....... .........A ..C....... .T.....C.. G.....T..A ......
```

Appendix 7.2.

Allozyme frequencies for five New Guinea taxa. Frequencies for *Boiga irregularis* are from L. Rawlings' unpublished honours thesis, frequencies for *Candoia aspera*, *C. carinata*, *Leiopython albertisii* and *Morelia viridis* were tested by S. Donnellan.
Boiga irregularis The following loci were invariant: *Aat-1, Acp, Ca, Gpdh, Guk, Idh-1, Lap, Ldh-2, Mdh-1, Mdh-2 ,Pgam, Pgk, Sod* and *Tpi*.

Locus	all	N	BcS 2	BcF 2	CHP 3	SHP 6	WSP1 1	WSP2 1	MAP 2	MP 2
Aat-2	b		100	100				100	100	100
	a				100	100	100			
Acoh	b				100	92	100	100	100	100
	a		100	100		8				
Ada	c				100	100	50	100	100	100
	b		100	100						
	a						50			
Alat	c		100							
	b			100				100	75	
	a				100	100	100		25	100
Aldh	b			100	100	100	-	-	100	-
	a		100							
Ap	c		100	100	67	60	100	100	100	100
	b					20				
	a				33	20				
Eno	b		100	100						
	a				100	100	100	100	100	100
Est-1	b							50		
	a		100	100	100	100	100	50	100	100
Est-2	c						-		50	
	b		100	100	83	50		100	50	100
	a				17	50				
Fbp	c		100	100	67	50	100	100	100	100
	b					8				
	a				33	42				
Fumh	b		100				-			
	a			100	100	100		100	100	100

Locus	all	N	BcS 2	BcF 2	CHP 3	SHP 6	WSP1 1	WSP2 1	MAP 2	MP 2
Gpi	b				17	10				
	a		100	100	83	90	100	100	100	100
Idh-2	b		100	100	100	100	50	100	100	100
	a						50			
Ldh-1	b				100	100	100	100	100	100
	a		100	100						
Lgl	b		25	100	100	100	100	100	100	100
	a		75							
Mpi	b				100	100	100	50	100	100
	a		100	100				50		
Pnp	b						100	100	100	
	a		100	100	100	100				100
PepA	b		100	100						
	a				100	100	100	100	100	100
PepB	c		100	75						
	b			25	100	90	100	100	100	100
	a					10				
PepC	c		100	100						
	b						100	100	100	100
	a				100	100				
PepD	c				33					
	b					8		50		
	a		100	100	67	92	100	50	100	100
Pgdh	b		75	100	100	100	100	100	100	100
	a		25							
Pgm	c									50
	b				100	100	100	100	100	50
	a		100	100						

Candoia aspera and *C. carinata*

The following loci were invariant: *Aat-1, Ada, Ca, Fbp, Fumh, Gda, G3pdh, Idh, Lap, Ldh-1, Ldh-2, Ndpk, PepA* and *Pgam.*

		C. aspera							*C. carinata*			
LOCUS	N	Mad1 3	Usi 2	Kar 2	Sue1 4	Nam 2	Nru 1	Yur 2	Sue2 1	Mad2 1	Fog 2	War 2
Aat-2	b	100	100	50	100	100	100	100				
	a			50					100	100	100	100
Acoh-1	b	100	100	50	100	100	100	100		-		
	a			50					100		100	100
Acoh-2	d						50			-		
	c					100	50	100				
	b	67	100	100	100				100		100	100
	a	33										
Ak-1	b	100	100	50	100	100	100	100	100		100	100
	a			50						100		
Eno	c	100	100	100	100	25	100	75	100	100	100	100
	b							25				
	a					75						
Gpi	c	50			25							
	b	50	100	50	75	100	100	100				
	a			50					100	100	100	100
Gsr	b		25	25	87	50		50	100		-	75
	a	100	75	75	13	50	100	50		100		25
Iddh	b	50	50									
	a	50	50	100	100	100	100	100	100	100	100	100

		C. aspera							*C. carinata*			
LOCUS	N	Mad1 3	Usi 2	Kar 2	Sue1 4	Nam 2	Nru 1	Yur 2	Sue2 1	Mad2 1	Fog 2	War 2
Lgl	b	100	100	50	100	100	100	100				
	a			50					100	100	100	100
Mpi	d			50					100	100	100	100
	c					100	100	100				
	b	50	100	50	87							
	a	50			13							
PepB	d	17										
	c	83	100	50	100	75			50			
	b					25	100	100			100	100
	a			50					50	100		
PepD	c	17	50	50	87							
	b					25						
	a	83	50	50	13	75	100	100	100	100	100	100
Pgdh	c	100	50	50	100	100	100	25				
	b							75				
	a		50	50					100	100	100	100
Pgm	b				12						50	75
	a	100	100	100	88	100	100	100	100	100	50	25
Tpi	b	100	100	50	100	100	100	100				
	a			50					100	100	100	100

Morelia viridis and *Leiopython albertisii*

The following loci were invariant: *Acoh-1, Acp, Ak-2, Ca, Fbp, Fumh, Idh-2, Lap, Ldh-1, Ldh-2, Mdh-1, Ndpk, Pgam, Pgk* and *Pgdh.*

		Morelia viridis					*Leiopython albertisii*			
Locus	N	Wau 2	Nor 1	Nam 2	Doi 3	Nru 2	Doi 1	Wau 2	Mad 2	Kab 1
Aat-1	b	100	100	100	100	100				
	a						100	100	100	100
Aat-2	c		100	100	100	100	100	100		
	b								100	100
	a	100								
Acoh-2	c	75								
	b		100	100	100	75	100	100	100	100
	a	25				25				
Ada	c		100	100	100	100				
	b	100						25		
	a						100	75	100	100
Ak-1	b						100	100	100	100
	a	100	100	100	100	100				
Eno	b	100	100	100	17	75	100	100	100	100
	a				83	25				
Est	b	100	100	100	100	100	50	100	100	100
	a						50			
Gda	b	100	100	100	100	100				
	a						100	100	100	100
Gpi	c					25				
	b	100	100	100	100	75				
	a						100	100	100	100
Gsr	b	50	50	100	17					
	a	50	50		83	100	100	100	100	100
Iddh	b						100	100	100	100
	a	100	100	100	100	100				

		Morelia viridis					*Leiopython albertisii*			
Locus	N	Wau 2	Nor 1	Nam 2	Doi 3	Nru 2	Doi 1	Wau 2	Mad 2	Kab 1
Idh-1	c	25								
	b	75	50				100	100	100	100
	a		50	100	100	100				
Lgl	b						100	25	100	100
	a	100	100	100	100	100		75		
Mdh-2	b	100	100	100	100	100	100	100	75	100
	a								25	
Mpi	c	100	100	100	100	100				
	b								100	100
	a						100	100		
PepA	c	75	100	100	100	100		100	100	100
	b						100			
	a	25								
PepB1	c	100					100	100	100	100
	b		50	50	83	50				
	a		50	50	17	50				
PepB2	b	100								
	a		100	100	100	100	100	100	100	100
PepD	b						50			
	a	100	100	100	100	100	50	100	100	100
Pgm-1	c							25		
	b	25		25	83	25	100	75	100	100
	a	75	100	75	17	75				
Pgm-2	b		100	100	100	100	100	100	100	100
	a	100								

Conclusion

The scope of the molecular studies of pythons covered in the present study has been very wide, ranging from intra-individual to inter-genera. Pythons share a mitochondrial gene duplication of the *control region* that has been found only in other families of snakes. However, some species of pythons, (particularly the Australo-Papuan and eastern Asian species), have some additional elements such as hairpins and a paralogous, partial tRNA sequence in the *control region* that are not found in viperid, colubrid or boid snakes studied to date. Likewise, there are a number of large tandem repeat sequences in the python *control region* that were not present in other snakes. Repeat sequences were found in the boid snake *Candoia aspera* and further investigation of snake *control regions* might find much more variation than originally thought. Concerted evolution of *control region* sequences has now found in four snake families (boids, colubrids, viperids and file snakes) and the duplication is also present in elapids (*Acanthophis*). This phenomenon would be interesting to pursue to discover the mechanism involved.

Fourteen of fifteen genera of pythons are found in Australia and New Guinea and most of the focus of this study was on the Australo-Papuan species. African and Asian species of the genus *Python* were included to assess the genetic relationships within the family Pythonidae. For completeness, it would have been useful to include *Morelia carinata* and *Python anchiete* in the python phylogeny. Future studies could investigate the genetic relationships within genera such as *Morelia* or *Python* to fill in the gaps left by the present study. The species complex of *Antaresia* could also be investigated further. A 15 bp indel present in *A. childreni/stimsoni* and absent in *A. perthensis/maculosa* suggests a common ancestry for *A. childreni/stimsoni* after the divergence of *A. perthensis* and *A. maculosa*.

In order to extend the findings of this study of phylogenetic relationships of pythons beyond a gene tree, the addition of nuclear gene sequences would be advantageous. The challenge will be to find a nuclear gene that has a rate of sequence evolution sufficient to

contain phylogenetic signal that will reflect the evolutionary history of the Pythonidae. The rate of molecular evolution of the mitochondrial genes is quite slow relative to mammals and even the medium-fast evolving *cytochrome b* gene did not provide sufficient signal on its own to differentiate between python lineages.

Control region and *cytochrome b* sequences were able to differentiate between distinct populations of green pythons (*Morelia viridis*). The present study found strong evidence to suggest that there are two cryptic species of green pythons distributed either side of the New Guinea central mountain range. The sampling of green pythons for this study was not sufficiently comprehensive to discern the geographic boundaries of the two species. Further investigation of the geographic boundaries for the "northern" and "southern" forms of green pythons could be undertaken by sampling from the coastal regions at the eastern and western ends of the central mountain range. As a pattern of "north-south" genetic sub-structuring is also present in other mid-montane New Guinea taxa (presented in chapter 7), delineation of the green python species boundaries may give a clue to zoo-geographic regions that became established as the mountain building of the Plio-Pleistocene divided faunal populations.

From this very preliminary data, it would seem that phylogenetic signal within the *control region* of green pythons is sufficient to differentiate between the Australian population and the New Guinea populations. Further testing is necessary to determine the amount of variation present in the New Guinea populations and to test whether the Australian haplotypes are also present in New Guinea. This population data would have a useful application in customs and fauna protection to provenance suspected illegally-imported animals. As the poaching of native specimens from the wild and the illegal import of exotic specimens is an on-going fauna management problem for Australian Parks and Wildlife officers, further molecular research on pythons should incorporate genetic markers that can be used to identify individuals that are captive bred and distinguish between wild populations (*e.g.* microsatellites).

Control region sequences have also been useful to investigate Australian populations

of water pythons. Results from the present study suggested that there might be two distinct populations, a northern territory group and an east coast population. These findings need to be confirmed further and a cost-effective, first place to start would be with allozymes. Genetic analysis supported the current division of Liasis olivaceus into two subspecies, the Pilbara distributed *L. o. barroni* and the northern Australian *L. o. olivaceus.*

There are many other python population studies that could be pursued. As many species have wide distributions that incorporate quite diverse habitats, some of which could have existed in isolation for significant periods, there were likely to have been opportunities for genetic sub-division to occur. In the pursuit of preserving genetic diversity, it is important to identify unique populations. Population studies of *Aspidites ramsayi, Morelia amethistina* and *M. spilota* would each be interesting to pursue for the above reasons. As these are also attractive species that are popular in the domestic market, investigations that lead to the development of molecular markers would also benefit the monitoring of the pet trade.

It is hoped that the findings presented in this study will be built on in future investigations into python population genetics.

References

Allen G.R. and Cross N.J. (1982) *Rainbowfishes of Australia and Papua New Guinea.* Angus and Robertson Publishers, Australia.

Anderson S., De Bruijn M.H.L., Coulson A.R., Eperon I.C., Sanger F. and Young I.G. (1982) Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *Journal of Molecular Biology,* **156**: 683-717.

Arctander P. (1995) Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proceedings of the Royal Society London B,* **262**: 13-19.

Austin C.C. (2000) Molecular phylogeny and historical biogeography of Pacific island boas (*Candoia*). *Copeia,* **2**: 341-352.

Avise J.C. (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos,* **63**: 62-76.

Avise J.C. (1994) *Molecular Markers, Natural History and Evolution.* Chapman & Hall, New York, USA.

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

Avise J.C. and Ball R.M. (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology,* **7**: 45-67.

Avise J.C., Ball R.M. and Arnold J. (1988) Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution,* **5**: 331-344.

Axelrod D.I. and Raven P.H. (1982) Paleobiogeography and origin of the New Guinea flora. In: *Biogeography and Ecology of New Guinea,* (ed. Gresset J.L.) pp. 919-941 W. Junk, The Hague.

Ayliffe M.A. and Timmis J.N. (1992) Tobacco nuclear DNA contains long tracts of homology to chloroplast DNA. *Theoretical and Applied Genetics,* **85**: 229-238.

Banks C. (1998) *Action plan for reptiles and amphibians in Australasian zoos (Draft).*

Banks C.B. and Meikle W. (1994) *ASMP Taxon Advisory Group Plan for Reptiles and Amphibians in Australasian Zoos.* Australasian Regional Association of Zoological Parks and Aquaria Inc., Mosman, Aus.

Barker D.G. and Barker T.M. (1994) *Pythons of the World: Australia.* Advanced Vivarium Systems Inc Ca.

Bensch S. and Härlid A. (2000) Mitochondrial genomic rearrangements in songbirds. *Molecular Biology and Evolution,* **17**: 107-113.

Berg T., Moum T. and Johansen S. (1995) Variable numbers of simple tandem repeats make birds of the order Ciconiiformes heteroplasmic in their mitochondrial genomes. *Current Genetics,* **27**: 257-262.

Birky Jr C.W., Fuerst P. and Maruyama T. (1989) Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics,* **121**: 613-627.

Boulenger G.A. (1893) *Catalogue of Snakes in the British Museum (Natural History).* British Museum, London.

Bowen B.W., Nelson W.S. and Avise J.C. (1993) A molecular phylogeny for marine turtles: trait mapping, rate assessment and conservation relevance. *Proceedings of the National Academy of Science USA,* **90**: 5574-5577.

Bremer B. (1994) Branch support and tree stability. *Cladistics,* **10**: 295-304.

Bromham L., Phillips M.J. and Penny D. (1999) Growing up with dinosaurs: molecular dates and the mammalian radiation. *Trends in Ecology and Evolution,* **14**: 113-118.

Brongersma L.D. (1956) On two species of boid snakes from the Lesser Sunda Islands. **59**: 290-300.

Brower A.V.Z., DeSalle R. and Vogler A. (1996) Gene trees, species trees and systematics: a cladistic perspective. *Annual Review of Ecology and Systematics*, **27**: 423-450.

Brown G.G. (1986) Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *Journal of Molecular Biology*, **192**: 503-511.

Brown J.R., Beckenbach A.T. and Smith M.J. (1992) Mitochondrial DNA length variation and heteroplasmy in populations of white sturgeon (*Acipenser transmotanus*). *Genetics*, **132**: 221-228.

Brown W.M., Prager E.M., Wang A. and Wilson A.C. (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution*, **18**: 225-239.

Bull J.J., Huelsenbeck J.P., Cunningham C.W., Swofford D.L. and Waddell P.J. (1993) Partitioning and combining data in phylogenetic analysis. *Systematic Biology*, **42**: 384-397.

Bulmer M., Wolfe K.H. and Sharp P.M. (1991) Synonymous nucleotide substitution rates in mammalian genes: implications for the molecular clock and the relationship of mammalian orders. *Proceedings of the National Academy of Science USA*, **88**: 5974-5978.

Buroker N.E., Brown J.R., Gilbert T.A., O'Hara P.J., Beckenbach A.T., Thomas W.K. and Smith M.J. (1990) Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. *Genetics*, **124**: 157-163.

Caccone A., Milinkovitch M.C., Sbordoni V. and Powell J.R. (1997) Mitochondrial DNA rates and biogeography in European newts (Genus *Euproctus*). *Systematic Biology*, **46**: 126-144.

Caldwell M.W. (1998) Squamate phylogeny and the affinities of snakes. *Zoological Journal of the Linnean Society*, **125**: 115-147.

Cantatore P. and Saccone C. (1987) Organisation, structure and evolution of mammalian mitochondrial genes. *International Review of Cytology,* **108**: 149-202.

Catzeflis F.M., Aquilar J.-P. and Jaeger J.-J. (1992) Muroid rodents: phylogeny and evolution. *Trends in Ecology and Evolution,* **7**: 122-127.

Chenoweth S.F., Hughes J.M., Keenan C.P. and Lavery S. (1998) When oceans meet: a Teleost shows secondary intergradation at an Indian-Pacific interface. *Proceedings of the Royal Society of London - Series B,* **265**: 415-420.

Christianson T.W. and Clayton D.A. (1988) A tridecamer DNA sequence supports human mitochondrial RNA 3'-end formation *in vitro*. *Molecular and Cellular Biology,* **8**: 4502-4509.

Clayton D.A. (1992) Transcription and replication of animal mitochondrial DNAs. *International Review of Cytology,* **141**: 217-232.

Cogger H.G. (1996) *Reptiles and Amphibians of Australia*. Reed Books, Melbourne, Australia.

Cogger H.G., Cameron E.E. and Cogger H.M. (1983) *Zoological Catalogue of Australia.* Australian Government Publishing Service, Canberra.

Colgan D., Flannery T.F., Trimble J. and Aplin K. (1993) Electrophoretic and morphological analysis of the systematics of the *Phalanger orientalis* (Marsupalia) species complex in Papua New Guinea and the Solomon Islands. *Australian Journal of Zoology,* **41**: 355-378.

Collura R.V. and Stewart C. (1995) Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominoids. *Nature,* **378**: 485-489.

Cooper A. and Penny D. (1997) Mass survival of birds across the cretaceous tertiary boundary: molecular evidence. *Science,* **275**: 1109-1112.

Cracraft J. (1983) Species concepts and speciation analysis. *Current Ornithology,* **1**: 159-187.

Crozier R.H. (1990) From population genetics to phylogeny: uses and limits of mitochondrial DNA. *Australian Systematic Botany,* **3**: 111-124.

Cunningham C.W. (1997) Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution,* **14**: 733-740.

Cunningham C.W., Blackstone N.W. and Buss L.W. (1992) Evolution of king crabs from hermit crab ancestors. *Nature,* **355**: 539-542.

da Silva M.N.F. and Patton J.L. (1998) Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Molecular Ecology,* 7: 475-486.

de Queiroz A. (1993) For consensus (sometimes). *Systematic Biology,* **42**: 368-372.

de Queiroz A., Donoghue M.J. and Kim J. (1995) Separate versus combined analysis of phylogenetic evidence. *Annual Review of Ecology and Systematics,* **26**: 657-681.

de Rooij N. (1917) *The reptiles of the Indo-Australian archipelago II Ophidia.* Brill, E. J., Leiden.

deGrouchy J. and Turleau C. (1984) *Clinical Atlas of Human Chromosomes.* Wiley, New York.

Densmore III L.D., Rose F.L. and Kain S.J. (1992) Mitochondrial DNA evolution and speciation in water snakes (Genus *Nerodia*) with special reference to *Nerodia harteri. Herpetologica,* **48**: 60-68.

Desjardin P. and Morais R. (1990) Sequence and gene organization of the chicken mitochondrial genome. *Journal of Molecular Biology,* **212**: 599-634.

Dobzhansky T. (1937) *Genetics and the origin of species.* Columbia University Press, New York.

Doda J.N., Wright D.T. and Clayton D.A. (1981) Elongation displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proceedings of the National Academy of Science USA,* **78**: 6116-6120.

Donnellan S.C. and Aplin K.P. (1989) Resolution of cryptic species in the New Guinean lizard *Sphenomorphus jobiensis* (Scincidae) by electrophoresis. *Copeia,* **1989**: 81-88.

Donnellan S.C. and Aplin K.P. (in press) Species boundaries in the *Rana arfaki* species group (Anura: Ranidae) and phylogenetic relationships to other New Guinean ranid frogs. *Records of the Australian Museum,*

Donnellan S.C., Hutchinson M.N. and Saint K.M. (1999) Molecular evidence for the phylogeny of Australian gekkonoid lizards. *Biological Journal of the Linnean Society,* **67**: 97-118.

Dow D.B. (1977). Geological synthesis of Papua New Guinea, *Geological synthesis of Papua New Guinea.*

Dowling T.E., Moritz C., Palmer J.D. and Reiseberg L.H. (1996) Nucleic acids III: analysis of fragments and restriction sites. In: *Molecular Systematies,* 2nd ed. (eds. Hillis D.M., Moritz C. and Mable B.K.) pp. 249-320 Sinauer Associates, Inc., USA.

Dujon B. (1981) Mitochondrial genetics and functions. In: *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance,* (eds. Strathern J.N., Jones E.W. and Broach J.R.) pp. 505-635 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Eriksson T. (1997) *Autodecay. 2.9.9.* Botaniska institutionen, Stockholm University.

Faber J.E. and Stepien C.A. (1998) Tandemly repeated sequences in the mitochondrial DNA control region and phylogeography of the pike-perches *Stizostedion. Molecular Phylogenetics And Evolution,* **10**: 310-322.

Farris J.S., Kallersjo M., Kluge A.G. and Bult C. (1995) Testing significance of incongruence. *Cladistics,* **10**: 315-319.

Felsenstein J. (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution,* **17**: 368-376.

Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution,* **39**: 783-791.

Felsenstein J. (1992) Estimating effective population size from samples of sequences: inefficiency of pair-wise and segregating sites as compared to phylogenetic estimates. *Genetical Research,* **59**: 139-147.

Fitch W.M. (1970) Distinguishing homologous from analogous proteins. *Systematic Zoology,* **19**: 99-113.

Flannery T. (1995) *The Mammals of New Guinea*. Robert Brown and Associates, Qld, Australia.

Flannery T.F., Colgan D. and Trimble J. (1994) A new species of *Melomys* from Manus Island, Papua New Guinea, with notes on the systematics of the *M. rufescens* complex (Muridae: Rodentia). *Proceedings of the Linnean Society of NSW,* **114**: 29-43.

Fyfe G. (1990) Notes on the central carpet python *Morelia spilota bredli*. *Herpetofauna,* **20**: 11-14.

Galloway R.W. and Löffler E. (1972) Aspects of geomorphology and soils in the Torres Strait region. In: *Bridge and Barrier: the Natural and Cultural History of Torres Strait,* (ed. Walker D.) pp. 11-28 Australian National University, Canberra.

Glickman B.W. and Ripley L.S. (1984) Structural intermediates of deletion mutagenesis: a role for palindromic DNA. *Proceedings of the National Academy of Science USA,* **81**: 512-516.

Goldberg T.L. and Ruvolo M. (1997) The geographic apportionment of mitochondrial genetic diversity in East African chimpanzees, *Pan troglodytes schweinfurthii*. *Molecular Biology and Evolution,* **14**: 976-984.

Goodwin R.E. (1979) The bats of Timor: systematics and ecology. *Bulletin of the American Museum of Natural History,* **163**: 73-122.

Greer A.E. (1979) A new species of *Lerista* (Lacertilia: Scincidae) from northern Queensland, with remarks on the origin of the genus. *Records of the Australian Museum*, **32**: 383-388.

Greer A.E. (1985) The relationships of the lizard genera *Anelytropsis* and *Dibamus*. *Journal of Herpetology*, **19**: 116-156.

Groves C.P. (1990) The centrifugal pattern of speciation in Meganesian rainforest mammals. *Memoirs of the Queensland Museum*, **28**: 325-328.

Haffer J. (1969) Speciation in Amazonian forest birds. *Science*, **165**: 131-137.

Haig D.W. and Medd D. (1996) Latest Miocene to early Pliocene bathymetric cycles related to tectonism, Puri anticline, Papuan basin, Paupa New Guinea. *Australian Journal of Earth Sciences*, **43**: 451-465.

Hall P.M. (1989) Variation in geographic isolates of the New Guinea Crocodile (*Crocodylus novaeguineae* Schmidt) compared with the similar, allopatric, Philippine Crocodile (*C. mindorensis* Schmidt). *Copeia*, **1989**: 71-80.

Harpending H.C. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**: 591-600.

Harpending H.C., Sherry S.T., Rogers A.R. and Stoneking M. (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**: 483-496.

Hartl D. and Clark A. (1997) *Principles of population genetics*. Sinauer Associates, Inc., Sunderland, Mass.

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

Heaney L.R. (1991) A synopsis of climatic and vegetational change to southeast Asia. *Climatic Change*, **19**: 53-61.

Hedges S.B. and Poling L.L. (1999) A molecular phylogeny of reptiles. *Science*, **283**: 998-1001.

Hess J.F., Parisi M.A., Bennett J.L. and Clayton D.A. (1991) Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature*, **351**: 236-239.

Hewitt G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**: 247-276.

Hill K.C. and Gleadow A.J.W. (1989) Uplift and thermal history of the Papuan Fold Belt, Papua New Guinea: Apatite fission track analysis. *Australian Journal of Earth Sciences*, **36**: 515-539.

Hill K.C., Grey A., Foster D. and Barrett R. (1993) An alternative model for the Oligo-Miocene evolution of northern PNG and the Sepik-Ramu basins. In: *Petroleum Exploration and Development in Papua New Guinea: Proceedings of the Second PNG Petroleum Convention*, (eds. Carman G.J. and Carman Z.) pp. 241-259 , Port Moresby.

Hillis D.M. (1991) Discriminating between phylogenetic signal and random noise in DNA sequences. In: *Phylogenetic analysis of DNA sequences*, (eds. Miyamoto M.M. and Cracraft J.) pp. 278-294 Oxford University Press, New York.

Hillis D.M. and Dixon M.T. (1991) Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology*, **66**: 411-446.

Hillis D.M. and Huelsenbeck J.P. (1992) Signal, noise and reliability in molecular phylogenetic analyses. *Journal of Heredity*, **83**: 189-195.

Hoffstetter R. (1964) Les serpentes du Néogène du Pakistan (couches des Siwaliks). *Bulletin Societé du Géologie de France*, **7**: 467-474.

Honeycutt R.L., Nedball M.A., Adkins R.M. and Janecek L.L. (1995) Mammalian mitochondrial DNA evolution: a comparison of the *cytochrome b* and *cytochrome c oxidase II* genes. *Journal of Molecular Evolution*, **40**: 260-272.

Hooijer D.A. (1975) Quaternary mammals west and east of Wallace's line. *The Netherlands Journal of Zoology*, **25**: 46-56.

Hope G.S. (1996a) History of Nothofagus in New Guinea and New Caledonia. In: *The Ecology and Biogeography of Nothofagus Forests,* (eds. Veblen T.T., Hill R.S. and Read J.) pp. 257-270 Yale University Press, New Haven and London.

Hope G.S. (1996b) Quaternary change and the historical biogeography of Pacific islands. In: *The origin and evolution of Pacific Island biotas, New Guinea to Eastern Polynesia: patterns and processes.,* (eds. Keast A. and Miller S.E.) pp. 165-190 SPB Academic Publishing, Amsterdam, The Nederlands.

Hope G.S. and Golson J. (1995) Late Quaternary change in the mountains of New Guinea. *Antiquity,* **69**: 818-830.

Huelsenbeck J.P. and Bull J.J. (1996) A likelihood ratio test to detect conflicting phylogenetic signal. *Systematic Biology,* **45**: 92-98.

Irwin D.M., Kocher T.D. and Wilson A.C. (1991) Evolution of the *cytochrome b* gene of mammals. *Journal of Molecular Evolution,* **32**: 128-144.

Jackman T.R., Larson A., de Queiroz K. and Losos J.B. (1999) Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. *Systematic Biology,* **48**: 254-285.

Janke A. and Árnason U. (1997) The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent Archosauria (birds and crocodiles). *Molecular Biology and Evolution,* **14**: 1266-1272.

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

Janke A., Xu X.F. and Árnason U. (1997) The complete mitochondrial genome of the Wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia, and Eutheria. *Proceedings of the National Academy of Sciences USA,* **94**: 1276-1281.

Johns R.J. (1982) Plant zonation. In: *Monographiae Biologiae,* Vol. 42 (ed. Gressitt J.L.) pp. 309-330 Dr. W. Junk, The Hague.

Johnson M.S. and Joll L.M. (1993) Genetic subdivision of the pearl oyster *Pinctada maxima* (Jameson, 1901) (Mollusca: Pteriidae) in northern Australia. *Australian Journal of Marine and Freshwater Research,* **44**: 619-626.

Jones M.R. and Torgersen T. (1988) Late Quaternary evolution of Lake Carpentaria on the Australia-New Guinea continental shelf. *Aurstralian Journal of Earth Sciences,* **35**: 313-324.

Jukes T.H. and Cantor C.R. (1969) Evolution of protein molecules. In: *Mammalian protein molecules,* (ed. Munro H.) pp. 21-132 Academic Press, New York.

Keenan C.P. (1994) Recent evolution of population structure in Australian Barramundi, *Lates Calcarifer* (Bloch): an example of isolation by distance in one dimension. *Australian Journal of Marine & Freshwater Research,* **45**: 1123-1148.

Kimura M. (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* **16**: 111-120.

Kishino H. and Hasegawa M. (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolution,* **29**: 170-179.

Kitchener D.J. and Suyanto A. 1996. Intraspecific morphological variation among island populations of small mammals in southern Indonesia In "Proceedings of the first international conference on eastern Inonesian-Australian vertebrate fauna.", Manado, Indonesia: 7-13.

Kluge A.G. (1989) A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae, Serpentes). *Systematic Zoology,* **38**: 7-25.

Kluge A.G. (1991) Boine snake phylogeny and research cycles. *Miscellaneous Publications, Museum of Zoology, University of Michigan,* **178**: 1-58.

Kluge A.G. (1993) *Aspidites* and the phylogeny of Pythonine snakes. *Records of the Australian Museum (Supplement)*, **19**: 1-77.

Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Pääbo S., Villablanca F.X. and Wilson A.C. (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Science USA*, **86**: 6196-6200.

Kornet D.J. and Turner H. (1999) Coding polymorphism for phylogeny reconstruction. *Systematic Biology*, **48**: 365-379.

Kraus F. and Brown W.M. (1998) Phylogenetic relationships of Colubroid snakes based on mitochondrial DNA sequences [Review]. *Zoological Journal of the Linnean Society*, **122**: 455-487.

Kumazawa Y. and Nishida M. (1993) Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *Journal of Molecular Evolution*, **37**: 380-398.

Kumazawa Y. and Nishida M. (1995) Variations in mitochondrial tRNA gene organization of reptiles as phylogenetic markers. *Molecular Biology and Evolution*, **12**: 759-772.

Kumazawa Y., Ota H., Nishida M. and Ozawa T. (1996) Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control region-like sequences duplicated and inserted into a tRNA gene cluster. *Molecular Biology and Evolution*, **13**: 1242-1254.

Kumazawa Y., Ota H., Nishida M. and Ozawa T. (1998) The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics*, **150**: 313-329.

Lamb T. and Avise J.C. (1992) Molecular and population genetic aspects of mitochondrial DNA variability in the Diamond Terrapin, *Malaclemys terrapin*. *Journal of Heredity*, **83**: 262-269.

Lambeck K. and Nakada M. (1990) Late Pleistocene and Holocene sea-level change along the Australian coast. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **89**: 143-176.

Lanave C., Preparata G., Saccone C. and Serio G. (1984) A new method for the calculation of evolutionary substitution rates. *Journal of Molecular Evolution,* **20**: 86-93.

Larget B. and Simon D.L. (1999) Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology & Evolution,* **16**: 750-759.

Lee M.S.Y., Bell G.L. and Caldwell M.W. (1999) The origin of snake feeding. *Nature,* **400**: 655-659.

Lee M.S.Y., Caldwell M.W. and Caldwell (1998) Anatomy and relationships of *Pachyrhachis problematicus*, a primitive snake with hind limbs. *Philosophical Transactions of the Royal Society of London B,* **353**: 1521-1552.

Lee W.-J. and Kocher T.D. (1995) Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. *Genetics,* **139**: 873-887.

Levinson G. and Gutman G.A. (1987) Slipped-strand mis-pairing: A major mechanism for DNA sequence evolution. *Molecular Biology and Evolution,* **4**: 203-221.

Leviton A.E., Gibbs R.H.J., Heal E. and Dawson C.E. (1985) Standards in herpetology and ichthyology: Part I. Standard symbolic codes for institutional resource collections in herpetology and ichthyology. *Copeia,* **1985**: 802-832.

Li W.-H., Wu C.-I. and Luo C.-C. (1985) Evolution of DNA sequences. In: *Molecular Evolutionary Genetics,* (ed. MacIntyre J.) pp. 1-94 Plenum Press.

Lopez J.V., Culver M., Stephens J.C., Johnson W.E. and O'Brien S.J. (1997) Rates of nuclear and cytoplasmic mitochondrial DNA sequence divergence in mammals. *Molecular Biology and Evolution,* **14**: 227-286.

Lopez J.V., Yuhki N., Masuda R., Modi W. and O'Brien S.J. (1994) Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *Journal of Molecular Evolution,* **39**: 174-190.

Lunt D.H. and Hyman B.C. (1997) Animal mitochondrial DNA recombination. *Nature,* **387**: 247.

Macey J.R., II J.A.S., Larson A. and Papenfuss T.J. (1998) Tandem duplication via Light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. *Molecular Biology and Evolution,* **15**: 71-75.

Macey J.R., Larson A., Ananjeva N.B., Fang Z. and Papenfuss T.J. (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Molecular Biology and Evolution,* **14**: 91-104.

Maddison W.P. (1989) Reconstructing character evolution on polytomous cladograms. *Cladistics,* **5**: 365-377.

Maddison W.P. and Maddison D.R. (1992) *MacClade: Analysis of phylogeny and character evolution. Vsersion 3.0.* Sinauer Associates.

Marshall H.D. and Baker A.J. (1997) Structural conservation and variation in the mitochondrial control region of Fringilline finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). *Molecular Biology and Evolution,* **14**: 173-184.

Martin A.P., Naylor G.J.P. and Palumbi S.R. (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature,* **357**: 153-155.

Martin A.P. and Palumbi S.R. (1993) Body size, metabolic rate, generation time and the molecular clock. *Proceedings of the National Academy of Science USA,* **90**: 4087-4091.

Marx H. and Rabb G.B. (1970) Character analysis: an empirical approach applied to advanced snakes. *Journal of Zoology, London,* **161**: 528-548.

Mayr E. (1969) *Principles of systematic zoology.* McGraw-Hill, New York.

McDiarmid R.W., Campbell J.A. and Touré T.A. (1999) *Snake Species of the World: A Taxomonic and Geographic Reference*. The Herpetologists' League, Washington, DC.

McDowell S.B. (1975) A catalogue of the snakes of New Guinea and the Solomons, with special reference to those in the Bernice P. Bishop Museum. Part II. Aniliodea and Pythoninae. *Journal of Herpetology,* **9**: 1-79.

McDowell S.B. (1987) Systematics. In: *Snakes: ecology and evolutionary biology,* (eds. Seigel R.A., Collins J.T. and Novak S.S.) pp. 3-50 Macmillan, New York.

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

Mindell D.P., Sorenson M.D. and Dimcheff D.E. (1998) Multiple independent origins of mitochondrial gene order in birds. *Proceedings of the National Academy of Sciences of the United States of America,* **95**: 10693-10697.

Minton S.A. and Minton M.R. (1973) *Giant Reptiles*. Charles Scribner's Sons, New York.

Miyamoto M.M. and Fitch W.M. (1995) Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology,* **44**: 64-76.

Moritz C. (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology,* **3**: 401-411.

Moritz C. and Brown W.M. (1986) Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. *Science,* **233**: 1425-1427.

Moritz C. and Brown W.M. (1987) Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. *Proceedings of the National Academy of Science USA,* **84**: 7183-7187.

Moritz C., Dowling T.E. and Brown W.M. (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics,* **18**: 269-292.

Murphy R.W., Sites Jr J.W., Buth D.G. and Haufler C.H. (1996) Proteins: Isozyme electrophoresis. In: *Molecular Systematics,* 2nd ed. (eds. Hillis D.M., Moritz C. and Mable B.K.) pp. 51-120 Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Muse S.V. and Weir B.S. (1992) Testing for equality of evolutionary rates. *Genetics,* **132**: 269-276.

Nei M. (1987) *Molecular Evolutionary Genetics*. ColumbiaUniversity Press, New York.

Nix H.A. and Kalma J.D. (1972) Climate as a dominant control in the biogeography of northern Australia and New Guinea. In: *Bridge and Barrier: the Natural and Cultural History of Torres Strait,* (ed. Walker D.) pp. 61-92 Australian National University, Canberra.

Norman J.A., Moritz C. and Limpus C.J. (1994) Mitochondrial DNA control region polymorphisms: genetic markers for ecological studies of marine turtles. *Molecular Ecology,* **3**: 363-373.

O'hUigin C. and Li W.H. (1992) The molecular clock ticks regularly in Muroid rodents and hamsters. *Journal of Molecular Evolution,* **35**: 377-384.

O'Shea M. (1996) *A Guide to the Snakes of Papua New Guinea*. Independent Publishing, Independent Group Pty Ltd., Port Moresby, PNG.

Pääbo S., Thomas W.K., Whitfield K.M., Kumazawa Y. and Wilson A.C. (1991) Rearrangements of mitochondrial transfer RNA genes in marsupials. *Journal of Molecular Evolution,* **33**: 426-430.

Parker F. (1982) *The Snakes of the Western Province*. PNG Dept. of Lands and Environment 82/1.

Pasveer J.P. and Aplin K.P. (1998) Quaternary vertebrate faunal succession on the Ayamaru Plateau, Bird's Head of Irian Jaya, Indonesia. In: *Proceedings of the 1996 I.S.I.R. Conference, Leiden,* pp. 891-930 Rodopi Press, Amsterdam.

Patton J.L. and da Silva M.N.F. (1998) Rivers, refuges, and ridges: the geography of speciation of Amazonian mammals. In: *Endless Forms: Species and Speciation,* (eds. Howard D.J. and Berlocher S.H.) pp. 202-213 Oxford University Press, New York.

Penny D. and Hendy M.D. (1985) Testing methods of evolutionary tree construction. *Cladistics,* **1**: 266-272.

Pesole G., Gissi C., De Chirico A. and Saccone C. (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of Molecular Evolution,* **48**: 427-434.

Pigram C.J. and Davies H.L. (1987) Terranes and the accretion history of the New Guinea orogen. *Bureau of Mineral Resources Journal of Australian Geology and Geophysics,* **10**: 193-211.

Polhemus D.A. and Polhemus J.T. (1998) Assembling New Guinea: 40 million years of island arc accretion as indicated by the distributions of aquatic Heteroptera (Insecta). In: *Biogeography and Geological Evolution of SE Asia,* (eds. Hall J. and Holloway J.D.) pp. 1-14 Backbuys Publishers, Amsterdam, The Nederlands.

Posada D. and Crandall K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**: 817-818.

Pratt T.K. (1982) Biogeography of birds in New Guinea. In: *Monographieae Biologicae,* Vol. 42 (ed. Gressit J.L.) pp. 815-836 Dr W Junk Publishers, The Hague.

Quinn T.W. (1992) The genetic legacy of Mother Goose - phylogeography patterns of lesser snow goose *Chen caerulescens* maternal lineages. *Molecular Ecology,* **1**: 105-117.

Quinn T.W. and Mindell D.P. (1996) Mitochondrial gene order adjacent to the control region in crocodile, turtle and tuatara. *Molecular Phylogenetics and Evolution,* **5**: 344-351.

Rage J. (1987) Fossil History. In: *Snakes: Ecology and Evolutionary Biology,* (eds. Seigel J.T. and Novak S.S.) pp. 51-76 McGraw-Hill Publishing Co.

Rage J.-C. (1984) The fossil snake *Cheilophis huerfanoensis* Gilmore, 1938 from Eocene of Colorade: redescription and reappraisal of relationships. *Journal of Vertebrate Paleontology,* **3**: 219-222.

Rage J.C. (1976) Les squamates du Miocène de Beni Mellal, Maroc. *Géologie Méditerranéenne,* **3**: 57-69.

Rage J.-C. and Richter A. (1994) A snake from the Lower Cretaceous (Barremian) of Spain: The oldest known snake. *Neues Jahrbuch für Geologie und Paläontologie Monatshefte,* **9**: 561-565.

Rambaut A. and Bromham L. (1998) Estimating divergence dates from molecular sequences. *Molecular Evolution and Biology,* **15**: 442-448.

Rand D.M. (1993) Endotherms, ectotherms and mitochondrial genome-size variation. *Journal of Molecular Evolution,* **37**: 281-295.

Rand D.M. (1994) Thermal habit, metabolic rate and the evolution of mitochondrial DNA. *Trends in Ecology and Evolution,* **9**: 125-131.

Read J. and Hope G. (1996) Ecology of Nothofagus forests of New Guinea and New Caledonia. In: *The Ecology and Biogeography of Nothofagus Forests,* (eds. Veblen T.T., Hill R.S. and Read J.) pp. 200-256 Yale University.

Richardson B.J., Baverstock P.R. and Adams M. (1986) *Allozyme Electrophoresis: A Handbook for Animal Systematics and Population Structure.* Academic Press, Sydney.

Riddle B.R. (1996) The molecular phylogeographic bridge between deep and shallow history in continental biotas. *Trends in Ecology and Evolution,* **11**: 207-211.

Rieppel O. (1984) Miniaturization of the lizard skull: its functional and evolutionary implications. In: *Symposium No 52 of the Zoological Society of London,* (ed. Ferguson M.W.J.) Academic Press, London.

Ripley L.S. and Glickman B.W. (1983) Unique self-complementarity of palindromic sequences provides DNA structural intermediates for mutation. *Cold Spring Harbor Symposium on Quantitative Biology,* **47**: 851-861.

Rodrigo A.G., Kelly-Borges M., Bergquist P.R. and Bergquist P.L. (1993) A randomisation test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. *New Zealand Journal of Botany,* **31**: 257-268.

Rodriguez F., Oliver J.F., Marín A. and Medina J.R. (1990) The general stochastic model of nucleotide substitutions. *Journal of Theoretical Biology,* **142**: 485-501.

Roe B.A., Ma D.-P., Wilson R.K. and Wong J.F.-H. (1985) The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *Journal of Biology and Chemistry,* **260**: 9759-9774.

Rogers A.R. (1995) Genetic evidence for a Pleistocene population explosion. *Evolution,* **49**: 608-615.

Rogers A.R. (1998) *Mismatch ver. 4.2. Ver. 4.2.*

Rogers A.R. and Harpending H. (1992) Population growth makes waves in the distribution of pair-wise genetic differences. *Molecular Biology and Evolution,* **9**: 552-569.

Rozas J. and Rozas R. (1999) DnaSP ver3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics,* **15**: 174-175.

Saccone C., Pesole G. and Sbisa E. (1991) The main regulatory region of mammalian mitochondrial DNA structure-function model and evolutionary pattern. *Journal of Molecular Evolution,* **33**: 83-91.

Saint K.M., Austin C.C., Donnellan S.C. and Hutchinson M.N. (1998) *C-mos*, a nuclear marker useful for Squamate phylogenetic analysis. *Molecular Phylogenetics and Evolution,* **10**: 259-263.

Saitou N. and Nei M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4**: 406-425.

Sambrook E., Fritsch F. and Maniatis T. (1989) *Molecular Cloning*. Cold Spring Harbor Press, New York.

Sarich V.M. (1985) Rodent macromolecular systematics. In: *Evolutionary Relationships among Rodents. A Multidisciplinary Analysis,* Vol. 92 (eds. Luckett W.P. and Hartenberger J.L.) pp. 423-452 Plenum Press, New York.

Sarich V.M. and Wilson A.C. (1967) Rates of albumin evolution in primates. *Proceedings of the National Academy of Science USA,* **58**: 142-148.

Scanlon J.D. (1996) Studies in the palaentology and systematics of Australian snakes. University of New South Wales, Australia. Thesis.

Scanlon J.D. and Lee M.S.Y. (2000) The Pleistocene serpent *Wonambi* and the early evolution of snakes. *Nature,* **403**: 416-420.

Scanlon J.D., Lee M.S.Y., Caldwell M.W. and Shine R. (1999) The palaeoecology of the primitive snake *Pachyrhachis*. *Historical Biology,* **13**: 127-152.

Schneider C.J., Cunningham M. and Moritz C. (1998) Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Molecular Ecology,* **7**: 487-498.

Schwaner T.D. and Dessauer H.C. (1981) Immunodiffusion evidence for the relationships of Papuan boids. *Journal of Herpetology,* **15**: 250-253.

Seddon J.M., Baverstock P.R. and Georges A. (1998) The rate of mitochondrial *12S rRNA* gene evolution is similar in freshwater turtles and marsupials. *Journal of Molecular Evolution,* **46**: 460-464.

Seutin G., Lang B.F., Mindell D.P. and Morais R. (1994) Evolution of the WANCY region in amniote mitochondrial DNA. *Molecular Biology and Evolution,* **11**: 329-340.

Shine R. (1991) *Australian Snakes, a Natural History*. Reed Books Pty Ltd.

Shine R. and Slip D.J. (1990) Biological aspects of the adaptive radiation of Australian pythons (Serpentes: Boidae). *Herpetologica,* **46**: 283-290.

Shoemaker J.S., Painter I.S. and Weir B.S. (1999) Bayesian statistics in genetics - a guide for the uninitiated [Review]. *Trends in Genetics*, **15**: 354-358.

Simpson G.G. (1951) The species concept. *Evolution*, **5**: 285-298.

Slatkin M. and Hudson R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**: 555-562.

Smith L.A. (1981a) A revision of the *Liasis olivaceus* species-group (Serpentes: Boidae) in Western Australia. *Records of the Western Australian Museum*, **9**: 227-233.

Smith L.A. (1981b) A revision of the python genera *Aspidites* and *Python* (Serpentes: Boidae) in Western Australia. *Records of the Western Australian Museum*, **9**: 211-226.

Smith L.A. (1985) A revision of the *Liasis childreni* species-group (Serpentes: Boidae). *Records of the Western Australian Museum*, **12**: 257-276.

Smith M.F., Thomas W.K. and Patton J.L. (1992) Mitochondrial DNA-like sequence in the nuclear genome of an Akodontine rodent. *Molecular Biology and Evolution*, **9**: 204-215.

Smith M.J. and Plane M. (1985) Pythonine snakes (Boidae) from the Miocene of Australia. *Journal of Australian Geology and Geophysics*, **9**: 191-195.

Sorenson M.D. (1999) *TreeRot. 2a*. Boston University.

Sorenson M.D. and Fleischer R.C. (1996) Multiple independent transpostions of mitochondrial DNA control region sequences to the nucleus. *Proceedings of the National Academy of Science, USA*, **93**: 15239-15243.

Springer M.S. (1995) Molecular clocks and the incompleteness of the fossil record. *Journal of Molecular Evolution*, **41**: 531-538.

Stanton D.J., Daehler L.L., Moritz C.C. and Brown W.M. (1994) Sequences with the potential to form stem-and-loop structures are associated with coding-region duplications in animal mitochondrial DNA. *Genetics*, **137**: 233-241.

Storr G.M., Smith L.A. and Johnstone R.E. (1986) *Snakes of Western Australia*. Western Australia Museum.

Streisinger G., Okada Y., Emrich J., Newton J., Tsugita A., Terzaghi E. and Inouye M. (1966) Frameshift mutations and the genetic code. *Cold Spring Harbor Symposium Quantitative Biology,* **31**: 77-84.

Strimmer K. and von Haeseler A. (1996) Quartet Puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution,* **13**: 964-969.

Stuhl G.M. (1932) Five new subspecies of the Family Boidae. *Occasional Papers of the Boston Society of Natural History,* **8**: 25-30, pls. 21-22.

Swofford D.L. (1999) *PAUP: phylogenetic analysis using parsimony. beta test version 4.0b2a.* Illinois Natural History Survey.

Swofford D.L., Olsen G.J., Waddell P.J. and Hillis D.M. (1996) Phylogenetic inference. In: *Molecular Systematics,* (eds. Hillis D.M., Moritz C. and Mable B.K.) Sinauer Associates, Inc., Sunderland, Massachusetts USA.

Szyndlar Z. and Böhme W. (1993) Die fossilen Schlangen Deutschlands: Geschichte der Faunen und ihrer Erforschung. *Mertensiella,* **3**: 381-431.

Taberlet P. (1998) Biodiversity at the intraspecific level: the comparative phylogeographic approach. *Journal of Biotechnology,* **64**: 91-100.

Taberlet P., Fumagalli L. and Wust-Saucy A. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology,* **7**: 453-464.

Tajima F. (1993) Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics,* **135**: 599-607.

Templeton A.R. (1983) Phylogenetic inference from restriction endonuclease cleavage maps with particular reference to the evolution of humans and the apes. *Evolution,* **37**: 221-244.

Thompson J.D., Higgins D.G. and Gibson T.J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research,* **22**: 4673-4680.

Thorpe R.S., McGregor D.P., Cumming A.M. and Jordan W.C. (1994) DNA evolution and colonization sequence of island lizards in relation to geological history: mtDNA RFLP, cytochrome b, cytochrome oxidase, 12SrRNA sequence and nuclear RAPD analysis. *Evolution,* **48**: 230-240.

Thyagarajan B., Padua R.A. and Campbell C. (1996) Mammalian mitochondria possess homologous DNA recombination activity. *Journal of Biological Chemistry,* **271**: 27536-27543.

Torgersen T., Hutchinson M.F., Searle D.E. and Nix H.A. (1983) General bathymetry of the Gulf of Carpentaria and the Quaternary physiography of Lake Carpentaria. *Palaeogeography, Palaeoclimatology, Palaeoecology,* **41**: 207-225.

Torgersen T., Jones M.R., Stephens A.W., Searle D.E. and Ullman W.J. (1985) Late Quaternary hydrological changes in the Gulf of Carpentaria. *Nature,* **313**: 785-787.

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

Torrance A. (1997) Evolution, phylogeny, biogeography and diversification of the Australasian Hydromyinae (Rodentia: Muridae). Harvard University. Unpublished PhD thesis.

Truswell E.M. (1993) Vegetation changes in the Australian Tertiary in response to climatic and phytogeographic forcing factors. *Australian Systematic Botany,* **6**: 533-557.

Underwood G. (1976) A systematic analysis of boid snakes. In: *Morphology and Biology of Reptiles. Linnean Society of London, Symposium Ser. 3,* (eds. Bellairs A.d.A. and Cox C.B.) Academic Press, London.

Underwood G. and Stimson A.F. (1990) A classification of pythons (Serpentes, Pythoninae). *Journal of Zoology London,* **221**: 565-603.

Walberg M.W. and Clayton D.A. (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Research,* **9**: 5411-5421.

Webb R.G. (1995) Redescription and neotype designation of *Pelochelys bibroni* from Southern New Guinea (Testudines: Trionychidae). *Chelonian Conservation and Biology,* **1**: 301-310.

Webb T. and Bartlein P.J. (1992) Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics,* **23**: 141-173.

Welter C., Dooley S. and Blin N. (1989) A rapid protocol for the purification of mitochondrial DNA suitable for studying restriction fragment length polymorphisms. *Gene,* **83**: 169-172.

Wilson A.C., Carlson S.S. and White T.J. (1977) Biochemical evolution. *Annual Review of Biochemistry,* **46**: 573-639.

Winter J.W. (1997) Responses of non-volant mammals to late Quaternary climatic changes in the wet tropics region of North-eastern Australia. *Wildlife Research,* **24**: 493-511.

Wolfe K.H. (1988) Mutation rates differ among regions of the mammalian genome. *Nature,* **337**: 283-285.

Wu C.I. and Li W.H. (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of the National Academy of Science USA,* **82**: 1741-1745.

1745.

Yang Z. (1996) Phylogenetic analysis using parsimony and likelihood methods. *Journal of Molecular Evolution,* **42**: 294-307.

Yang Z.H. and Rannala B. (1997) Bayesian phylogenetic inference using DNA sequences - a Markov Chain Monte Carlo method. *Molecular Biology & Evolution,* **14**: 717-724.

Zhang D.X. and Hewitt G.M. (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution,* **11**: 247-251.