Molecular Systematics of New Zealand skinks of the genus *Cyclodina*.

A thesis presented in partial fulfillment of the requirements

for the degree of

Master of Science in Genetics

at Massey University

New Zealand

Kerryn Elizabeth Slack

1996

Abstract.

The taxonomic status, relationships and possible origin(s) of eight skinks from the genus *Cyclodina* were investigated using molecular systematics - one from each of the six recognised and two proposed New Zealand species. DNA sequence data from the mitochondrial 12S ribosomal RNA gene was obtained using the polymerase chain reaction (PCR) and a thermocycling-based sequencing procedure. Phylogenetic analysis was carried out using spectral analysis, which utilizes new and sophisticated algorithms, and the maximum parsimony, minimum evolution and maximum likelihood options of PAUP* Version 4.0, a new test version of PAUP. New Zealand members of the skink genus *Leiolopisma* and three overseas skinks were also included in analyses.

A single resolved tree was not produced, which may indicate that the New Zealand *Cyclodina* diverged rapidly. The eight *Cyclodina* taxa form genetically distinct lineages, supporting the separate taxonomic status of each of the recognised and proposed species. The suggestion that the *C. aenea* population from the Poor Knights Islands is a separate species is well supported, the two *C. aenea* taxa being separated by and/or pairing with other taxa in most trees. However, the possibility of subspecific status cannot be excluded for *C. oliveri* from the Mokohinau Islands (the second proposed new species).

While the eight *Cyclodina* taxa form a closely related group, *L. fallai* pairs with *C. alani* and *L. zelandicum* with the Poor Knights *C. aenea* in many or all of the phylogenies. *L. moco* and the Mauritian skink *L. telfairi* also appear to have links with *Cyclodina*. The level of spectral analysis and bootstrap support is low for most of these relationships; nevertheless a monophyletic origin for *Cyclodina* with regards to the New Zealand *Leiolopisma* is not supported under the present taxonomic classification. Longer sequences from additional genes and a larger, more diverse set of skinks are required (in conjunction with other molecular, morphological and ecological information) before the exact relationships of these taxa and the origin(s) and divergence times of the New Zealand *Cyclodina* can be accurately determined. However, the results of this study do suggest that *Cyclodina* is older than previously thought, possibly even Gondwanan in origin.

Acknowledgments.

My profound thanks to David Penny (supervisor extraordinaire) for his patience and inspiration and from whom I have learned an enormous amount.

I am also grateful to Barry Scott (Head of the Molecular Genetics Unit) and Tim Brown (Head of the Department of Microbiology and Genetics) for taking me on and for their patience (and for Tim's funny stories).

Lab work was financed by Massey University, the Graduate Research Fund, the Department of Microbiology and Genetics, the Molecular Genetics Unit and FARSIDE.

Thanks also to Charlie Daugherty of Victoria University for supplying the skink samples without which this study would not have been possible, and for several helpful discussions.

This thesis could not have been achieved without Robert Hickson, Peter Lockhart and Trish McLenachan who helped me with lab procedures, computer programmes, evolutionary theory and numerous other matters. Thanks guys!

Kay Rutherfurd, Mike Collett and Sirinda Yunchalard provided several extremely useful lab tips, and Alan Cooper's advice on lab procedures and how to use the programme ESEE was also thoroughly appreciated. Peter Waddell introduced me to PAUP* and I am also grateful to David Swofford for his help.

Many thanks to Rod Hitchmough and Gina Lento - always willing to help, often at very short notice. And to Mike Charleston, for a number of reasons. I also wish to thank Kath Hurr and Bennet McComish for the interesting discussions and useful comments.

Thanks to Geoff Patterson, Dave Towns, Tony Whitaker, Peter Anderson and Tony Jewell and other members of SRARNZ for interesting and informative discussions about the New Zealand skinks and to Trevor Worthy for osteological information. Thanks also to Geoff for the hospitality, and for letting me drive the minivan!

I am grateful to Ron Tucker, Paul Hocquard and Catherine Norman, who tracked things down, arranged to have things fixed up and supplied me with 'out of hours' passes at regular intervals. I also wish to offer my grateful thanks to Jackie McDonald and Laura McKenzie, always friendly, always helpful.

Thanks also to members of the Palmerston North Aikido Club, for helping me to stay sane.

Finally, with my deepest thanks for their patience and support, I dedicate this thesis to my family and to Curtis with all my love.

Table of Contents.

List of Figures.

List of Tables.

Chapter One: Introduction.

This study uses molecular systematics to investigate New Zealand members of the skink genus *Cyclodina*. Specifically, the polymerase chain reaction (PCR) and a thermocycling-based sequencing procedure were used to obtain a short stretch of DNA sequence (approximately 385 base pairs from the mitochondrial 12S ribosomal RNA gene) for eight skinks - one from each of the six currently recognised and extant New Zealand *Cyclodina* species and one from each of two populations proposed (Vos, 1988) to be 'cryptic' species[1] (the *C. aenea* population from the Poor Knights Islands and the *C. oliveri* population from the Mokohinau Islands). The sequence data obtained from the eight individuals was then analysed using the latest phylogenetic analysis programmes (the Hadamard conjugation and a new test version of PAUP [phylogenetic analysis under parsimony] - PAUP* Version 4.0) and phylogenetic inferences made about the taxonomic status, relationships and origin(s) of the New Zealand *Cyclodina*.

With the results of this study, preliminary investigations of the New Zealand skinks using molecular data - both allozymes (Vos, 1988; Daugherty *et al.*, 1990b; Patterson and Daugherty, 1990, 1994 and unpublished data) and DNA sequences (Hickson *et al.*, 1992; Hickson, 1993; this study) - are all but complete. The individuals examined using DNA sequences form a subset of those investigated with allozyme data - allowing direct comparison of the results of analysis of the two datasets. The conclusions drawn from the compared datasets can be used for conservation purposes, while the use of molecular genetic information to study model groups such as the skinks should aid in addressing long-standing questions about evolutionary processes in New Zealand and in general, particularly when examined in conjunction with other information such as ecological data (Hickson, 1993). In this chapter, the literature is reviewed and the molecular systematic approach used in this study and the issues necessitating it are discussed.

Molecular Systematics.

Molecular systematics combines the use of molecular data (such as DNA sequences) with methods to analyse this information (such as parsimony or distance methods). In recent years, developments and/or refinements in techniques such as PCR and DNA sequencing and in analytic programmes have firmly established molecular systematics as a powerful tool for the study of evolution and for taxonomic and conservation purposes (for example, see Hillis *et al.*, 1996 and Moritz and Hillis, 1996).

---

[1] 'Cryptic' species are groups which are virtually indistinguishable from other species on morphological grounds (and so tend to be misclassified when morphological data alone is used) but which are definitely distinct non-interbreeding species.

DNA Sequences.

'Complete sequencing of homologous DNA fragments from different organisms provides the most powerful and direct method for obtaining information on amount of genetic variation or extent of genetic divergence [among individuals, populations, species or higher taxa]' (Harrison, 1989; see also Kocher *et al.,* 1989; Hickson, 1993, for discussions on limitations of other methods). This is because DNA sequences represent discrete character state data, the number of characters in even the smallest genome is huge, and the basis of genetic variation can be directly investigated.

Because both the extent of genetic diversity between taxa and the specific types of nucleotide substitution that have occurred can be directly examined, sequences are useful for both taxonomic purposes and for the study of evolutionary relationships and the testing of hypotheses about evolutionary processes (Kocher and White, 1989; Kocher *et al.,* 1989; White *et al.,* 1989; Arnheim *et al.,* 1990; Simon *et al.,* 1991). Indeed, sequence data has now been used to investigate both relatively recent events (for example Cann *et al.,* 1987; Vigilant *et al.,* 1991; Ward *et al.,* 1991) and very ancient ones (for example Woese, 1987; Lake, 1988, 1990, 1991; Rivera and Lake, 1992). In addition, distinct populations requiring further study and/or conservation strategies can be identified using sequence analysis and comparison to allozyme data, morphology and other biological datasets (Arnheim *et al.,* 1990; Bowen *et al.,* 1991). Sequence data can also be used for other conservation purposes (for example, see Baker and Palumbi, 1994, for a discussion of the use of molecular data for monitoring whaling).

The Polymerase Chain Reaction (PCR) and Sequencing.

Until the development of PCR (Saiki *et al.,* 1988; Arnheim and Erlich, 1992) obtaining DNA sequences (using approaches such as the construction and screening of clone libraries for each of the individuals which need to be examined) was difficult, labour-intensive, time-consuming and expensive (Kocher *et al.,* 1989; Sambrook *et al.,* 1989). However, with PCR, sequencing has the potential to be as simple, rapid and efficient as allozyme analysis (Gyllensten and Erlich, 1988; Kocher and White, 1989; White *et al.,* 1989; Arnheim *et al.,* 1990; Innis and Gelfand, 1990) and the sequence obtained would seem to be at least as reliable as that produced by conventional cloning and sequencing (Pääbo and Wilson, 1988; Kocher *et al.,* 1989). The process is also versatile (White *et al.,* 1989; Arnheim *et al.,* 1990; Hickson *et al.,* 1992) and relatively inexpensive (Arnheim *et al.,* 1990; Hickson *et al.,* 1992).

Essentially, PCR is the enzymatic amplification of a specific region of DNA many times over. The process involves three steps (see Palumbi *et al.,* 1989; Innis and Gelfand, 1990; White *et al.,* 1989). Firstly, the template DNA (from which copies are to be made) is denatured by means of a high temperature (normally 90-94°C). Secondly, the temperature is lowered to allow the primers - two short (20-30 nucleotides), single-stranded pieces of DNA - to bind to sequences on complementary strands of the DNA. The temperature used at this stage is between 37 and 60°C depending the specificity of the primers. Finally, the temperature is raised slightly (normally to 72°C) allowing a heat-stable DNA polymerase (such as *Taq* polymerase - Saiki *et al.,* 1988) to bind to the primers and, in the presence of all four deoxynucleotide triphosphates (dNTPs), produce copies of the DNA between the primers. As each new copy acts as a template for further duplication, with the amount of primers in sufficient excess the process is exponential, and by repeating the above steps 20-40 times millions of copies of the desired region can be produced in a few hours.

The amplification products produced by PCR are concentrated enough that they can be sequenced directly (single-stranded sequencing - see Gyllensten and Erlich, 1988; Kocher *et al.,* 1989; double-stranded sequencing - see Casanova *et al.,* 1990; Hickson *et al.,* 1992; Anon., 1992). PCR has several other advantages. DNA can be obtained from frozen tissue, dried or ethanol-stored specimens (for example, see Kocher *et al.,* 1989; Hickson *et al.,* 1992), subfossil[2] bones (Cooper *et al.,* 1992) and museum skins (Wayne and Jenks, 1991) as well as from fresh material.

In addition, only a small amount of tissue is required - a few nanograms of fresh tissue, less than one milligram for old specimens (Kocher *et al.,* 1989). And preparations of total genomic DNA can be used rather than purified mitochondrial DNA providing the primers are of sufficient specificity (Kocher *et al.,* 1989). Furthermore, sequence from specific regions of DNA can now be obtained from organisms for which there is little or no background genetic information by using 'universal' primers (Kocher *et al.,* 1989; see also below).

With the popularity of sequencing growing following the development of PCR, DNA sequence databases (such as Genbank, EMBL) have expanded rapidly. Access to a wide range of sequences means that newly obtained sequences can be checked relatively quickly and easily to ensure that they do represent the correct group and are not the result of contamination and, in addition, larger, more diverse datasets are available for analysis.

---

[2] Nonfossilised bones no more than 30, 000 years old (Daugherty *et al.,* 1993)

'Universal' Primers and the Use of the Mitochondrial 12S Ribosomal RNA (rRNA) Gene.

The lack of sequence data for *Cyclodina* and, indeed, for many other lizards when this study began necessitated the use of 'universal' primers for PCR and sequencing. These are primers which anneal to conserved regions of the genome and so can be used to obtain homologous sequences from a wide variety of taxa for which there is little or no genetic information. In 1989, Kocher *et al.* identified three such pairs of primers for mitochondrial DNA (mtDNA) - two for cytochrome b, two for the 12S rRNA gene and three for the control region.

Hickson *et al.* (1992; see also Hickson, 1993) found that the 12S rRNA primers, used by Kocher *et al.* (1989) to successfully amplify a 386 base pair (bp) segment of the small rRNA from humans, rodents, marsupials, crocodiles, insects, spiders, fishes, amphibians and birds, were also suitable for amplifying 12S rRNA sequences from skinks of the genus *Leiolopisma*. Therefore the 12S rRNA primers were chosen for use in this study.

Animal mtDNA has several features which make it useful for the purposes of analysis. These include that there is rarely more than one type of mtDNA in an organism, that recombination appears to be relatively infrequent, and that mitochondrial inheritance is predominantly maternal (Harrison, 1989). In addition, mtDNA appears to evolve at a more rapid rate than nuclear genes in vertebrates, facilitating the study of populations and other closely related taxa (Harrison, 1989), although older divergences are also being investigated using the more conserved positions among the coding regions (Simon *et al.,* 1991; Cooper *et al.,* 1992; Simon *et al.*, 1994).

Phylogenetic Analysis of DNA Sequence Data.

Aligned DNA sequences contain signals - patterns of similarity and difference between taxa at discrete sites - which are evaluated by phylogenetic analysis methods and used to infer evolutionary relationships. Different methods use differing amounts of this information and different search strategies and optimality criteria (including parsimony - see Swofford and Olsen, 1990; maximum likelihood - Felsenstein, 1981; and minimum evolution - as in neighbor-joining - Saitou and Nei, 1987) to produce an inferred pattern of evolutionary relationships or a phylogeny. A large number of analytic techniques for constructing phylogenies from sequence data and investigating taxonomy are now available - all with their own advantages and limitations (see Felsenstein, 1988; Nei, 1991; Penny *et al.*, 1992; Charleston *et al.,* 1994; Swofford *et al.*, 1996, for reviews).

One method used in this study is spectral analysis, which uses new algorithms: the Hadamard conjugation and the closest tree optimality criterion (Hendy and Penny, 1989, 1993; Penny *et al.*, 1992, 1993a; Lento *et al.*, 1995). Standard methods used include the maximum parsimony, minimum evolution and maximum likelihood options of a new test version of PAUP (phylogenetic analysis using parsimony; see Swofford, 1985, 1993) - PAUP* Version 4.0 (1995, 1996) - and, to a lesser extent, the parsimony and neighbor-joining options of PHYLIP (phylogeny inference package; see Felsenstein, 1993).

Limitations in Using PCR and mtDNA Sequences and the Use of Other Datasets.

While mtDNA has several advantages for phylogenetic analysis, it represents only a small amount of the organism's genome, and this can be exacerbated by uniparental inheritance and lack of recombination. The use of sequences several thousand bp in length from several unlinked loci is preferred (Nei, 1987; Martin *et al.*, 1990), however the use of PCR generally restricts studies to shorter sequences (Kocher *et al.*, 1989). In addition, the entire mtDNA genome represents a single genetic marker unlinked to nuclear genes, thus patterns of variation for nuclear and mtDNA markers may not be concordant (Harrison 1989). However, isolation and direct sequencing of nuclear loci is hampered by diploidy and multiple copies of genes. Allelic variants can be separated by various techniques (Sheffield *et al.*, 1990; Gyllensten and Erlich, 1988; Jeffreys *et al.*, 1990), but these are demanding. Single-copy, conservative nuclear genes can be rapidly sequenced, but are not much use for population studies unless they span intergenic regions (see Palumbi and Baker, 1994).

Furthermore, as no current algorithm for phylogenetic reconstruction based on sequence data is simultaneously efficient, consistent, powerful, robust and falsifiable (Penny *et al.*, 1992), most phylogenetic trees are probably incorrect (Penny *et al.*, 1992). Some examples, among many, of the difficulties in phylogenetic reconstruction are revealed in Faith (1990), Thomas *et al.* (1990), Holmes *et al.* (1993a and b), Milinkovich *et al.* (1993), Novacek (1993) and Árnason and Gullberg (1994). Statistical methods to assess the reliability of phylogenies are being developed (for example, see Li and Gouy, 1991), but 'phylogenetic trees should be regarded as hypotheses and subject to error' (Hickson, 1993).

Investigating evolutionary processes requires knowledge of the rate of sequence evolution in the organism being studied. This poses certain problems - rates are lineage-dependant (Thorpe, 1982; Goddard *et al.*, 1990; Palmer, 1990), thus caution must be taken when applying rates determined for one group to another (Wilson *et al.*, 1985, 1987); single genes or proteins may be unreliable in the provision of a temporal framework (Wilson *et al.*, 1987; Easteal, 1990) and the models used to describe the process of sequence evolution may not be appropriate (Gillespie, 1986; Wilson *et al.*, 1987).

Moreover, molecular data is only part of the picture in understanding how evolution operates. It must be remembered that morphological variation, the basis for natural selection, forms the heart of the theory of evolution (Darwin, 1859).

The use of other datasets can ease some of these problems. The combination of diverse datasets, each with their own limitations and levels of resolution, increases confidence in phylogenetic conclusions and inconsistencies can identify false assumptions or limitations in models or data (Hillis, 1987; Patterson, 1987; Hillis and Moritz, 1990; Sytsma, 1990). The use of combined datasets can also suggest theories that it would not be possible to infer from the individual datasets alone - for example, use of both allozyme and mtDNA sequence data allowed the identification of a potential case of hybridisation between an *L. maccanni* population at Gorge Burn and a sympatric *L. n. polychroma* population (Hickson *et al.,* 1992; Hickson, 1993). However, for different datasets to be useful, the results of one should be generally predictive of those for another.

In addition, one of the two main forms of phylogenetic analysis used in this study, the Hadamard conjugation (Hendy and Penny, 1989, 1993; Penny *et al.,* 1992, 1993a; Lento *et al.,* 1995), provides a quantitative measure of the support in the data for and against conflicting associations of taxa, allowing confidence in the resulting phylogeny to be assessed.

Molecular Systematics and the New Zealand Fauna and Flora.

The application of molecular systematics should aid in addressing long-standing questions about the origins and evolution of New Zealand's fauna and flora and about evolutionary processes in general. New Zealand's biota is often compared to that of other isolated Pacific islands with amazing fauna and flora, such as Hawaii and the Galapagos (Daugherty *et al.,* 1993). However, unlike these islands, New Zealand is of ancient continental origin, has a temperate climate and has been subjected to repeated glaciations and drowning events (Daugherty *et al.,* 1993). As such it offers a wonderful opportunity for studying evolution in action (for example, see Fleming, 1958). However, in the past this potential has not been realised due to lack of suitable investigative techniques (Hickson, 1993). Phylogenetic analysis has, by default, been concentrated on morphological taxonomy and biogeographic studies (see Hickson, 1993). The development of biochemical and molecular genetic techniques and analysis and the use of concerted approaches now allow these possibilities to be explored (for example, see Cooper *et al.*, 1992; Hickson, 1993; Paterson *et al.*, 1993, 1995a and b).

In addition, the use of PCR and sequencing in conjunction with other molecular techniques (such as allozyme electrophoresis) aids in the identification of cryptic species (see Daugherty *et al.,* 1993) and in the clarification of relationships within traditionally 'difficult' taxonomic groups. New Zealand has a responsibility under the Biodiversity Convention to conserve unique groups. This involves more than just legal protection - research to identify which groups are most in need of protection, and to determine their conservation needs (for example, distribution, ecology and vulnerability to predation and/or habitat disturbance) is also required.

The New Zealand Herpetofauna.

The herpetofauna of New Zealand (see Table 1.1) are a fascinating group. Amongst them are numbered the tuatara (sole survivor of the reptile order Rhynchocephalia) and New Zealand's endemic *Leiopelma* frogs - believed to be one of the oldest groups in the order Salientia (frogs and toads) (Bull and Whitaker, 1975; Hudson, 1994). The geckos, too, are now thought to be of ancient origin (Hudson, 1994), and it is possible that the skinks also have Gondwanaland origins (Daugherty *et al.,* 1993; Hickson, 1993).

There are more lizard species (skinks and geckos) in New Zealand than in most temperate, continental areas of similar size (Daugherty *et al.,* 1993). Several of the world's most southerly lizards are found in New Zealand - the harlequin gecko (*Hoplodactylus rakiurae*) and the southern and small-eared skinks (*Leiolopisma notosaurus* and *L. stenotis* respectively) (Hudson, 1994; Higham, 1995). New Zealand's geckos are uncommon in that all give birth to live young - all other gecko species are oviparous except for one New Caledonian species (Bull and Whitaker, 1975; Hudson, 1994).

Molecular Investigations into the New Zealand Herpetofauna.

Molecular investigations into the New Zealand herpetofauna are required for several reasons. Firstly, the evolutionary history of our terrestrial reptile and amphibian fauna is difficult to determine due to lack of a fossil record (Daugherty *et al.,* 1993). This is complicated by the fact that species within a group are often very similar morphologically, leading to difficulties in taxonomic classification. Use of allozyme electrophoresis has already revealed cryptic species or potential cryptic species within all groups of the endemic New Zealand herpetofauna (tuatara - Daugherty *et al.,* 1990a; frogs - Bell, B. D., Daugherty, C. H. and May, J. H., unpublished data; geckos - Hitchmough, R. A., unpublished data; skinks - Vos, 1988; Daugherty *et al.,* 1990b; Patterson and Daugherty, 1990, 1994, unpublished data). Taken in conjunction with the reputation of New Zealand's endemic fauna as a whole for elusiveness, restricted distribution and morphological similarity (Towns, 1985; Pickard and Towns, 1988), this large number

Table 1.1. The extant herpetofauna of New Zealand[†].

Terrestrial/Aquatic		
Tuatara	2 endemic species	genus *Sphenodon*
Frogs	3 endemic species (plus 1*)	genus *Leiopelma*
	3 introduced species	genus *Litoria*
Geckos	16 endemic species[‡] (plus 13*)	genus *Hoplodactylus* (9 species plus 13*)
		genus *Naultinus* (7 species)
Skinks	26 endemic species[ƒ] (plus 4*)	genus *Leiolopisma* (20 species plus 2*)
		genus *Cyclodina* (6 species plus 2*)
	1 introduced species	genus *Lampropholis*
Turtles	2 introduced freshwater turtles	one *Chelodina* species
		one *Chrysemys* subspecies
Marine		
Turtles	4 regularly visiting marine turtles	one *Caretta* subspecies
		one *Chelonia* species
		one *Dermachelys* species
		one *Eretmochelys* species
Sea snakes	1 relatively regular visitor	genus *Pelamis*
	2 occasional visitors	genus *Laticauda*

[†] Table constructed from information in Daugherty *et al.* (1994) (tuatara, frogs, geckos and endemic skinks), and Pickard and Towns (1988) and Hudson (1994) (introduced skink species, turtles and sea snakes).

* Not yet formally recognised (proposed new species) - see Daugherty *et al.* (1994).

[‡] Excludes *N. poecilochlorus* which Daugherty *et al.* (1994) do not believe to be a valid species. Includes *H. nebulosus* which Daugherty *et al.* (1994) do believe to be a valid species.

[ƒ] Excludes *L. gracilicorpus* which Daugherty *et al.* (1994) believe to be either extinct or synonymous with another species. Includes *L. waimatense* (formerly *L. otagense* form "waimatense") which is now generally accepted to be a separate species (G. B. Patterson, pers. comm.).

suggests that there may yet be unidentified species. In spite of the many advantages provided by sequence data, however, PCR and sequencing have only been used to investigate members of the skink genus *Leiolopisma* and one individual from the genus *Cyclodina* (*C. aenea*) to date (Hickson *et al.,* 1992; Hickson, 1993) among the New Zealand herpetofauna.

Conservation.

Today, many of the endemic species of herpetofauna which once existed on mainland New Zealand are either extinct (Worthy, 1987a and b, 1991; Hudson, 1994) or restricted to offshore islands (Worthy, 1987c; Pickard and Towns, 1988; Daugherty *et al.,* 1994; Higham, 1995). Many are considered to be rare, threatened or endangered (Daugherty *et al.,* 1994; see also SRARNZ Notes, 1995b). This is believed to be due mostly to the effects of predation by introduced mammals (particularly rodents, but also mustelids and cats), with vulnerable species only surviving on pest-free islands or in a few areas on the mainland which (for various reasons) offer protection from predation (Barwick, 1959; Bull and Whitaker, 1975; Hardy, 1977; Towns *et al.,* 1985; Worthy, 1987b and c; Towns, 1991; Taylor and Thomas, 1993; Higham, 1995). However, habitat destruction by humans (Bull and Whitaker, 1975; Daugherty *et al.,* 1993; Hudson, 1994; Higham, 1995) and other introduced animals, such as rabbits (*Oryctolagus cuniculus*) and possums (*Trichosurus vulpecula*) (Hudson, 1994; but see also Towns, 1991), undoubtedly also contributes.

While most endemic amphibian and reptile species and visiting marine turtles and sea snakes are protected under the Wildlife Act (1953) and an Order in Council (1981) amending this Act, four species of lizards are not: *Leiolopisma nigriplantare maccanni*, *Cyclodina aenea*, *Hoplodactylus maculatus* and *H. granulatus*. Of these, *L. n. maccanni* no longer exists as a name, but apparently *L. n. polychroma* - a newly erected species (Patterson and Daugherty, 1990) comprising a large proportion of the skinks formerly under *L. n. maccanni* - is now regarded as unprotected (SRARNZ Notes, 1995a).

Collection and trade in the four currently unprotected lizard species does occur (Hudson, 1994; SRARNZ Notes, 1994, 1995a) and while the collection of 'common' skinks will almost certainly not lead to extinction, it may well lead to declines in local population numbers and will undoubtedly cause habitat degradation. (SRARNZ Notes, 1994). And in light of the existence of possibly unidentified cryptic species among our herpetofauna, collection and trade in the unprotected species (and the consequent dangers of possibly exporting very rare animals as opposed to 'common' ones) is particularly of concern. However, it is likely that the legislation will be changed to provide these species with full protection in the near future.

The influence of still other factors have not yet been evaluated. For example - the New Zealand Tourism Board would like to see international tourist numbers rise to three million by 2004 (Sage, 1995). Considering that many of these overseas visitors come because of New Zealand's 'clean green' image and visit parks and reserves where these species are found, the impacts that the increased numbers will undoubtedly have (Sage, 1995), particularly on the rarer species, need to be assessed.

As mentioned earlier, conservation of unique populations requires research into a variety of areas. Conservation management strategies for species on the Department of Conservation's conservation priority list have been implemented or are being designed (SRARNZ Notes, 1993, 1994, 1995a and b, 1996; Daugherty *et al.,* 1994) and this should aid in protecting the designated groups. However, it remains crucial to identify any further cryptic species: '.....taxonomy can kill when distinct species are not accorded specific status' (Gittleman and Pimm, 1991; see also Daugherty *et al.,* 1990a).

The New Zealand Skinks (Reptilia: Lacertilia: Scincidae).

Within the extant endemic New Zealand herpetofauna (excluding marine turtles and sea snakes), 30 out of 65 (46%) of the currently recognised or proposed species are skink species (see Table 1.1). These species are divided into two genera: *Leiolopisma* Duméril and Bibron 1839 and *Cyclodina* Girard 1857. *Leiolopisma* contains the greater number of species and much of the information available for the New Zealand skinks deals with these species. Currently there are twenty-one recognised species of *Leiolopisma* endemic to New Zealand (one of which is considered to be either extinct or synonymous with another species), two more have been proposed and several others may exist (see Table 1.2; Daugherty *et al.,* 1994; SRARNZ Notes 1994, 1995a and b, 1996; G. B. Patterson, pers. comm.).

The world-wide distribution and total number of *Leiolopisma* species is at present unclear. Prior to 1974, the genus *Leiolopisma* included species from Asia, Africa, Mid and Central America and Australasia (Greer, 1970). Subsequently it was restricted to a South Pacific and Mauritius distribution - with species found on Mauritius, in Australia, New Caledonia and New Zealand and on Lord Howe and Norfolk Islands (Greer, 1974). This was extended to include Fiji in 1985 with the recognition of a *Leiolopisma* species there (*L. alazon*; Zug, 1985). Currently, the genus consists of species on Mauritius (the type species *L. telfairi*) and in New Zealand, Fiji (Zug, 1985) and New Caledonia (Bauer and Vindum, 1990; Sadlier, 1986) - the Lord Howe/Norfolk Island species having been reclassified as *Cyclodina* and then *Pseudemoia* (see below), and the Australian species having been reclassified into five new genera in 1990

based on immunological data (Hutchinson *et al.,* 1990). It has been suggested several times (for example Sadlier, 1986; Hutchinson *et al.,* 1990) that *Leiolopisma* should become monotypic - with *L. telfairi* as the only *Leiolopisma* species. However, before any meaningful generic classification of the '*Leiolopisma*' species can be achieved further research is needed, using molecular data as well as other datasets such as immunological studies. As yet, few members of the '*Leiolopisma* assemblage' have been investigated using sequence data - only members of the New Zealand *Leiolopisma* and one individual each from *C. aenea*, *L. telfairi*, *Lampropholis guichenoti* (an Australian skink) and *Tropidoscincus rohssii* (a New Caledonian skink) (Hickson, 1993). This is discussed further in Chapter Five.

At present, only six extant New Zealand *Cyclodina* species are formally recognised (see Table 1.2), most of which are limited mainly to islands and thought to be rare or endangered (Daugherty *et al.,* 1994). However, allozyme data suggests that there are at least two more currently unrecognised New Zealand species (Vos, 1988; see Table 1.2). A seventh species, *C. northlandi*, was identified from subfossil bones in Northland, New Zealand in 1991 (Worthy, 1991) and appears to be extinct. Until recently, *Cyclodina* was considered to be unique to New Zealand. However, in 1986, Cogger reclassified the Lord Howe/Norfolk Island species (*L. lichenigerum)* as *Cyclodina*. Subsequently, he placed it in the genus *Pseudemoia* (Cogger, 1992 - not sighted; G. B. Patterson, pers. comm.). The status of this species remains unclear (it is referred to as *L. lichenigerum* throughout this study) and the classification of the Fijian species as belonging to *Leiolopisma* rather than *Cyclodina* is also uncertain (Zug, 1985).

The nomenclature used in this study for the formally recognised New Zealand skink species is that of Hardy (1977), Patterson and Daugherty (1990, 1994) and Worthy (1991), while that used for proposed or possible species is that of Daugherty *et al.* (1994) and SRARNZ Notes (1994, 1995a) (see Table 1.2). Vos (1988)'s nomenclature of *C. oliveri* (Mokohinau) has been retained, but her nomenclature of *C. aenea* (Aorangi) has been replaced by that of Daugherty *et al.* (1994). New Zealand has one introduced species of skink (*Lampropholis delicata*) - this species is not discussed further in this study.

New Zealand's *Leiolopisma* species are distributed throughout the country while the *Cyclodina* species are restricted to the North Island and northern offshore islands (Pickard and Towns, 1988; Patterson and Daugherty, 1990, 1994; see Table 1.2). These skinks live in a variety of habitats (Bull and Whitaker, 1975; Towns *et al.,* 1985; Hudson, 1994), are mostly terrestrial, include both nocturnal and diurnal species and are insectivores, frugivores or scavengers (Hardy, 1977; Bull and Whitaker, 1975; Hudson, 1994). In size, they range from approximately 130 mm to 350 mm in total length (Hudson, 1994) and all but one (*L. suteri*) bear their young live (Bull and Whitaker, 1975; Hudson, 1994). Half of New Zealand's skink species are restricted mainly to islands and over half are now classified as rare, vulnerable or endangered (Daugherty *et al.,* 1994).

Table 1.2. New Zealand *Leiolopisma* and *Cyclodina* species - formally recognised (including extinct), recognised, proposed and possible species[‡]. Distributions for those species given in Pickard and Towns (1988) are only given here as North Island (indicating species found only in the North Island and/or northern offshore islands), South Island (indicating species found only in the South Island and/or southern offshore islands), or North/South (indicating found throughout New Zealand). More detailed distribution data is given for species recognised or proposed since 1988.

***Cyclodina* Girard 1857**

Formally recognised species:	Distribution:
C. aenea Girard 1857	North Island
C. alani (Robb 1970)	North Island
C. macgregori (Robb 1975)	North Island
C. northlandi Worthy 1991*	North Island
C. oliveri (McCann 1955)	North Island
C. ornata (Gray 1843)	North Island
C. whitakeri Hardy 1977	North Island
Proposed species:	
'*C. aenea*' (Poor Knights Islands) (Vos, 1988)	Poor Knights Islands
'*C. oliveri*' (Mokohinau Islands) (Vos, 1988)	Mokohinau Islands and Hen and Chicken Islands

***Leiolopisma* Duméril and Bibron 1839**

Formally recognised species:	Distribution:
L. acrinasum Hardy 1977	South Island
L. chloronoton Hardy 1977	South Island
L. fallai McCann 1955[ƒ]	North Island
L. gracilicorpus Hardy 1977*[†]	
L. grande (Gray 1845)	South Island
L. homalonotum (Boulenger 1906)	North Island
L. inconspicuum Patterson and Daugherty, 1990	Otago and Southland
L. infrapunctatum (Boulenger 1887)	North/South
L. lineoocellatum (Duméril and Duméril 1851)	North/South
L. maccanni Patterson and Daugherty, 1990	Southern half of the South Island, east of the main divide
L. microlepis Patterson and Daugherty, 1990	Central North Island
L. moco (Duméril and Bibron, 1839)	North Island
L. nigriplantare (Peters, 1873)	
- subspecies *nigriplantare* (Peters, 1873)	Chatham Islands
- subspecies *polychroma* Patterson and Daugherty, 1990	Central North Island south to Stewart Island (inclusive)
L. notosaurus Patterson and Daugherty, 1990	Stewart Island and Codfish Island
L. otagense (form "otagense") Hardy 1977	South Island
L. smithi (Gray 1845)	North Island
L. stenotis Patterson and Daugherty, 1994	Stewart Island
L. striatum (Buller 1871)	North Island
L. suteri (Boulenger 1906)	North Island
L. zelandicum (Gray 1843)	North/South

..... continued next page

Recognised (but not yet formally) species:	
L. waimatense[§] (formerly *L. otagense* form "waimatense") Hardy 1977	South Island
Proposed species:	
'*L. longtoes*' (Patterson and Daugherty, unpublished data)	South Canterbury to Nelson/Marlborough
'*L. Open Bay Island*' (Patterson and Daugherty, unpublished data)	Open Bay Island (south of Hokitika, north of Haast)
Other possible species:	
'Big Bay skink' (looks to be a distinct species, most closely related to *L. notosaurus*; Daugherty, Patterson and Hitchmough, unpublished data)	South Westland
'West Coast skink 1' (looks to be a distinct species, very closely related to *L. infrapunctatum*; Daugherty, Patterson and Hitchmough unpublished data)	Chesterfield (near Hokitika)
'West Coast skink 2' (looks to be different to the skinks from the Hokitika area; Daugherty, Patterson and Hitchmough, unpublished data)	Reefton

[‡] Table constructed from information in Hardy (1977), Pickard and Towns (1988), Vos (1988), Patterson and Daugherty (1990, 1994), Daugherty *et al.* (1990b, 1994), Worthy (1991), SRARNZ Notes (1994, 1995a) and from G. B. Patterson (pers. comm.).

* almost certainly extinct.

[ƒ] See also McCann (1972).

[†] Daugherty *et al.* (1994) state that they believe this species to either be extinct or to be synonymous with another species.

[§] now regarded as a separate species (Daugherty *et al.*, 1994; G. B. Patterson, pers. comm.), but has not been formally recognised as such yet.

A large body of morphological, ecological and physiological data has now been accumulated for the New Zealand skinks (see Hardy, 1977 - p. 223, for a brief review of literature and Whitaker and Thomas, 1989 and SRARNZ Notes, 1993, 1994, 1995a and b, 1996 for bibliographies). However, molecular data is also required, for - as with the rest of New Zealand's herpetofauna - the skinks lack a fossil record and are morphologically very similar. In fact, while the skinks have been studied for over 150 years they are still recognised as a 'difficult' taxonomic group. Molecular studies involving allozyme electrophoresis have now been under way for some time (Hardy, 1977; Towns *et al.*, 1985; Vos, 1988; Daugherty *et al.*, 1990b; Patterson and Daugherty, 1990, 1994, unpublished data) and with this study, a preliminary investigation of most of the New Zealand skinks using sequence data becomes complete (see also Hickson, 1993).

Historical Taxonomy of the New Zealand Skinks and the Status of Cyclodina as a Genus.

As mentioned above, the New Zealand skinks have a long and troubled taxonomic history (Hardy, 1977) and while New Zealand is currently considered to have two genera of skinks (*Leiolopisma* and *Cyclodina*), the status of the genus *Cyclodina* was questioned until as late as 1980 (Robb, 1980).

'Throughout the 19th century and the first few years of the 20th, descriptions of new species and compilation of species lists for the country were emphasised' (Hardy, 1977). However, the descriptions were generally inaccurate and many type specimens were lost (McCann, 1955; Robb, 1973). Little was recorded about the habits, life history or ecology of any New Zealand lizard at this stage (Barwick, 1959). During this time, the first non-*Leiolopisma* New Zealand species was described (Gray, 1843) - *Tiliqua ornata* - now *C. ornata* (Gray 1843) (see Tables 1.2 and 1.3), and the name *Cyclodina* was first applied - to *C. aenea* Girard 1857 (see Tables 1.2 and 1.3; Hardy, 1977). Between 1906 and 1955, very little taxonomic work was done (Hardy, 1977) and no further non-*Leiolopisma* skinks were identified - however, the recognised ones were renamed many times (see Hardy, 1977 for synonymies).

In 1955, McCann published a revision of the New Zealand lizards (see also McCann, 1972). In this, he erected another of the species now recognised as belonging to the genus *Cyclodina* (*C. oliveri*; see Tables 1.2 and 1.3), but placed all of the species in *Leiolopisma* except for *Sphenomorphus pseudornatus* (which consisted of part of what we now know as *C. ornata* and part of *C. aenea* - see Table 1.3; Hardy, 1977). This division was based on an eyelid character - the divided or undivided condition of the lower eyelid (McCann, 1955; Hardy, 1977).

After McCann's 1955 revision (see also McCann, 1972), there was an upsurge in publications on various topics, including occurrence and distribution, zoogeography, histology, parasitology, life histories, oviparity, ecology, thermal relations and thermophysiology, for a variety of species (see Hardy, 1977 for review of literature). In 1970 and 1975, another two of the currently recognised *Cyclodina* species were diagnosed (*C. alani* and *C. macgregori)*, but were again placed in *Leiolopisma* (Robb, 1970, 1975; see Tables 1.2 and 1.3) and during this period, *Sphenomorphus pseudornatus* was reclassified as *Leiolopisma* by Greer (1974) and Robb (1974, 1975). However, it was noted at this time that certain species (most of those now recognised as *Cyclodina* species) formed a natural grouping within the *Leiolopisma* genus (Bull and Whitaker, 1975).

Hardy (1977), separated the New Zealand skinks into *Leiolopisma* and *Cyclodina* in his revision of the New Zealand skinks (see Table 1.3) and erected a sixth *Cyclodina* species (*C. whitakeri*) on the basis of both morphological data and electrophoretic studies of lactate dehydrogenase (LDH) isozymes and haem-containing proteins. He also reduced *L. pachysomaticum* (Robb 1975) to synonymy with *C. oliveri* (see Table 1.3). He noted that the eyelid character used by McCann (1955; among others) was known to be unreliable in some populations, but felt that this character was useful when used in conjunction with other features, and that there were enough differences between the two groups to warrant a separation into two genera (see Table 1.4).

Originally, there was some argument over this division (Robb 1980), mainly because of the limited number of differences which Hardy (1977) formally diagnosed as separating the two genera (see Table 1.4). Scale counts and body ratios overlap continuously throughout the New Zealand species so only species at opposite ends of the total range can be distinguished from one another using counts or ratios alone (Hardy, 1977).

Several osteological differences have been formally diagnosed since Hardy (1977)'s review (Worthy, 1991; see Table 1.4). And apart from the formally defined diagnostic characters, there do appear to be other differences. Hardy (1977) also commented that 'the skull is generally lower and more sharply pointed' in *Leiolopisma* but did not include this feature in his diagnosis for the two genera. *Cyclodina* also has a characteristic 'teardrop-shaped' mark under the eye (except for *C. aenea*) (Hardy, 1977). Interestingly, *L. homalonotum* has something similar (Hardy, 1977).

Table 1.3. Reproduced from Hardy, 1977, providing a cross-reference between the species recognised in his study and those recognised by McCann (1955; see also McCann, 1972) and Robb (1970, 1975). Taxa erected by Hardy (1977) are denoted by bold type.

Species recognised by Hardy	Species recognised by McCann and Robb
Leiolopisma fallai	*Leiolopisma fallai*
Leiolopisma suteri	*Leiolopisma suteri*
Leiolopisma smithi	*Leiolopisma smithi smithi*
	Leiolopisma smithi numerale
Leiolopisma homalonotum	*Leiolopisma homalonotum*
Leiolopisma gracilicorpus	*Leiolopisma homalonotum*
Leiolopisma moco	*Leiolopisma moco*
Leiolopisma infrapunctatum	*Leiolopisma infrapunctatum*
Leiolopisma striatum	*Leiolopisma latilinearum*
Leiolopisma zelandicum	*Leiolopisma ornatum*
Cyclodina ornata	*Sphenomorphus pseudornatus*
Cyclodina aenea	*Leiolopisma aeneum*
Leiolopisma nigriplantare nigriplantare	*Leiolopisma dendyi*
	Leiolopisma turbotti
Leiolopisma nigriplantare maccanni	*Leiolopisma zelandica*
Leiolopisma grande	*Leiolopisma grande grande*
Leiolopisma lineoocellatum	*Leiolopisma festivum*
Leiolopisma chloronoton	*Leiolopisma lineoocellatum*
Leiolopisma otagense	*Leiolopisma grande otagense*
	Leiolopisma grande waimatense
Leiolopisma acrinasum	Not recognised
Leiolopisma (?) *fasciolare*	Not listed
Cyclodina oliveri	*Leiolopisma oliveri*
	Leiolopisma pachysomaticum
Cyclodina whitakeri	Not recognised
Cyclodina alani	*Leiolopisma alani*
Cyclodina macgregori	*Leiolopisma macgregori*

Table 1.4. Formally defined similarities and differences between New Zealand *Leiolopisma* and *Cyclodina* species[†].

1. Characteristics shared by both genera:

a. Morphological (Hardy, 1977):
- movable lower eyelid
- lack supranasal scales (some overseas *Leiolopisma* have supranasals)
- well developed prefrontal scales
- frontoparietal scales distinct (some overseas *Leiolopisma* have fused frontoparietals)
- limbs well developed

b. Osteological (Worthy, 1991)*:
- lygosomine (have fused frontal bones)
- alpha palatal pattern
- 11 premaxillary teeth
- no postorbital bone
- meckelian canal completely overlapped by dentary

2. Similar/overlapping characteristics (Hardy, 1977 and Hudson, 1994)

Leiolopisma:
- small to relatively large (snout to vent length up to 150 mm)
- mostly terrestrial (*L. striatum* is likely to be more arboreal than most skinks; *L. moco* is sometimes found up manuka; *L. acrinasum*, *L. smithi* and *L. suteri* are known to fish in rock pools for invertebrates)
- clutch or brood size variable (up to ten)
- mostly diurnal

Cyclodina:
- small to medium-sized (snout to vent length up to 125 mm)
- terrestrial
- brood size variable (up to eight)
- crepuscular or nocturnal

3. Differences between the two genera

a. Morphological (Hardy, 1977):
- Lower eyelid character: *Leiolopisma* - lower eyelid with a transparent palpebral disc (a well-rounded, clearly defined, transparent and convex central scale)
  Cyclodina - lower eyelid scaly or at least covered by one to two large opaque scales not clearly differentiated from surrounding scales

..... continued next page

- Subocular scale row: *Leiolopisma* - suboculars interrupted by an enlarged supralabial scale (a few overseas species have a continuous row)
  Cyclodina - suboculars in a continuous series (except in *C. aenea* from Aorangi Island in the Poor Knights Islands group)
- External ear opening: *Leiolopisma* - fairly well developed, usually with one or more projecting scales on the anterior margin (*L. moco* and *L. n. maccanni* show a large amount of inter-population variation)
  Cyclodina - well developed, without projecting scales on the anterior margin
- Body in cross-section: *Leiolopisma* - oval
  Cyclodina - squarish
- Digits: *Leiolopisma* - long
  Cyclodina - shortened (particularly on front limbs)

b. Osteological (Worthy, 1991)*:

- Ridge on braincase (occipital capsule): *Leiolopisma* - well developed ridge extends over the entire prootic-supraoccipital suture
  Cyclodina - well developed ridge extends only part way to anterior end of prootic-supraoccipital suture
- Quadrate: *Leiolopisma* - quadrate without an anteriomedial ridge arising dorsally and aligned vertically
  Cyclodina - quadrate has pronounced anteriomedial ridge arising dorsally and aligned vertically
- Dentary: *Leiolopisma* - lower posterior notch on dentary extends anterior of upper notch
  Cyclodina - lower posterior notch on dentary does not extend anterior of upper notch

† Table constructed from information in Hardy (1977), Worthy (1991) and Hudson (1994).

* Worthy has investigated only a few *Leiolopisma* species and therefore his diagnostic characters for *Leiolopisma* should be treated with caution (G. B. Patterson, pers. comm.)

There also appear to be differences in habitat preferences and in activity cycles between the two genera - Hardy (1977) noted that 'although little is known about the behaviour and habitat requirements of the majority of New Zealand skinks, it appears that the two groups are in several respects ecologically separated. The non-*Leiolopisma* species appear to be crepuscular [active at dawn or dusk] or nocturnal, preferring a damper, more heavily vegetated environment. Although *L. suteri* is nocturnal ..... in general, New Zealand members of the [*Leiolopisma*] genus appear to be diurnal and to favour drier more open habitats.' Hardy (1979) also noted that *C. oliveri* males were heterogametic for chromosome pair six, while no *Leiolopisma* species studied showed this. The generic division is supported by the study of Vos (1988; see later) and is now generally accepted (G. B. Patterson, pers. comm.).

Cryptic Species among the New Zealand Skinks.

Classification of the New Zealand skinks has been made even more difficult by the presence of cryptic species in both genera. This is further complicated by the fact that some species show variable colourations depending not only on their habitat and geographical location but also on the presence of other skink species (Daugherty *et al.,* 1990b; D. R. Towns, pers. comm.; G. B. Patterson, pers. comm.). There also seem to be clines in morphological features, so that individuals at the extremes of the species range can be very different but overall, fit within the one species (D.R. Towns, pers. comm.; G. B. Patterson, pers. comm.; R. A. Hitchmough, pers. comm.) which confuses the issue further still. The use of molecular data, in conjunction with other datasets can aid in resolving these problems.

Several cryptic species have now been identified among the small brown skinks of the genus *Leiolopisma* and others have been proposed in both *Cyclodina* and *Leiolopisma* (see Table 1.2). Originally, when McCann revised *Leiolopisma* (1955; see also 1972), he included a species he called *L. zelandica*. In 1976, Gill used ecological and morphological data to identify a new species, sympatric with McCann (1955)'s *L. zelandica*, which he believed to be conspecific with the holotype of *Tiliqua zelandica*. On these grounds he stated that this species should actually be called *L. zelandicum*, but, as this name was unavailable, left it unnamed. In his 1977 revision, Hardy renamed McCann (1955)'s *L. zelandica*, calling it *L. nigriplantare maccanni*, and divided McCann (1955)'s *L. ornatum* into two species - *L. zelandicum* (Gill [1976]'s unnamed species) and *C. ornata* (part) (see Table 1.3).

Recently, more extensive allozyme electrophoresis revealed *L. n. maccanni* to be a complex of cryptic taxa (Daugherty *et al.,* 1990b; Patterson and Daugherty 1990, 1994). In 1990, Patterson and Daugherty formally identified this 'subspecies' as comprising three species and one subspecies: *L. maccanni*, *L. inconspicuum*, *L. notosaurus*, and *L. n. polychroma* (see Table 1.2).

They also identified another morphologically very similar, but unrelated species: *L. microlepis* (see Table 1.2). Hardy (1977) had examined specimens of *L. maccanni*, *L. notosaurus* and *L. n. polychroma*, and included them under *L. n. maccanni* (Patterson and Daugherty, 1990). However, *L. inconspicuum* and *L. microlepis* had not formerly been unequivocally identified (Patterson and Daugherty, 1990).

In 1994, Patterson and Daugherty formally identified another new species from the *L. n. maccanni* complex: *L. stenotis* from Stewart Island (see Table 1.2). Specimens of this species were available to Hardy, but were not included in his formal analysis (Patterson and Daugherty, 1994). Analysis of DNA sequences (Hickson *et al.*, 1992; Hickson, 1993) supports the identity of these new taxa. Since then, two more *Leiolopisma* species have been proposed on the basis of allozyme analysis (in conjunction with morphological data) and the existence of still more species seems likely (see Table 1.2).

With regards to *Cyclodina*, Hardy (1977) noted considerable morphological differences between some populations and the other members of the species to which they were assigned: the *C. oliveri* and *C. aenea* populations from the Poor Knights Islands and the *C. ornata* population from the Three Kings Islands. He also mentioned other populations which differed to some extent from the rest of their species: the *C. ornata* population from the Poor Knights Islands and the *C. macgregori* population from Sail Rock. However, studies of these populations (Hardy, 1977) using electrophoresis of haem compounds and LDH did not reveal any evidence supporting reclassification of these groups.

Hardy (1977) also felt that his electrophoretic data did not justify the existence of *L. pachysomaticum* (Mercury Islands and Alderman Islands; Robb, 1975) and reclassified individuals in this group as *C. oliveri*. However, he did note that the *C. oliveri* populations from the Hen and Chicken and Mokohinau Islands and from Little Barrier Island, along with the Poor Knights individuals, did appear somewhat different morphologically from those populations which had formerly comprised *L. pachysomaticum*. He also found that there was a small amount of difference (both in terms of morphology and haem electrophoresis) between the Little Barrier population and both the Poor Knights and Mercury Islands populations. He summarised these findings by commenting that further knowledge of range and variability among this species might support further taxonomic subdivision(s).

In 1988, a study of *Cyclodina* using both morphological and more detailed allozyme data (Vos, 1988) led to the proposal that at least two of the populations Hardy (1977) had noted as being different from other members of these species on morphological grounds were cryptic species - the *C. oliveri* population from the Mokohinau Islands and the *C. aenea* population from the Poor Knights Islands. Vos (1988)'s study - which is the most recent and complete study of *Cyclodina* to date - is summarised and discussed in detail later this chapter.

Hypotheses about the Origin(s), Divergence Times and Time(s) of Arrival of the New Zealand Skinks.

Because of the absence of a fossil record for the New Zealand lizards, the use of a 'molecular clock' based on molecular data is the only direct way of calculating divergence times for these taxa (Daugherty *et al.*, 1990b).

Without molecular data, the traditional view (Bull and Whitaker, 1975; see also Towns *et al.*, 1985) divided New Zealand's herpetofauna into two groups - Recent and Archaic. The Archaic group was believed to be of Gondwanan origin (80 million years ago [MYA] or more) and consisted of the tuatara (*Sphenodon punctatus*) and the *Leiopelma* frogs (Bull and Whitaker, 1975; see also Towns *et al.*, 1985). The skinks and geckos however, were regarded as Recent arrivals (Miocene/ Pliocene/ Pleistocene) by one or more transoceanic invasions (Bull and Whitaker, 1975; Hardy, 1977; see also Towns *et al.*, 1985) - presumably by rafting (Bull and Whitaker, 1975; Hardy, 1977). However, this assumption left unresolved whether earlier species were replaced by the new arrivals (as lizards have been in existence for much longer than 80 million years).

At this time, the geckos were thought (Bull and Whitaker, 1975) to have arrived via New Caledonia in the Miocene, about 20 MYA, and to have subsequently diverged in New Zealand. In the case of the skinks, *Leiolopisma* was believed (Bull and Whitaker, 1975; Hardy, 1977) to be related to certain New Caledonian and Australian species and to have arrived by the late Pliocene (approximately five MYA), although it was suggested that *L. suteri* may have arrived more recently (Bull and Whitaker, 1975). *Cyclodina* was suggested to be more closely related to New Caledonian (Hardy, 1977; Bull and Whitaker, 1975) or Australian (Bull and Whitaker, 1975) species than to *Leiolopisma* and to have arrived at the same time as *Leiolopisma* or somewhat later, in the early Pleistocene (up to two and a half MYA; Bull and Whitaker, 1975; Hardy, 1977). Because the skinks were generally regarded as Pliocene arrivals, the identification of distinctive elements within the fauna tended to result in hypotheses involving several invasions to account for their presence (Towns *et al.*, 1985; for example, see Towns, 1974; Hardy, 1977).

In recent years, molecular data - allozymes (Towns *et al.*, 1985; Daugherty *et al.*, 1990b; Patterson and Daugherty, 1990, 1994 and unpublished data) and DNA sequences (Hickson *et al.*, 1992; Hickson, 1993) - has been used to investigate the New Zealand skinks and suggests (Hickson, 1993) that the genetic diversity of the New Zealand *Leiolopisma* is considerably greater than their morphological diversity and that the skinks might therefore be much older than previously suggested. Based on allozyme analysis of several *Leiolopisma* species (Towns

et al., 1985; Daugherty *et al.*, 1990b) and the increasing number of recognised New Zealand *Leiolopisma* species, Towns *et al.* (1985) and Daugherty *et al.* (1990b) both suggested a Miocene time of arrival (like the geckos) for *Leiolopisma* at least. Daugherty *et al.* (1990b) applied the commonly used calibration of Maxson and Maxson (1979; D = 1 accumulates every 14 million years) to their allozyme data and found that this indicated divergence dates among the skinks studied of 10 - 14 MYA. Vos (1988) used the same calibration to estimate the time of divergence between *Cyclodina* and *Leiolopisma* based on her allozyme studies (see later) and found that it appeared to be in the region of 19 MYA (Miocene).

However, as there is no fossil record, the molecular clock cannot be independently calibrated - which is required due to lineage-dependant rates of evolution (Daugherty *et al.*, 1990b; Hickson, 1993). With allozyme data, the range of calibrations used by different workers varies by a factor of 22 (Daugherty *et al.*, 1990b). This factor is almost certainly in error and more work is required to eliminate the more extreme values. It does however, demonstrate one of the advantages sequences have over allozymes in the examination of genetic diversity - the properties of nucleotides are better characterised than those of allozymes and while sequences also display lineage-dependant rates of evolution, the rate estimates can be more accurate than for allozyme data. In addition, only a low number of electrophoretic loci could be sampled for the New Zealand skinks, meaning the results may be less reliable than results based on sequence data (Nei *et al.*, 1983; Nei, 1987; Hickson, 1993). Therefore the above estimates of divergence times must be treated with caution, and the possibility that the present diversity of the New Zealand skinks is of recent origin cannot be dismissed (Daugherty *et al.*, 1990b).

Towns *et al.* (1985) suggested that if the skinks did arrive in the Miocene, long distance oceanic dispersal may not necessarily have occurred - as New Zealand and New Caledonia retained their land links until near the beginning of the Miocene period (27-37 MYA; Stevens, 1980). Therefore common ancestry of the New Zealand lizard fauna through direct land links would be possible instead. Daugherty *et al.* (1990b) did not discuss origin(s) or mechanism(s) of arrival. A recent review of hypotheses for skink evolution and dispersal is presented by Bauer (1993), but this does not discuss the New Zealand skinks in detail.

The most recent, and simplest, hypothesis is that the skinks have a Gondwanan origin (Hickson, 1993). As mentioned previously, Hickson (1993) obtained sequence data for members of most New Zealand *Leiolopisma* species and for one individual from each of *C. aenea*, *L. telfairi* (Mauritius), *Lampropholis guichenoti* (a related Australian skink species) and *Tropidoscincus rohssii* (a New Caledonian skink species from a related genus). While a lineage-specific rate of mtDNA evolution for the skinks has not yet been calculated, he found no evidence to suggest that the obtained *Leiolopisma* sequences are changing any faster than other vertebrate groups

for which mtDNA rates have been calculated (see Hickson, 1993) and on this basis, analysis of the obtained data suggested that the *Leiolopisma* species diverged in excess of 15 MYA. He then reviewed the available allozyme data (Daugherty *et al.,* 1990b; Patterson, G. B. and Daugherty, C. H., unpublished data), immunological data (Hutchinson *et al.,* 1990), biogeographic information (see Hickson, 1993) and sequence data (Hickson, 1993) and proposed three hypotheses for the origins and time of arrival of the New Zealand *Leiolopisma.*

Each of these hypotheses involves a Gondwanan origin for the skinks - but they differ in time of arrival and divergence of the New Zealand taxa. The first proposes a Gondwanan origin and divergence of the New Zealand *Leiolopisma* - followed by Gondwanan break up approximately 80 MYA (Cretaceous) and subsequent continental drift. The second proposes that the New Zealand *Leiolopisma* resulted from diversification within New Zealand during the Miocene (15-25 MYA) from one or a few taxa which survived the Oligocene (36-25 MYA), in response to increasing land area and habitat diversity (the Oligocene Drowning theory - see Cooper and Cooper, 1995). In this case, the common ancestor(s) may have been present in New Zealand since the Gondwanaland break up, or reached New Zealand before the Oligocene. The third involves diversification some time after the break up of Gondwanaland followed by several independent colonisations of New Zealand, either via pre Miocene land links (see Stevens, 1980) or by some mechanism of transoceanic travel.

Hypotheses for the Current Distribution Patterns of the New Zealand Skinks, particularly Cyclodina.

As mentioned earlier, *Leiolopisma* species are found throughout New Zealand while *Cyclodina* species are restricted to the North Island and northern offshore islands (see Table 1.2; Pickard and Towns, 1988). McCann (1955; see also McCann, 1972) identified two main faunal breaks in the distribution patterns of the New Zealand lizards (see also Bull and Whitaker, 1975). The first is formed by Cook Strait and Bull and Whitaker (1975) noted that if the '*Sphenomorphus*' species (now *Cyclodina*) were in New Zealand during the last glaciation, the only suitable habitat for them (based on ecology) would have been north of Auckland, with subsequent dispersal south interrupted by this barrier. The second is in the central North Island and affects *Leiolopisma* species only. Bull and Whitaker (1975) suggested that this second break may have originated during the Pleistocene. With regards to specific patterns of distribution, Bull and Whitaker (1975) described three categories - widespread, restricted and relict.

Hardy (1977) believed that the New Zealand *Leiolopisma* formed three main geographic groups - a northern group, a group found in the south of the South Island and a group centred around the Cook Strait area. He suggested that this distribution pattern was the result of three main

invasions and of Pleistocene glaciations. With regards to *Cyclodina* he felt that his electrophoretic analyses indicated three main lines of divergence from the ancestral stock and suggested that 'The initial divergence of *Cyclodina* species apparently occurred in the northern half of the North Island, with subsequent dispersal southward. The majority of species have probably dispersed over much of the North Island in Recent times, although this is now particularly evident only for *C. ornata* and *C. aenea*. It is thought that such dispersal must have been variously interrupted by fluctuating Pleistocene conditions, Cook Strait becoming an important Recent barrier to continued southward movement'.

Towns *et al.* (1985) reviewed Hardy (1977)'s biogeographic regions in the light of new distributional information and suggested that they were of doubtful validity. Upon examining all available distributional data in conjunction with ecology, they defined the New Zealand skinks as falling into two categories: widely spread and restricted (either as relicts, local endemics or by specific ecological requirements).

Towns *et al.* (1985) defined twelve species of lizards as widely distributed, including *C. ornata* and *C. aenea* which are found throughout the North Island and on many of its offshore islands. *C. alani* and *C. oliveri* were defined as 'north-eastern island relicts' - found only on north-eastern islands (from the Three Kings Islands in the north to East Island in the south). *C. whitakeri* and *C. macgregori* were described as 'disjunctive relicts' - found on north-eastern islands and islands off the south-western coast of the North Island (localities 500 km apart). *C. whitakeri* is also found at one location on the mainland (Hardy, 1977; Pickard and Towns, 1988). *C. alani*, *C. oliveri*, *C. whitakeri* and *C. macgregori* do not appear to be true island endemics, but instead appear to be relict populations of once widely distributed species (Towns *et al.*, 1985; Worthy, 1987c).

This odd distribution pattern would seem to be the result of predation, particularly by rodents (for example, see Bull and Whitaker, 1975; Hardy, 1977; Towns *et al.*, 1985; Worthy, 1987c; Towns, 1991). In other countries, lizard predators include many kinds of snakes, birds, small mammals and even other lizards. As New Zealand has no indigenous predatory mammals, land-dwelling snakes or very large lizards, for a long time the only lizard predators were a few species of birds - including raptors (such as the harrier hawk and the bush hawk), kingfishers, rails (including the weka), owlet-nightjars, owls (including the morepork), the red-billed gull, and the pukeko (Barwick, 1959; Bull and Whitaker, 1975; Worthy, 1987c) - and the tuatara (Bull and Whitaker, 1975). These predators probably had a negligible effect on lizard numbers (Barwick, 1959; Bull and Whitaker, 1975; Worthy, 1987c), although there is strong circumstantial evidence that weka can 'impact severely on lizard faunas' (SRARNZ Notes 1995a).

Over the last 1000 years, however, man has introduced a variety of rodents (*Mus musculus*, *Rattus rattus*, *R. norvegicus* and *R. exulans*), mustelids (*Mustela erminea, M. nivalis* and *M. putorius*) and also feral and domestic cats (*Felis catus* and *F. domesticus*) (Barwick, 1959; Bull and Whitaker, 1975; Hardy, 1977; Towns *et al.,* 1985; Worthy, 1987b and c). Indeed, the kiore (*R. exulans*) could have been here for as long as 2000 years (T. H. Worthy, pers. comm.).

Bull and Whitaker (1975) briefly discussed the effect of the kiore on lizards, and Hardy (1977) suggested that the absence of some *Cyclodina* species from the mainland and from many of the offshore islands was probably due to predation by this rodent, as did Worthy (1987c). Reviews of lizard distribution data by Towns *et al.* (1985) and Towns (1991) confirmed that where kiore are present, the habitat range of most lizards (including *C. aenea*) is reduced, and that the large, nocturnal, ground-dwelling *Cyclodina* species (*C. oliveri*, *C. alani*, *C. whitakeri* and *C. macgregori*) are almost completely incompatible with rodents.

Establishing a reasonable classification for the New Zealand skinks and using further molecular data to test the various hypotheses on origins and arrival times should aid in clarifying how (and whether) the New Zealand skinks diverged in this country. More detailed population studies at the molecular genetic level are also required. The effects of factors such as glaciation, 'drownings', natural barriers and predation on this divergence to produce the observed distribution patterns could then be evaluated.

Cyclodina.

As mentioned above, *C. oliveri*, *C. alani*, *C. whitakeri* and *C. macgregori* are large (total length between 220 and 255 mm; Hudson, 1994), nocturnal and ground-dwelling (Towns *et al.*, 1985; Hudson, 1994). The latter three are considered to be rare, vulnerable or endangered (Daugherty *et al.,* 1994; see also SRARNZ Notes, 1995b). *C. alani* is New Zealand's heaviest skink - weighing 30-50 grams (Towns, 1991; Hudson, 1994). *C. aenea* is the smallest of the New Zealand skinks (total length approximately 130mm; Towns *et al.*, 1985; Hudson, 1994). It has a wide range of habitats, but is principally found in forested areas (Towns *et al.,* 1985).

Biochemical, Morphological and Osteological Data for Cyclodina (Vos, 1988).

Vos (1988) investigated members of each of the *Cyclodina* species using biochemical (allozyme), morphological and osteological data. These included individuals from some of the populations mentioned earlier as potential cryptic species: *C. oliveri* from the Mokohinau and the Poor Knights Islands and *C. aenea* from Aorangi Island (the Poor Knights Islands). She also examined one individual from each of *L. nigriplantare* (presumably *L. nigriplantare maccanni, sensu* Hardy 1977), *L. zelandicum* and *L. telfairi*. Her results are summarised as follows.

Allozyme Electrophoresis.

Vos (1988) investigated allozyme variation at 17 loci. One of these showed no change for both *Leiolopisma* and *Cyclodina*, while one was unvarying for the *Cyclodina* genus. She found that only three of the 16 *Cyclodina* populations examined (see Figure 1.1) showed any heterozygosity (although she notes that this was probably due to the small sample sizes used); that polymorphism appears to be low within species; and that apart from the three heterozygotes, all polymorphism observed was in the form of fixed allelic differences between the populations. Fixed allelic differences are alleles at allozyme loci which are unique to each species and which result from lack of interbreeding between species: one fixed allelic difference is sufficient to show that sympatric populations are not interbreeding (Vos, 1988). With allopatric species, it must be decided if they would interbreed if they occurred sympatrically. Generally, if allopatric populations show a greater genetic distance from each other than that shown by the two most closely related species occurring sympatrically, they are considered to be different species (Vos, 1988).

Only one population of *C. ornata* was available. Allozyme analysis showed that it was separated from all the other *Cyclodina* species by one fixed difference (this is a specific difference found when *C. ornata* is compared to all of the other *Cyclodina* species; there will also be additional specific differences between *C. ornata* and each of the other species). *C. ornata* appeared to be most closely related to *C. oliveri*, with only two fixed differences in total between these species (except the Mokohinau Islands *C. oliveri* population, from which *C. ornata* is only separated by one fixed difference in total).

Two populations of *C. alani* were examined, no intra-specific genetic variation was found and they were separated from all the other *Cyclodina* species by three fixed differences. *C. alani* appears to be most closely related to *C. macgregori*, with only five fixed differences between the two species in total.

Only one population of *C. macgregori* was available. This was separated from all the other *Cyclodina* species by one fixed difference and appeared to be most closely related to *C. alani* (see above).

Three populations of *C. whitakeri* were examined and a small amount of intra-specific genetic variation was found (two heterozygotes). These populations were separated from all the other *Cyclodina* species by one null allele and one fixed difference, and appeared to be most closely related to *C. oliveri* with only four fixed differences in total between the two species (except the Mokohinau Islands population, from which the *C. whitakeri* populations are separated by only two fixed differences in total).

Four populations of *C. oliveri* were investigated - all were separated from all the other *Cyclodina* species by one fixed difference and appeared to be most closely related to:

1. *C. ornata* (see above).
2. *C. whitakeri* (see above).

Intra-specific genetic variation was apparent - the population from the Mokohinau Islands was separated from the other *C. oliveri* populations by three fixed differences.

Five populations of *C. aenea* were examined - all were separated from all the other *Cyclodina* species by three fixed differences - but at one of these, the Poor Knights Islands population and the other *C. aenea* populations carry different alleles. All except the Poor Knights Islands population were separated by five fixed differences in total. Intra-specific genetic variation was apparent - the population from the Poor Knights Islands was separated from the other *C. aenea* populations by four fixed differences. *C. aenea* does not appear to very closely related to any of the other *Cyclodina* species.

Analysis of the Resulting Allozyme Data.

Vos (1988) used both weighted pair-group method with arithmetic average (WPGMA; Sneath and Sokal, 1973) clustering of Nei's unbiased genetic distance (D) (Nei, 1978) and phylogenetic analysis under parsimony (PAUP; Swofford, 1985) to investigate the obtained allozyme data. *L. telfairi*, *L. zelandicum* and *L. nigriplantare* were used as outgroups. Three equally parsimonious trees were obtained from the parsimony analysis (Vos, 1988). Derived character states (apomorphies) and repeated or parallel changes (homoplasies) were the same in all three trees, and all three were identical with respect to the pattern of inferred relationships and differed in branch length only slightly. Therefore Vos (1988) treated them as the same for discussion purposes.

The WPGMA and parsimony trees presented by Vos (1988) are almost identical (see Figures 1.1 and 1.2). Both support the differentiation of eight distinct groups among the *Cyclodina* species - four representing *C. ornata*, *C. whitakeri*, *C. alani* and *C. macgregori*, two representing *C. aenea* and two representing *C. oliveri*. Of the two *C. aenea* lineages, one represents the *C. aenea* population from the Poor Knights Islands while the second represents the remaining *C. aenea* populations. The two *C. oliveri* lineages consist of the *C. oliveri* population from the Mokohinau Islands, and the remaining *C. oliveri* populations. Both trees also suggest three main groups of species within *Cyclodina* :

- *C. ornata*, *C. oliveri*, the *C. oliveri* population from the Mokohinau Islands and *C. whitakeri*.
- *C. aenea* and the *C. aenea* population from the Poor Knights Islands.
- *C. alani* and *C. macgregori*.

The only differences lie in the divergence of *C. alani* and *C. macgregori* after *C. aenea* in the WPGMA tree (Figure 1.1) and before *C. aenea* in the parsimony tree (Figure 1.2), and in *L. nigriplantare* being more closely related to *L. telfairi* than to *L. zelandicum* in the parsimony tree, as opposed to *L. nigriplantare* and *L. zelandicum* grouping together in the WPGMA tree. Vos (1988) comments that 'this just highlights the different assumptions used in the construction of the two different trees'.

These analyses support the validity of the six *Cyclodina* species recognised by Hardy (1977). Each species is defined by at least one fixed allelic difference in the WPGMA tree (Figure 1.1), while there is at least one apomorphy between taxa in the parsimony tree (Figure 1.2). Allopatric populations within species show high genetic similarity, except for one *C. oliveri* population and one *C. aenea* population.

The *C. aenea* population from the Poor Knights Islands and the *C. oliveri* population from the Mokohinau Islands form separate lineages from their assigned species in both trees and show a high number of fixed differences both from other populations in their given species, and from the other *Cyclodina* species. Considering the results of both trees, the *C. oliveri* clusters were separated by three fixed allelic differences and three apomorphies, while the two *C. aenea* clusters were separated by four fixed differences and three apomorphies. In addition, the value of D between the *C. aenea* population from the Poor Knights Islands and the other *C. aenea* populations and between the *C. oliveri* population from the Mokohinau Islands and the other *C. oliveri* populations both exceed the value of D between the two most closely related sympatrically occurring species. The two *C. oliveri* lineages were separated by a D value of 0.21, the two *C. aenea* lineages were separated by a D value of 0.27, while the two most closely related sympatrically occurring species (*C. ornata* and *C. oliveri*, with the exception of the Mokohinau Islands population) were separated by a D value of 0.16. On the basis of these differences, Vos (1988) proposed the *C. aenea* (Poor Knights) and *C. oliveri* (Mokohinau Islands) populations to be separate species.

Both trees support a monophyletic origin for *Cyclodina* with regards to *Leiolopisma*. The *Cyclodina* species are separated from the *Leiolopisma* species by five characters in the parsimony tree, two fixed allelic differences in the WPGMA tree and a D value of 1.4. As mentioned earlier, Vos (1988) used Maxson and Maxson (1979)'s calibration (D = 1 accumulates every 14 million years) to estimate a possible divergence date of 19 MYA between *Cyclodina* and *Leiolopisma* from this D value.

Figure 1.1 WPGMA tree from Vos (1988) showing the inferred relationships between the New Zealand *Cyclodina* species (16 populations), *Leiolopisma nigriplantare* (presumably *L. nigriplantare maccanni*, *sensu* Hardy 1977), *L. zelandicum* and *L. telfairi*. Nei (1978)'s unbiased genetic distances were used.

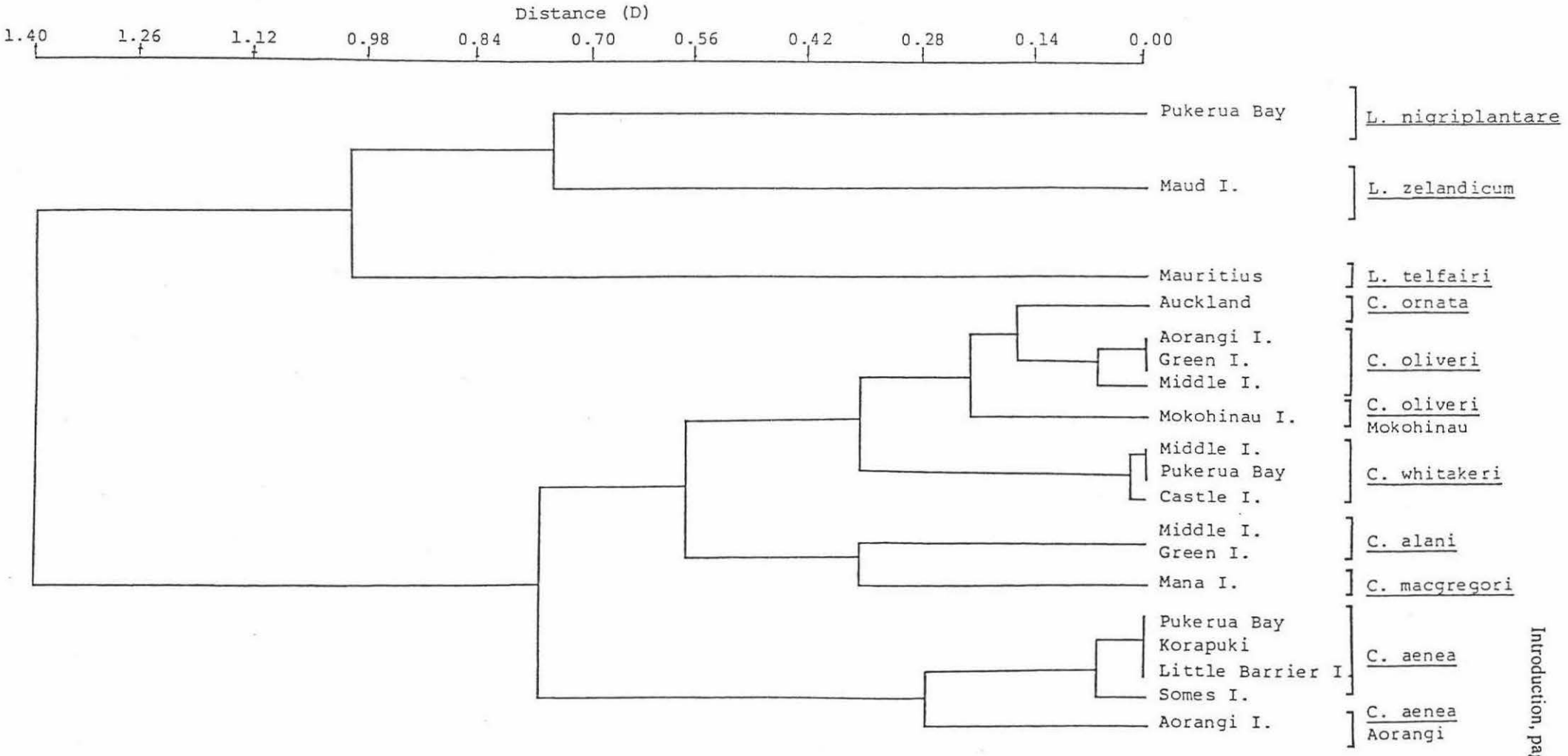
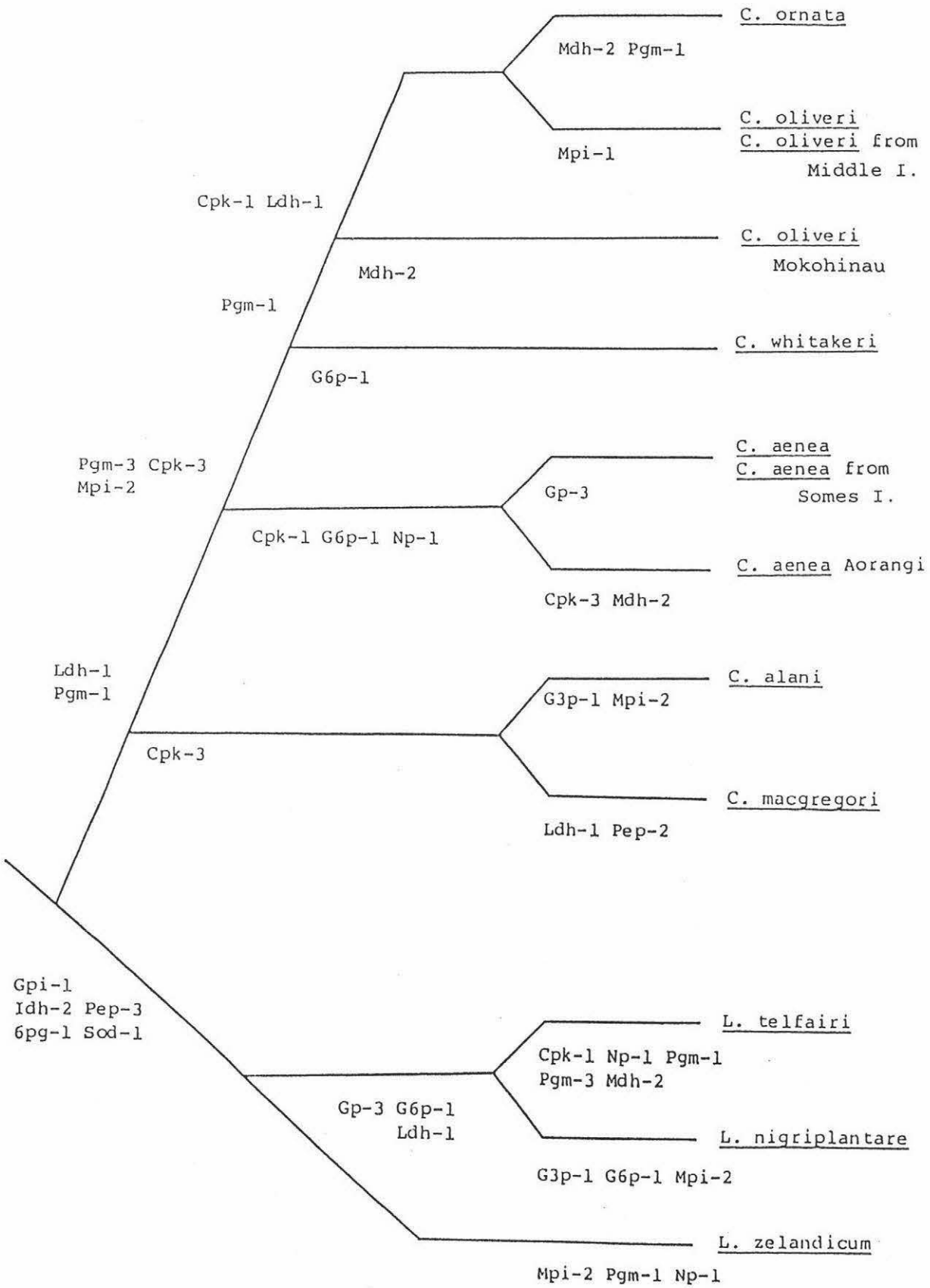


Figure 1.2 Parsimony tree from Vos (1988) showing the inferred relationships between the New Zealand *Cyclodina* species, *Leiolopisma nigriplantare* (presumably *L. nigriplantare maccanni, sensu* Hardy 1977), *L. zelandicum* and *L. telfairi*. Apomorphies are shown.



Morphology.

At the species level, the skinks of the *Cyclodina* genus are remarkably similar morphologically (Vos, 1988). However, some populations are markedly different from the rest of their species (Hardy, 1977; Vos, 1988). While molecular data has been shown to be of value in distinguishing different species among the New Zealand skinks (for example, see Vos, 1988; Daugherty *et al.,* 1990b), it is of no use to workers in the field, thus it is desirable to be able to distinguish individuals into species using morphological characters as well.

Vos (1988) used various types of statistical analysis (including discriminant analysis) to examine a large number of morphological characters. She found that while species can be distinguished morphologically, several characters have to be used - with the most informative being the numbers of midbody scales, ventral scales and subdigital lamellae; position of the ear opening; adult body size; and 'teardrop' coloration under the eye. While Vos (1988) did not included colour pattern (ventral and sides) in the discriminant analysis, she noted that it is probably the best way to distinguish the *Cyclodina* species morphologically. Vos (1988)'s findings on morphology are summarised below.

C. alani and *C. macgregori* are best discriminated from the other *Cyclodina* using the position of the ear opening (set further back than in the other *Cyclodina*), high number of midbody scales and large adult body size. To distinguish between them, subdigital lamellae and body colour pattern (sides and belly) are the most useful characters as *C. alani* has a distinctive colour pattern and a smaller number of subdigital lamellae than *C. macgregori.*

C. oliveri and *C. whitakeri* are very similar morphologically, and can really only be distinguished from each other and from the other *Cyclodina* species on the basis of their colouration (ventral and sides). The *C. oliveri* population from the Mokohinau Islands can also be distinguished on the basis of colouration. In addition, *C. oliveri* (Mokohinau Islands) individuals have a higher midbody scale row number than *C. whitakeri*, but while they also have a higher scale count than many *C. oliveri*, this is not conclusive evidence for distinguishing the two. The body pattern of the *C. oliveri* population from the Mokohinau Islands is shared by the *C. oliveri* populations from Little Barrier and the Hen and Chicken Islands.

C. aenea and *C. ornata* are again very similar. They can be distinguished from the other *Cyclodina* by their small adult body size and low midbody and ventral scale row counts. The main ways of distinguishing between them included colour pattern, the fact that *C. aenea* (including the *C. aenea* population from the Poor Knights Islands) lacks the 'teardrop'

colouration under the eye observed in all the other *Cyclodina*, while the *C. aenea* population from the Poor Knights Islands displays an interrupted subocular scale row not seen in any of the other *C. aenea* populations, nor in any of the other *Cyclodina* species. Individuals in this population also have higher numbers of midbody and ventral scales than those in the other *C. aenea* populations.

Vos (1988) also noted that more specimens needed to be examined before a definitive morphological key could be constructed for members of the *Cyclodina* genus, as the high range of variation within each species needed to be investigated more thoroughly.

Osteology and Scale Microarchitecture.

Vos (1988) also investigated the New Zealand *Cyclodina* using osteological features and scale microarchitecture. She found that each species appears to have some unique characteristics, but the sample size was too small to determine whether these reflect true inter-specific differences, or whether they fall within the normal range of intra-specific variation. These characteristics include a possible generic difference in scale surface microarchitecture between *Cyclodina* and *Leiolopisma* (*Cyclodina* species having a corrugated inner area on the scale surface which is not observed in *Leiolopisma* species); vertebrate numbers (*Leiolopisma* species appear to have higher numbers, but again a larger ample size is necessary to determine whether this is a true generic difference); and the dentary (shape is genus-specific; see also Worthy, 1987c; Worthy, 1991).

Taxonomic Status and Relationships of the New Zealand Cyclodina.

Vos (1988) concluded that the six *Cyclodina* species recognised by Hardy (1977) are valid. She found that intra-specific allozyme variation in *Cyclodina* is low, while inter-specific variation is high; that all *Cyclodina* species are differentiated by at least one fixed allelic difference and one apomorphy; and that morphologically, the different species are very similar but identifiable when several characters are used.

Based on both morphological and allozyme data examined in her study, Vos (1988) proposed the *C. oliveri* population from the Mokohinau Islands and the *C. aenea* population from the Poor Knights Islands to be cryptic species. She noted that further work needed to be done on other *Cyclodina* populations which Hardy (1977) recognised as being morphologically different from other members of their species, namely the *C. ornata* populations from the Poor Knights Islands (Aorangi Island) and the Three Kings Islands, and the *C. oliveri* populations from Little Barrier and the Hen and Chicken Islands (which are very similar morphologically to the *C. oliveri* population from the Mokohinau Islands; Vos, 1988).

Morphologically, Vos (1988) found the validity of *L. pachysomaticum* (Robb, 1975) to be unclear, with the representatives examined falling into the lower range of that seen in *C. oliveri*. She commented that without examining more specimens, the validity of this species on morphological grounds could not be determined. However, allozyme analysis indicated that there was no difference between the population of *C. oliveri* from Green Island, Mercury Islands (formerly part of *L. pachysomaticum*) from the other *C. oliveri* populations examined (except the one from the Mokohinau Islands) and on the basis of this support, Vos (1988) agreed with Hardy (1977) that *L. pachysomaticum* was not a valid species.

On morphological grounds, *C. aenea*, the *C. aenea* population from the Poor Knights Islands and *C. ornata* are similar, as are *C. alani* and *C. macgregori* (Vos, 1988). The *C. oliveri* population from the Mokohinau Islands, the other *C. oliveri* populations and *C. whitakeri* also show morphological similarities (Vos, 1988).

Hardy (1977) had suggested that close phylogenetic relationships exist between *C. ornata* and *C. aenea* and between *C. oliveri*, *C. alani* and *C. whitakeri* based on his analysis of LDH isozymes. Vos (1988)'s analysis of her allozyme data suggests that the *Cyclodina* species form three groups of closely related species which are somewhat different to those proposed by Hardy (1977). The first group contains *C. ornata*, *C. oliveri*, the *C. oliveri* population from the Mokohinau Islands and *C. whitakeri* - with *C. ornata* and the majority of the *C. oliveri* taxa (not the Mokohinau Islands population) forming the closest relationship within this group. The second group consists of *C. alani* and *C. macgregori*, while *C. aenea* and the Poor Knights *C. aenea* population form the third (see Figures 1.1 and 1.2).

Vos (1988) felt that Hardy (1977)'s separation of *Cyclodina* from *Leiolopisma* was well-supported by her allozyme data, which showed the presence of two fixed allelic differences between the two genera. Phylogenetic analysis of this data gave a Nei's distance of 1.4 between the genera and supported a monophyletic origin for *Cyclodina* with regards to *Leiolopisma*. She also found that the five morphological differences defined by Hardy (1977) and the genus-specific shape of the dentary (Worthy, 1987c; see also Worthy, 1991) are best for differentiating between *Leiolopisma* and *Cyclodina* (see Table 1.4).

However, Hardy (1977) examined only two presumed loci and used very small sample sizes. Vos (1988) also used a fairly low sample size and while she used a much larger number of allozyme loci than Hardy (1977), the accuracy of trees inferred from less than 30 allozyme loci is low (Nei *et al.*, 1983; Nei, 1987). By investigating these relationships using sequence data and the Hadamard conjugation (Hendy and Penny, 1989, 1993; Penny *et al.*, 1992, 1993a; Lento *et al.*, 1995), the resolution of the resulting phylogenies and the confidence that can be placed in them can be more directly assessed.

Hypotheses about the Origin(s), Divergence Times and Time(s) of Arrival of the New Zealand Cyclodina Species.

In this section, three hypotheses for the origin(s) of *Cyclodina* in New Zealand are presented. Again, the simplest explanation is that the genus has been here since the break-up of Gondwanaland (80 MYA). New Zealand is a Gondwanaland fragment and contains many recognised ancient lineages, such as the leiopelmatid frogs, the ratites and the tuatara, among others (Daugherty *et al.,* 1993). As mentioned earlier, Hickson (1993) reviewed various datasets and found some support for a Gondwanan origin of the New Zealand *Leiolopisma*. The suggested close relationship between *Cyclodina* and *L. alazon* from Fiji (Zug, 1985) may suggest that the theory of a Gondwanan origin for the New Zealand *Cyclodina* is correct - as Fiji is also a continental remnant (Hickson, 1993). In addition, the one *Cyclodina* individual examined by Hickson (1993) appeared to be closely related to *L. telfairi* (Hickson, 1993; see also Chapter Four) providing some support for this theory. If this hypothesis is correct, one would expect analysis of further sequence data, and of other datasets such as immunological data, to suggest that *Cyclodina* is as closely related to overseas taxa as to the New Zealand *Leiolopisma*.

A second hypothesis is that *Cyclodina* diverged from *Leiolopisma* within New Zealand. Under this theory, the time of divergence of the New Zealand *Cyclodina* would obviously depend on the times of arrival/divergence of *Leiolopisma* in New Zealand which is presently unclear (see earlier; Hickson, 1993). The genus *Cyclodina* consists of a smaller number of species and has a more limited distribution than *Leiolopisma*, and until recently was believed to be unique to the North Island and the northern offshore islands of New Zealand. This would make the suggestion that the genus had evolved from *Leiolopisma* within New Zealand plausible. However, the postulated close relationship with *L. alazon* (Zug, 1985) and the reclassification of *L. lichenigerum* from Lord Howe and Norfolk Islands into *Cyclodina* (Cogger, 1986) do not support this theory. The status of *L. lichenigerum* is presently unclear though, as it has now been placed in *Pseudemoia* (Cogger, 1992 - not sighted; G. B. Patterson, pers. comm.) and indeed it may be more closely related to the New Caledonian species than to either the New Zealand *Leiolopisma* or *Cyclodina* (D. R. Towns, pers. comm.). Allozyme analysis to date provides some support for a divergence from *Leiolopisma* in New Zealand (G. B. Patterson, pers. comm.). In addition, the *Cyclodina* individual examined by Hickson (1993) also appeared to be closely related to the New Zealand *Leiolopisma* species *L. fallai* (Hickson, 1993; see also Chapter Four). If this theory is correct, one would expect analysis of further sequence and immunological data to indicate that *Cyclodina* is more closely related to the New Zealand *Leiolopisma* than to all overseas taxa.

A third possibility is that *Cyclodina* arrived independently of *Leiolopisma*. The postulated relationship with *L. lichenigerum* from Lord Howe and Norfolk Islands would support this theory, but as mentioned, the status of this species is presently unclear. At the moment, there is very little sequence data available with which to investigate this theory, and from where *Cyclodina* might have originated if this hypothesis is correct, as Hickson (1993) did not examine *L. lichenigerum* nor any Australian or New Caledonian skinks of the genus *Leiolopisma*. Phylogenetic analysis of the *La. guichenoti* and *Tropidoscincus rohssii* sequences (Hickson, 1993) provided some support for a close relationship between *L. infrapunctatum* and *La. guichenoti*, while the closest relationship suggested by sequence analysis for *T. rohssii* was with *L. fallai* (but this relationship was not strongly supported). If this is the correct hypothesis, one would expect the *Cyclodina* species to be more closely related to certain overseas taxa than to the New Zealand *Leiolopisma* and analysis of further sequence data and of other datasets should indicate this. Again, if a pre-Miocene/Miocene time of arrival is indicated, *Cyclodina* may have arrived via land links; a post-Miocene time of arrival would require some form of transoceanic travel.

Aims of this Thesis.

1. To obtain DNA sequence from the mitochondrial 12S ribosomal RNA gene for one individual from each of the six currently recognised *Cyclodina* species and for one individual from each of two populations whose taxonomic status within this genus is presently unclear ('*C. aenea*' from Aorangi Island, Poor Knights group and '*C. oliveri*' from the Mokohinau Islands group).

2. To analyse this sequence data and to examine the resulting phylogeny(ies) with respect to the amount of support in the data for the current taxonomic status of the recognised New Zealand *Cyclodina* species and the validity of the two proposed *Cyclodina* species (Vos, 1988) and to evaluate how much confidence can be placed in the pattern of inferred relationships.

3. To analyse the *Cyclodina* sequence data in conjunction with the sequence data of Hickson (1993) for members of the New Zealand *Leiolopisma* and three overseas skinks, and to evaluate the amount of support in the data for a monophyletic origin for the *Cyclodina* species and for the various hypotheses regarding the origin(s) of the New Zealand *Cyclodina* species.

4. To compare the resulting phylogeny(ies) and theories with the theories drawn from other datasets available for the genus (such as allozyme data).

A unique DNA sequence was successfully obtained for each of the eight individuals examined. Comparison to the *Leiolopisma* sequences obtained by Hickson (1993) indicated that these did indeed represent *Cyclodina* 12S rRNA sequences. The sequences were analysed using the Hadamard conjugation (Hendy and Penny, 1989, 1993; Penny *et al.*, 1992, 1993a; Lento *et al.*, 1995) - a method of spectral analysis which provides a quantitative measure of the resolution of the data and thus an indication of the reliability of the resulting phylogeny - and the maximum parsimony, minimum evolution and maximum likelihood options of a new test version of PAUP (phylogenetic analysis under parsimony; see Swofford, 1985, 1993) - PAUP * 4.0 (1995,1996).

A single resolved tree was not produced, indicating that the *Cyclodina* species may have rapidly diverged from each other. However, the six currently recognised *Cyclodina* species and the two proposed species (Vos, 1988) do form separate lineages - agreeing with the allozyme analysis of Vos (1988) and supporting the taxonomic status of these taxa. Furthermore, while sequence analysis indicates that the *Cyclodina* species do basically group together, a monophyletic origin for the genus *Cyclodina* with regards to *Leiolopisma* is not supported under the present taxonomic classification. More sequence data is required before the origin(s) and times of divergence of the New Zealand *Cyclodina* can be accurately assessed; however estimates from this study suggest that *Cyclodina* is older than previously thought. These conclusions agree with the finding of Hickson (1993) that 'the evolutionary history of the New Zealand skinks is more complex and fascinating than previously suspected'.

Chapter Two: Materials and Methods.

This chapter gives the species names, geographical origins and source of tissue for the skinks investigated in this study and describes the methods used to successfully obtain approximately 385 base pairs (bp) of DNA sequence for each of these individuals. These methods consisted of preparing genomic extractions containing high molecular weight (MW) DNA for each skink and then amplifying DNA from each of the extractions using the polymerase chain reaction (PCR) and 'universal' primers for the mitochondrial 12S ribosomal RNA (rRNA) gene (Kocher *et al.*, 1989). Templates for sequencing were prepared from the obtained PCR products by a centrifugal dialysis filtering method and sequenced using a direct double-stranded thermocycling-based sequencing procedure. As PCR is very sensitive to contamination, various steps were taken to ensure that the correct sequences were obtained - these are detailed in the procedures and discussed in the next chapter. Methods of alignment and phylogenetic analysis used in this study are also given.

Skink Material.

The tissue used in this study was frozen tail muscle, supplied from the reptile collection of the Museum of New Zealand. Tissue was provided for eight individuals - one from each of the six formally recognised and extant New Zealand *Cyclodina* species and one from each of the two populations which Vos (1988) proposed to be cryptic species (*C. aenea* from the Poor Knights Islands and *C. oliveri* from the Mokohinau Islands). The species names, geographical origins and Museum of New Zealand reptile collection catalogue numbers of these individuals are shown in Table 2.1 and Figure 2.1. The nomenclature used throughout this study to distinguish between the two *C. aenea* taxa and between the two *C. oliveri* taxa is also given in Table 2.1.

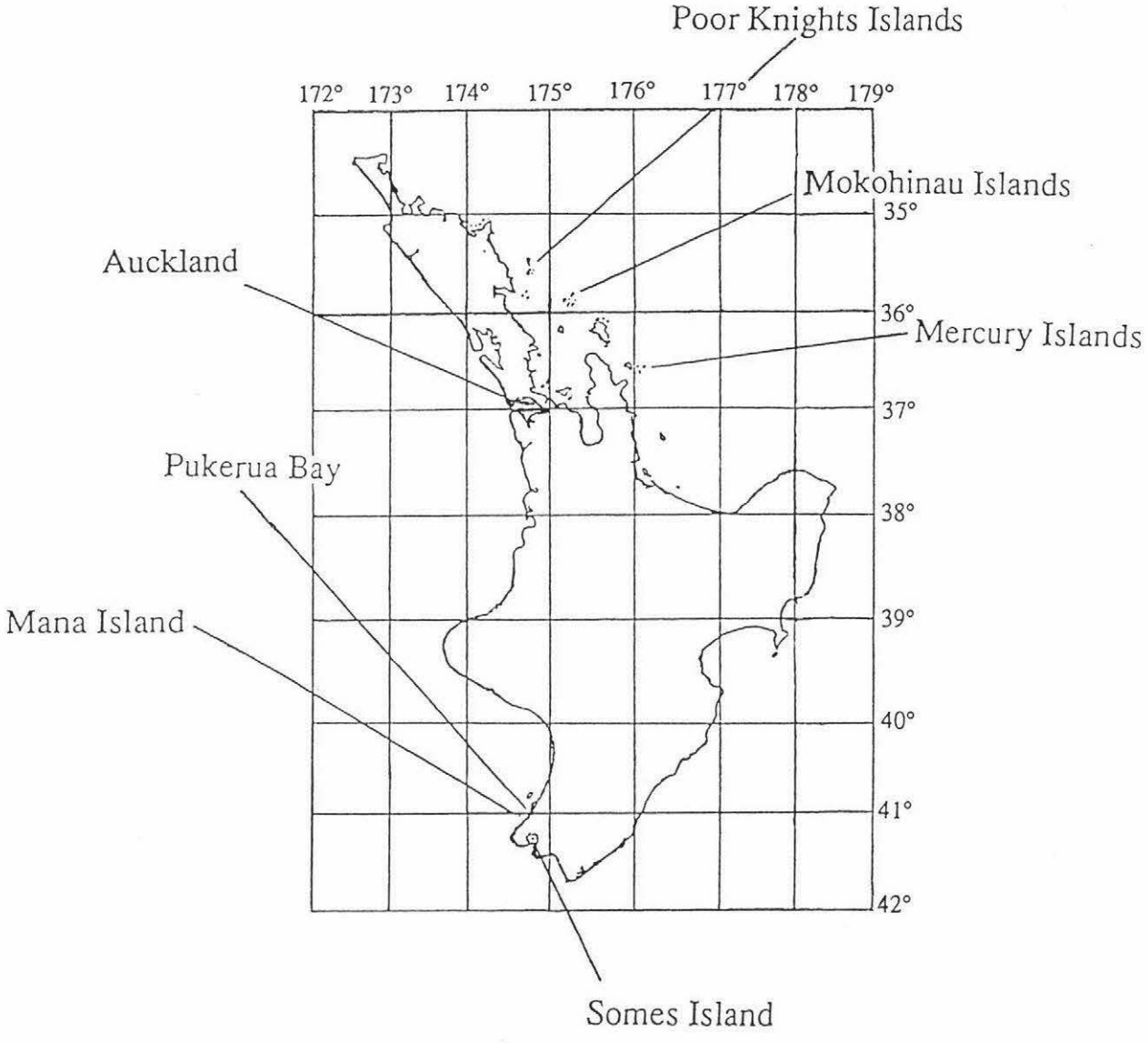
As mentioned in Chapter One, the individuals investigated in this study form a subset of those examined using allozyme data (Vos, 1988) - to allow direct comparison of the results of the two studies. While Vos (1988) investigated more than one individual (from geographically separated populations) for most of the *Cyclodina* species, only one population from each of the currently recognised or proposed (Vos, 1988) *Cyclodina* species needed to be examined here - as this study was designed as a preliminary investigation of the New Zealand *Cyclodina* using DNA sequence data, rather than as an exhaustive search for potential cryptic species or as a population study. And as the 12S rRNA gene is a relatively highly conserved mitochondrial gene, with much the same sequence expected from different individuals within the same population (Wilson *et al.*, 1985; Mindell and Honeycutt, 1990; Hickson *et al.*, 1992), only one individual from each of the selected populations was required.

Table 2.1. Species name, geographical origin and Museum of New Zealand reptile collection catalogue number for each of the eight *Cyclodina* taxa for which tissue was supplied and the nomenclature used throughout this study to distinguish between *C. oliveri* from the Mercury Islands and *C. oliveri* from the Mokohinau Islands and between *C. aenea* from Somes Island and *C. aenea* from the Poor Knights Islands.

Species	Origin	Cat. No.	Nomenclature
C. ornata	North Head, Devonport, Auckland	FT 188	
C. oliveri	Green I., Mercury Is., Coromandel	FT 137	*C. oliveri* (GM)
*C. oliveri**	Mokohinau Is., Northland	FT 182	*C. oliveri* (Mo)
C. whitakeri	Pukerua Bay, Wellington	CD 949	
C. alani	Green I., Mercury Is., Coromandel	FT 145	
C. macgregori	Mana I., Wellington	FT 3	
C. aenea	Somes I., Wellington Harbour	CD 1962	*C. aenea* (So)
*C. aenea**	Aorangi I., Poor Knights Is., Northland	CD 1037	*C. aenea* (PK)

* proposed to be a cryptic species (Vos, 1988).

Figure 2.1 Map of the North Island of New Zealand showing the geographical origins of the *Cyclodina* taxa investigated in this study.



Extraction of Total Genomic DNA.

Hickson (1993) evaluated the specificity of the 12S rRNA 'universal' primers (Kocher *et al.*, 1989) for the correct mitochondrial region in New Zealand skinks of the *Leiolopisma* genus. He found that the same unique PCR product and sequence was obtained from each individual whether genomic DNA or purified mitochondrial DNA (mtDNA) was used as the template. As it is simpler to prepare total DNA, purified mtDNA was accordingly not used in this study.

Two extractions were carried out for each individual, with a six month interval between the two, to allow the consistency of sequences and possibility of contamination to be evaluated. The basic procedure was that of Hickson *et al.* (1992) (see also Palumbi *et al.*, 1989; Hickson, 1993). A small amount (approximately 0.05 g) was aseptically removed from the supplied tail muscle (not including exposed [surface] tissue, which was discarded). This was placed in 300 µl of extraction buffer (50 µg/ml Proteinase K, 10 mM Tris-HCl pH 8.0, 100 mM EDTA, 1% SDS) and incubated for approximately fifteen minutes at 65°C. The tissue was then macerated briefly with a sterile macerator, and subsequently incubated overnight at 37°C (Sambrook *et al.*, 1989) or at 65°C for two to four hours.

After incubation, the solution was extracted using one volume (300 µl) of phenol plus one of chloroform. This step was repeated until the interface cleared (two extractions were usually sufficient). The solution was then extracted once using one volume of chloroform alone. The DNA was subsequently precipitated at 0°C for five minutes using 1/10 volume of 3M NaAc (pH 5.3) and two volumes of 95% ethanol, centrifuged for fifteen minutes at 0°C and washed with approximately 500 µl of cold 75% ethanol. This was followed by vacuum drying and resuspension at room temperature in 25 µl of sterile Milli-Q water. A negative control (containing no tissue) was set up simultaneously and using the same solutions to allow the possibility of contamination during extraction to be evaluated.

2 - 2.5 µl of the extraction product and of the extraction negative control were then checked on a 0.7% agarose (TBE) minigel along with 4 µl of 1 kb ladder (Gibco BRL, Life Technologies Inc., Gaithersburg, Maryland, USA; approximately 1 µg/µl) to ensure that high MW DNA had been successfully extracted, that the control was negative and to give a quantitative estimate of the amount of DNA present.

Several (three) variations in this experimental procedure were tried. The first involved a longer DNA precipitation time (five hours or overnight). The second involved taking more tissue and increasing the DNA precipitation time; while the third involved taking more tissue, lengthening

the incubation period prior to maceration to 25 - 30 minutes and increasing the DNA precipitation time. No systematic improvement was found using these variations. Using more tissue may have improved the yield somewhat, but this result was not consistent.

In addition, a direct comparison of the 37°C and 65°C incubation temperatures and times was carried out for two of the skinks. Twice the normal amounts of tissue and extraction buffer were used initially, then divided in half after the maceration step with one half incubated at 37°C overnight and the other at 65°C for two to four hours. No difference was noted between the resulting extraction products for either of the two skinks so tested.

Prior to PCR, the extraction product was diluted to ensure that the amount of DNA added to the PCR reaction would be within the range required for successful PCR and to attenuate any inhibitors which might be present. Dilutions were made empirically, according to the intensity of DNA on the gel. If little or no high MW DNA was visible - the extraction template was left undiluted or diluted 1 in 2, 1 in 10 or 1 in 25. Dilutions of 1 in 50 or 1 in 100 were made for extractions of mid intensity, while 1 in 250 or 1 in 500 dilutions were prepared from extractions with a large amount of DNA. Two to three dilutions were tried for most extraction products and whichever amplified most efficiently was used in subsequent PCR reactions. The extraction negative control was diluted 1 in 10, and 2 µl of both undiluted and diluted control were used in PCR reactions as a more sensitive test of whether contamination had occurred.

Amplification using the Polymerase Chain Reaction (PCR).

As mentioned in Chapter One, Hickson *et al.* (1992) found that the 12S rRNA 'universal' primers of Kocher *et al.* (1989) successfully amplified DNA sequence from the appropriate region of the mitochondria in *Leiolopisma* skinks. Accordingly they were the primers of choice for this study. The sequences of these primers are:
12SAR: 5'-AAACTGGGAT TAGATACCCC ACTAT-3' (L1091)
12SBR: 5'-GAGGGTGACG GGCGGTGTGT-3' (H1478).
L and H indicate the light and heavy strands of mtDNA while the numbers denote the 3' ends of the primers with regards to the complete human mtDNA sequence (Anderson *et al.*, 1981). The binding sites of these primers in the skink 12S rRNA gene are indicated in Tables 3.2 and 3.3.

The basic procedure followed was that of Hickson (1993). PCR reactions were carried out in 20 µl volumes and it was found that the following concentrations and amounts produced good amplifications: 1.25 - 2.5 µg/µl BSA, 100 µM of each deoxynucleotide triphosphate (dNTP), reaction buffer (Promega Corporation, Madison, Wisconsin, USA; 50 mM KCl, 10 mM Tris-HCl pH 9.0, 0.1% Triton X-100), 2.5 mM $MgCl_2$ (Promega), 0.05 - 0.1 µM of each primer, 2 - 4 µl of the appropriately diluted extraction product and 0.5 units of *Taq* DNA polymerase (Promega).

The range in values given for some of the PCR reagents indicates that, for some individuals, a better amplification was obtained by using a higher or lower concentration of these components. BSA was included because it increases the amount of product generated by the amplification reaction. Reactions were also set up using reaction buffer ($MgCl_2$-free), $MgCl_2$ and DNA polymerase from TFL™ (Epicentre Technologies) - no significant difference in amplification yield or, later, in clarity of sequence was noticed between amplification products from the two different brands.

Each reaction was set up in a 0.5 ml microcentrifuge tube and a drop of mineral oil was added to prevent evaporation. The PCR reaction was carried out using a DNA thermal cycler (Perkin Elmer Cetus, Connecticut, USA) and involved 35 cycles of:

94°C for 60 seconds for DNA strand separation
54°C for 60 seconds for primer annealing
72°C for 60 seconds for DNA copying

which Hickson (1993) identified as close to optimal for the 12S primers and skink DNA.

A negative and a positive control were set up with each PCR reaction. The same reagents were used in both experimental and control reactions, but the positive control contained a skink DNA source known to amplify successfully (initially *Leiolopisma*) to ensure that the reagents were working and that the conditions were correct, while the negative control contained no template to allow the possibility of contamination during PCR to be assessed.

After amplification, 2 µl from each reaction, including the PCR controls and the negative controls for the DNA extraction (diluted and undiluted), were checked on a 1.5% agarose (TBE) minigel along with 4 µl of 1 kb ladder (Gibco BRL; approximately 1 µg/µl) to ensure that a PCR product had been generated, to check the size of the product (approximately 400 bp), to give a quantitative estimate of the amount of product present and to ensure that DNA had not been amplified from either of the negative controls (DNA extraction or PCR).

Some initial trouble with contamination was experienced (evidenced by the occurrence of positive signals in the PCR negative control lanes). All tubes from these runs were discarded. Use of the following measures virtually eliminated the problem (suggesting cross-contamination/pipette carryover as the main source of contamination). Reactions were set up in 20 µl volumes, rather than the 80 µl volumes used by Hickson (1993). A set of new pipettes was obtained and used only for setting up the PCR reactions carried out in this study. Existing PCR stocks were discarded and new ones obtained. Aerosol resistant tips (such as those produced by Molecular BioProducts, Inc., San Diego, California, USA) were used for the

dilution of concentrated PCR reagents such as dNTPs (to ensure that cross-contamination of these stocks did not occur) and for the addition of DNA template to the PCR reaction tubes. As an extra precaution, all PCR reagents were aliquoted out into small volumes, again using aerosol resistant tips. Occasionally, a very faint band in the PCR negative control lane was observed even with these modifications - possibly due to airborne contamination during building renovations. This problem was solved by discarding the aliquots of PCR reagents currently in use (along with all tubes from these runs) each time this occurred.

Preparation of Templates for Sequencing.

The preparation of templates for sequencing involves the removal of substances which would otherwise inhibit sequencing (such as PCR oil and unused primer DNA and dNTPs) from the PCR reaction tubes. To allow the possibility of contamination to be assessed, several templates were prepared for each individual. These included templates from each of the two extractions and templates representing different amplifications of the same extraction. As 20 µl PCR reaction volumes were used as opposed to 80 µl ones, amplification products from several reactions were combined to produce enough template for sequencing. Only tubes having the same PCR template (that is, from the same extraction) were combined.

For each template, the contents of the appropriate PCR tubes were transferred to a sterile 1.5 ml Eppendorf tube (taking as little oil as possible) and combined. Following Hickson (1993), approximately 30 µl of chloroform was added for every 80 µl of PCR product, and the contents of the tube mixed and centrifuged briefly to remove any remaining mineral oil. Hickson (1993) assessed several methods of removing the rest of the inhibitory substances present and found that the use of centrifugal dialysis filters such as those produced by Promega (Magic PCR Preps™ DNA Purification System[1], Promega Corporation, Madison, Wisconsin, USA) was the most efficient and cost effective of those examined. Therefore, it was the method of choice here.

Following the procedure of Hickson (1993), 100 µl of Direct Purification Buffer (Promega) was added to the tube containing the oil free PCR product, followed by 1 ml of well suspended Magic PCR Preps™ DNA Purification Resin (Promega) and the solution was mixed by gently vortexing three times in the space of one minute. The DNA/resin mixture was then pipetted into a 5 ml syringe with a Magic minicolumn (Promega) attached to it and slowly pushed through the column. 2 ml of Magic PCR Preps™ Column Wash Solution (Promega; 80% isopropanol) was then added into the syringe and again the plunger was depressed slowly.

[1] Now Wizard™ PCR* Preps DNA Purification System

The minicolumn was then transferred from the syringe to a clean Eppendorf tube and centrifuged for 20 seconds to remove excess wash solution. The column was subsequently left at room temperature for five minutes to allow evaporation of any remaining isopropanol. Following this, 30 µl of sterile Milli-Q water was placed in the column, which was then left at room temperature for one minute (to allow elution of the DNA off the filter in the column into the water) and then centrifuged for 20 seconds. The resulting solution was collected and placed back into the column and the elution/centrifugation step repeated. This repetition increases the amount of product recovered (Hickson, 1993). 3 µl of the resulting product was run on a 1.5% agarose (TBE) minigel along with 4 µl of 1 kb ladder (Gibco BRL; approximately 1 µg/µl) to check the yield.

Direct Double-Stranded Sequencing using a Thermocycling-based Procedure.

Each template was sequenced using a direct double-stranded thermocycling-based sequencing approach. In his investigations into the New Zealand *Leiolopisma*, Hickson (1993) found that direct double-stranded sequencing (Casanova *et al.*, 1990) was more efficient than producing single-stranded PCR templates for sequencing (Gyllensten and Erlich, 1988; Kocher *et al.*, 1989). Accordingly, a direct double-stranded sequencing method was used in this study.

Thermocycling methods are a relatively recent development which have certain advantages over other sequencing methods - for example, less template is required, high denaturation temperatures and repeated cycles help prevent strand reannealing, and high polymerisation temperatures inhibit secondary structure formation in the template DNA - thereby facilitating polymerisation through highly structured regions (Anon., 1992). As the 12S rRNA molecule does form a secondary structure (see Chapter Four; Neefs *et al.*, 1993; Hickson *et al.*, 1996), this latter characteristic is of particular importance here. The method chosen in this study, the fmol™ DNA Sequencing System (Promega Corporation, Madison, Wisconsin, USA), also contains deaza nucleotide mixes which resolve compressions associated with G-C rich areas (Anon., 1992).

Four primers were used for sequencing - those used for PCR (12SAR and 12SBR) and two internal primers specific for skink DNA (SK12SL and SK12SR). The use of these four primers allowed the complete sequence of both strands of DNA between the 12SAR and 12SBR binding sites to be determined in skinks of the *Leiolopisma* genus (Hickson *et al.*, 1992; Hickson, 1993). The internal primers were designed from preliminary skink sequence data obtained using the PCR primers (Hickson *et al.*, 1992; Hickson, 1993). They are complementary and bind approximately halfway along the skink 12S PCR product (their binding position is shown in Table 3.3). Their sequences and location relative to the human mtDNA sequence are:
SK12SL: 5'-CTTCTTTCAT AAGGTAGGC-3' (L1408)
SK12SR: 5'-GCCTACCTTA TGAAAGAAG-3' (H1390).

The basic procedure followed was that given in the fmol™ DNA Sequencing System Technical Manual (Anon., 1992) for sequencing using direct incorporation. For each template, the following steps were performed (reagents - including water and mineral oil - should be kept on ice at all times and reactions should be set up on ice). Four 0.5 ml microcentrifuge tubes were labelled (G, A, T, C) and 2 µl of the appropriate d/dNTP Mix (Promega) was added to each tube. In a separate 0.5 ml tube, a 16µl 'cocktail' was set up containing between 2.5 and 7 µl of template DNA (depending on the intensity of the DNA in the preceding gel check step), 3 pmol of the appropriate primer, 5 µl of fmol™ Sequencing 5x Buffer (Promega; 250 mM Tris-HCl pH 9.0, 10 mM $MgCl_2$) and 0.5 µl of [α-^{35}S]dATP (1000 Ci/mmol, 10 µCi/µl). 1 µl of Sequencing Grade *Taq* DNA Polymerase (Promega; 5 u/µl) was then added to the 'cocktail' and the reagents mixed.

In early runs, the 'cocktail' reagents were mixed by pipetting the solution up and down several times. However, these early runs quite frequently produced sequence with one or two missing lanes. This problem was pinpointed to insufficient combination of the 'cocktail' contents and mixing by twice flicking and spinning briefly in a 0°C centrifuge was tried instead. Using this modification to the given procedure consistently resulted in four clear lanes.

4 µl of the 'cocktail' was then pipetted into each of the prepared d/dNTP tubes, one drop of mineral oil was added to each tube, the tubes were centrifuged briefly and then placed in a preheated DNA thermal cycler (Perkin Elmer Cetus) at 95°C for two minutes prior to thermocycling. Various thermocycling parameters were tried, with 45 cycles of the following being the most successful:

95°C for 30 seconds
50°C for 30 seconds
70°C for 30 seconds

On completion, 3 µl of fmol™ Sequencing Stop Solution (Promega; 10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanole) was added to each tube, and the tubes were briefly centrifuged.

A positive control was set up with each sequencing run to ensure that all reagents were working and that the conditions used were correct. This consisted of all of the components given above except that 500 fmol of pGEM-3Zf(+) (Promega) was added as the template and 3 pmol of pUC/M13 Forward Primer (Promega) as the primer.

Sequencing reactions were run on 6 % acrylamide, 7.5 M urea (TBE) sequencing gels, with a limiting parameter of 65 Watts. The gels were prerun for at least 20 minutes and samples denatured for two minutes at 75°C prior to loading. For sequencing reactions involving the 12SAR or 12SBR primers, one-third (3 µl) of the sequencing product was loaded for each of

three runs - usually one lasting approximately six hours (three dye fronts), one lasting approximately two hours (one dye front) and one lasting approximately one hour ('half' a dye front). For those involving the SK12SL or SK12SR primers, runs of approximately four hours (two dye fronts), two hours and one hour were normally carried out. These combinations of runs allowed all of the sequence generated by a given primer to be read clearly. The one hour runs actually provided little more information than the two hour runs - nevertheless, they were useful in clarifying some positions and in eliminating reading errors (see Chapter Three).

When using the Fmol sequencing procedure, gels with shorter length runs on them must be fixed before being exposed to film to get rid of the radioactive by-products of sequencing that will otherwise obscure the sequence (Anon., 1992). Fixing was carried out in 10% acetic acid, 10% ethanol for half an hour to two hours. Each gel was then dried and exposed twice to Kodak X-OMAT™ AR film (Eastman Kodak Company, Rochester, New York, USA) - with a longer exposure period for the second film. The production of two films for each gel facilitated the reading of sequence at the top (shorter exposure period) and the bottom (longer exposure period) of the gel and contributed immensely to the elimination of reading errors (see Chapter Three).

Alignment and Phylogenetic Analysis.

The obtained sequences were aligned (see Chapter Three) using the programme Esee Version 3.0 (E. Cabot; available from cabot@trog.mbb.sfu.ca). Phylogenetic analysis involved the two main programmes of the Hadamard conjugation (Hendy and Penny, 1989, 1993; Penny *et al.*, 1992, 1993a; Lento *et al.*, 1995; available from FARSIDE@massey.ac.nz) - the multifunction programme Prepare and the treebuilding programme Hadtree (see Chapter Four). The maximum parsimony, minimum evolution and maximum likelihood options of a new test version of PAUP (phylogenetic analysis using parsimony; see Swofford, 1985, 1993) - PAUP* Version 4.0 (1995; available from orders@sinauer.com) test versions 4.0.0d34 and 4.0.0d34 - d40 (1996) and, to a lesser extent, the parsimony and neighbor-joining options of PHYLIP (phylogeny inference package; see Felsenstein, 1993; available from joe@genetics.washington.edu) version 3.51c were also used for phylogenetic analysis (see Chapter Four).

Chapter Three: Results.

A unique DNA sequence was obtained for each of the eight skinks under study using the procedures described in Chapter Two. This chapter details the results of each of the methods, including the results of steps taken to ensure that the sequences did indeed represent 12S ribosomal RNA (rRNA) sequences from the individuals investigated, and presents the obtained sequences, both in full and in a consensus format.

Total Genomic DNA Extraction.

High molecular weight (MW) DNA was successfully extracted twice for each individual. None of the DNA extraction controls showed any traces of DNA on the initial gel check, nor was DNA amplified from either diluted or undiluted DNA extraction controls during the polymerase chain reaction (PCR). This indicates that the total DNA preparations were not contaminated during the extraction procedure.

Amplification using the Polymerase Chain Reaction (PCR).

PCR products of the correct size were obtained from all PCR runs except for the following: amplification did not occur with the undiluted extraction product from one of the *C. oliveri* (Mo) extractions, nor with a 1 in 2 dilution of one of the extractions from *C. whitakeri*. In both cases, PCR was successful when the extraction product was diluted further, probably because inhibitory substances present in the concentrated extraction products had been successfully diluted out (R. E. Hickson, pers. comm.). As mentioned in Chapter Two, a signal was occasionally observed in the PCR negatives - all tubes from such runs were discarded in case of contamination.

Occasionally more than one band was observed when the PCR products were checked on minigels. When compared to the size ladder, these extra bands were approximately 800, 1200, 1600 and 2000 base pairs (bp) in size (the latter two were only seen infrequently). Considering the sizes of these bands and the fact that only one sequence was obtained for each skink, the most likely explanation is that they represent multimers of the correct product (approximately 400 bp), probably resulting from the addition of too much DNA template to the PCR reaction (Palumbi *et al.*, 1989; A. C. Cooper, pers. comm.).

Preparation of Sequencing Templates.

Several highly concentrated DNA templates were successfully prepared for each skink from the obtained PCR products, including templates from the different extractions. However, sequence was not obtained from all of these templates due to initial problems with the sequencing procedure (in particular - missing sequence lanes, as mentioned in Chapter Two).

Thermocycling-based Direct Double-stranded Sequencing and Alignment and Comparison of the Resulting Partial Sequences.

An average of 6.63 partial sequences were obtained for each skink - with a partial sequence being defined as the set of sequencing gel runs (three runs in 88.68% of cases) generated by each specific primer/template combination (the sequences are partial because individual primers do not produce the entire 385 bp of sequence). A summary of these partial sequences is presented in Table 3.1 and details, for each skink, the total number of templates used, the primers used with each template, the extraction each template was derived from (Table 3.1a) and the number of partial sequences obtained using each primer and in total (Table 3.1b).

Points to note in Table 3.1a include: Sequence for all but one of the individuals (*C. alani*) was obtained from two or more different templates, and from all but two (*C. alani* and *C. oliveri* [GM]) using templates derived from each of the two extractions (although all four primers were not necessarily used for all templates). With the exception of *C. aenea* (PK) for which only three primers were successful (see later), each of the four primers was used to generate sequence at least once for each skink. Sequence was obtained using all four primers with one template from each extraction for one skink (*C. macgregori*).

The main point to note in Table 3.1b is that all but six of the 53 partial sequences obtained in total consisted of a set of three sequencing gel runs each (those that did not are indicated) - with a total of 151 sequencing gel runs carried out in all. Most of these sets of three consisted of the runs outlined in Chapter Two (a six hour, a two hour and a one hour run for the 12SAR and 12SBR primers and a four hour, a two hour and a one hour run for the SK12SL and 12SKR primers). In some cases, however, a different combination of runs was used (a six hour and two x two hour combination for the 12SAR and 12SBR primers and a six hour, a two hour and a one hour combination or a six hour, a four hour and a two hour combination for the SK12SL and 12SKR primers). As these alternative combinations still consisted of three runs and still allowed the complete sequence produced by the primer to be read clearly, they were regarded as equivalent to the 'normal' combinations (and thus are not differentiated in Table 3.1).

Table 3.1. Summary of the partial sequences obtained in this study - indicating, for each taxa, the total number of templates used, the primers used with each template, the extraction each template was derived from (**a.**) and the number of partial sequences obtained using each primer and in total (**b.**). Partial sequences which do not consist of three gel runs are indicated (**b.**). In total, 53 partial sequences, comprising 151 sequencing runs, were produced. Each of the runs was read as completely as possible twice - once from each of the duplicate films produced for each gel.

a. Summary of the total number of sequencing templates used for each taxon, which primers were used with each template (external - 12SAR and 12SBR; internal - SK12SL and SK12SR) and which extraction each template was derived from (first or second).

Taxon	Total number of templates used	Primers used with each template (template number: primers used)	Extraction
C. ornata	Three	#1: 12SAR+12SBR+SK12SL+SK12SR	Second
		#2: SK12SL+SK12SR	Second
		#3: SK12SL+SK12SR	First
C. oliveri (GM)	Three	#1: 12SAR+12SBR+SK12SL+SK12SR	Second
		#2: 12SAR+12SBR+SK12SL+SK12SR	Second
		#3: SK12SL+SK12SR	Second
C. oliveri (Mo)	Two	#1: 12SAR+12SBR+SK12SL+SK12SR	Second
		#2: SK12SL+SK12SR	First
C. whitakeri	Three	#1: 12SAR+12SBR+SK12SL+SK12SR	First
		#2: 12SAR+12SBR	First
		#3: SK12SL+SK12SR	Second
C. alani	One	#1: 12SAR+12SBR+SK12SL+SK12SR	Second
C. macgregori	Two	#1: 12SAR+12SBR+SK12SL+SK12SR	First
		#2: 12SAR+12SBR+SK12SL+SK12SR	Second
C. aenea (So)	Two	#1: 12SAR+12SBR+SK12SL+SK12SR	Second
		#2: 12SAR+12SBR	First
C. aenea (PK)	Two	#1: 12SAR+12SBR	Second
		#2: SK12SL[§]	First

b. Summary of the number of partial sequences (sets of sequencing gel runs generated by each specific primer/template combination) obtained for each taxon, showing the number of partial sequences generated using each primer and the total number of partial sequences obtained for each skink. Except where indicated otherwise (*, [+] and [x]), each partial sequence comprises three sequencing gel runs.

Taxon	Number of partial sequences generated using:				Total number of partial sequences.
	12SAR	12SBR	SK12SL	SK12SR	
C. ornata	1	1	3*	3[+]	8
C. oliveri (GM)	2	2	3[x]	3	10
C. oliveri (Mo)	1	1	2	2	6
C. whitakeri	2	2	2	2	8
C. alani	1	1	1	1	4
C. macgregori	2	2	2	2	8
C. aenea (So)	2	2	1	1	6
C. aenea (PK)	1	1	1	-[§]	3
All eight					53

* One of these consists of one sequencing gel run only; the other two consist of two runs each.

[+] One of these consists of two sequencing gel runs only; the other two consist of three runs each, but one run in each set is incomplete due to problems with the sequencing gel or film.

[x] One of these consists of one sequencing gel run only; one consists of two runs only.

[§] Although several attempts were made to obtain sequence for *C. aenea* (PK) using the 12SKR primer, the obtained sequence was always very faint and, when read, did not appear to be for the correct region (see text).

Each of the sequencing gel runs was read as completely as possible twice - once from each of the duplicate films produced for each gel. These duplicate sequences and the sequence overlaps between the different runs comprising each partial sequence were compared and reading errors eliminated. A partial sequence alignment was then created for each skink (data not shown) using the programme Esee Version 3.0 (E. Cabot; available from cabot@trog.mbb.sfu.ca) and the sequences in each alignment were compared. Overall, the partial sequences obtained for each individual using the different primers and templates (including templates from both extractions) were both clear and consistent - confirming one unique sequence for each skink. A small number of anomalies were noticed on each gel (see later) - however, all but a few of these were eliminated through comparison of the different primer and template sequences.

The eight complete, aligned sequences are presented in Tables 3.2 and 3.3. Table 3.2 shows the sequences in full, with the remaining few positions for which confidence is not absolute given in small case (nevertheless, confidence in these positions was still high enough to define them as small case as opposed to N's). Table 3.3 presents the same sequences in a consensus format. Given the extensive checking of all sequences through the comparison of duplicate films, of overlaps in sequence between the different runs (particularly between the one and two hour runs) and of the different partial sequences in each alignment, the eight sequences presented in Tables 3.2 and 3.3 are extremely unlikely to contain either reading or sequencing errors.

Placement of Gaps (Insertions/Deletions).

The *Cyclodina* dataset contains six gaps - *C. aenea* (So) and *C. ornata* each have a single base deletion between positions 25 and 30, *C. aenea* (PK) has two single base deletions in the same region while *C. alani* has two single base insertions - one at position 60 or 61, the other between positions 367 and 370. These gaps cannot be placed precisely as they either fall into a variable region or in a region with several identical bases or both (see Tables 3.2 and 3.3). Examination of the secondary structure of these sequences (see Chapter Four) indicates that all six gaps fall into unpaired (loop) regions, but does not help in exact placement. Therefore, two of the gaps have been placed in the most probable position (position 25 for *C. aenea* (PK)'s first deletion and position 29 for *C. ornata*'s deletion), while the remaining four, for which a logical placement cannot be made, have been placed in a default position (position 26 for *C. aenea* (PK)'s second deletion and *C. aenea* (So)'s deletion and positions 61 and 369 for *C. alani*'s two insertions). As these placements are not necessarily correct, variable regions containing gaps (columns 25-30 and 60-61) have been omitted from many of the phylogenetic analyses of *Cyclodina* carried out in this study (see Chapter Four).

Table 3.2. *Cyclodina* 12S rRNA sequences - 387 bases (including gaps), aligned and shown in full. Bases for which confidence is not 100% are shown in bold small case. Gaps are indicated by dashes. *C. aenea* (So) and *C. ornata* each have a single base deletion between positions 25 and 30, *C. aenea* (PK) has two single base deletions in the same region while *C. alani* has two single base insertions - one at position 60 or 61, the other between positions 367 and 370 (see text for explanation of the placements shown in this table). Columns containing gaps were omitted from many of the phylogenetic analyses presented in this study. Due to inability to place these gaps unambiguously, the remaining columns in regions 25-30 and 60-61 were also sometimes omitted. The binding positions of the 12SAR and 12SBR primers are not shown but lie immediately to the left and right, respectively, of the sequences. The SK12SL and SK12SR primers bind between positions 186 and 204 (see Table 3.3).

```
                        1         2         3         4         5         6
                123456789012345678901234567890123456789012345678901234567890
C.aenea(So)     GCCCAGCCGTCAAcAAAGACAGTAT-AAACACAACACTGTTCGCCAGAGAACTACAAGCT
C.aenea(PK)     GCTCAGCCGTTAACAAAGACAGTA--AAACACAATACTGTTCGCCAGAGAACTACAAGCA
C.whitakeri     GCTCAGCCGTCaAcAAAGACAGTATAAGATACAATACTGTTCGCCGGAGAACTACAAGCT
C.alani         GCTCAGCCGTCAACAAAGACAGTATAAAACACAACACTGTTCGCCAGAGAACTACAAGCT
C.macgregori    GCTCAGCCGTCAAcAAAGACAGTACAAAGTACAATACTGTTCGCCAGAGAACTACAAGCT
C.oliveri(Mo)   GCTCGGCCGTCAAcAAAGACAGTATAAGATACAACACTGTTCGCCAGAGAACTACAAGCT
C.oliveri(GM)   GCCCGGCCGTCAAcAAAgACAGTATAAGAcACAATACTGTTCGCCAGAGAACTACAAGCT
C.ornata        GCTCAGcCGTCAAcAAAGACAGTATAAG-TACAATACTGTTCGCCAGAGAACTACAAGCT

                        7         8         9        10        11        12
                123456789012345678901234567890123456789012345678901234567890
C.aenea(So)     -AAAGCTAAAAACCCCAAGGACTTGGCGGTGCTCCACATCACCCTAGAGGAGCCTGTCCT
C.aenea(PK)     -AAAGCTCGAAACTCCAAGGACTTGGCGGTGCTCCACATCACCCTAGAGGAGCCTGTCCT
C.whitakeri     -AAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCCCCCTAGAGGAGCCTGTCCT
C.alani         TAAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCCCCCTAGAGGAGCCTGTCCT
C.macgregori    -AAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCATCCTAGAGGAGCCTGTCCT
C.oliveri(Mo)   -AAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCACCCTAGAGGAGCCTGTCCT
C.oliveri(GM)   -AAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCACCCTAGAGGAGCCTGTCCT
C.ornata        -AAAGCTAAAAACTCCAAGGACTTGGCGGTGCTCCACATCACCCTAGAGGAGCCTGTCCT

                       13        14        15        16        17        18
                123456789012345678901234567890123456789012345678901234567890
C.aenea(So)     ATAATCGATACCCCCCGATCCACCTCACCACTTTTTGAAATTCAGCCTATATACCGCCGT
C.aenea(PK)     ATAATCGATACCCCCCGATCTACCTCACCACTTTTTGAAACTCAGCCTATATACCGCCGT
C.whitakeri     ATAATCGATACCCCACGATCTACCTCACCATTTTTTGAAACTCAGCCTATATACCGCCGT
C.alani         ATAATCGATACCCCCCGATCCACCTCACCATTTTTTGAAACTCAGCCTATATACCGCCGT
C.macgregori    ATAATCGATACCCCCCGATCCACCTCACCATTTTTTGAACCTCAGCCTATATACCGCCGT
C.oliveri(Mo)   ATAATCGATACCCCACGATCTACCTCACCATTTTTTGAaAcTCAGCCTATATACcGCCGT
C.oliveri(GM)   ATAATCGATACCCCACGATCTACCTCACCATTTTTTGAAACTCAGCCTATATACCGCCGT
C.ornata        ATAATCGATACCCCCCGATCTACCTCACCATTTTTTGAaACTCAGCCTATATACCGCCGT
```

..... continued next page

```
                        19         20         21         22         23         24
                        123456789012345678901234567890123456789012345678901234567890
C.aenea(So)             CGTCAGCCTACCTTGTGAAAGAAGCGTAGTAAGCAAAATAGTCACCAACTAAAACGTCAG
C.aenea(PK)             CGTCAGCCTACCTTATGAGAGAGACACAGTAAGCAAAATAGTCACCAACTAAAACGTCAG
C.whitakeri             CGTCAGCCTACCTTATGAaAGAAGCACAGTAAGCGAAATAGTCACCAACTAAAACGTCAG
C.alani                 CGACAGCCTACCTTATGAAAGAAGCACAGTAAGCAAAATAGTCACCAACTAAAACGTCAG
C.macgregori            CGTCAGCCTACCTTATGAAAGAaGCAAAGTAAGCgAAATAGTCACCAACTAGAACGTCAG
C.oliveri(Mo)           CGTCAGCCTACCTTATGAaAGAAGCACAGTAAGcaAAATAGTCACCAACTAAaACGTCAG
C.oliveri(GM)           CGTCAGCCTACCTTATGAAAGAAGCACAGTAAGCGAAATAGTCACCAACTAAAACGTCAG
C.ornata                CGTCAGCCTACCTTGTGAAAGAAGCACAGTAAGcAAAACAGTCACCAACTAAAACGTCAG

                        25         26         27         28         29         30
                        123456789012345678901234567890123456789012345678901234567890
C.aenea(So)             GTCAAGGTGTAGCACATAAAGCGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACG
C.aenea(PK)             GTCAAGGTGTAGCACATAAAGTGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACG
C.whitakeri             GTCAAGGTGTAGCACATAAAATGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACG
C.alani                 GTCAAGGTGTAGCACATAAAATGGAAGAGATGGGCTACACTCTCTCCCACAGAGAACACG
C.macgregori            GTCAAGGTGTAGCACATAAAATGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACA
C.oliveri(Mo)           GTCAAGGTGTAGCACATAAAATGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACG
C.oliveri(GM)           GTCAAGGTGTAGCACATAAAATGGAAGAGATGGGCTACACTCTCTCCCCCAGAGAACACG
C.ornata                GTCAAGGTGTAGCACATGAAATGGAAGAgATGGGCTACACTCTCTCCCCCAGAGAACACG

                        31         32         33         34         35         36
                        123456789012345678901234567890123456789012345678901234567890
C.aenea(So)             AACAGCACCAATGAAACACTGCTCAAAGGCGGATTTAGTAGTAAGATAAACAAGAGAACT
C.aenea(PK)             AATAGCATCAATGAAACACGGCTCGAAGGTGGATTTAGTAGTAAGATAAACAAGAGAACT
C.whitakeri             AACAGCATCAATGAAACACTGCTCAAAGGCGGATTTAGTAGTAAGATAAACAAGAGAGCT
C.alani                 AAGAGCACCAATGAAACACTGCTCAAAGGTGGATTTAGTAGTAAGATAAACAAGAGAACT
C.macgregori            AACAGCACCAATGAAATACTGCTCAAAGGCGGATTTAGTAGTAAGATAAACAAGAGAACT
C.oliveri(Mo)           AACAGCATCAATGAAACCCTGCTCAAAGGTGGATTTAGTAGTAAGATAAACAAGAGAACT
C.oliveri(GM)           AACAGCATCAATGAAACACTGCTCAAAGGTGGATTTAGTAGTAAGATAAACAAGAGAACT
C.ornata                AACAGCACCAATGAAACACTGCTCAAAGGCGGATTTAGTAGTAAGATAAACAAGAGAACT

                        37         38
                        12345678901234567890123456
C.aenea(So)             TATCTTAA-ACCAGCCCTGGAGCGCGC
C.aenea(PK)             TATCTTAA-ACCAGCCCTGGAGCGCGC
C.whitakeri             TATCTTAA-ACCAGCCCTGGAGCGcGC
C.alani                 TATCTTAAAACCAGCCCTGGAGCGCGC
C.macgregori            TATCTTAA-ACCAGCCCTGGAGCGCGC
C.oliveri(Mo)           TATCTTAA-ACCAGCCCTGGAGCGCGC
C.oliveri(GM)           TATCTTAA-ACCAGCCCTGGAGCGCGC
C.ornata                TATCTTAA-ACCAGCCCTGGAGCGCGC
```

Table 3.3. *Cyclodina* 12S rRNA sequences - 387 bases (including gaps), aligned and shown with the consensus sequence at top. Conserved positions are in upper case and variable ones in lower case. A dot indicates that the sequence for that taxa is the same as the consensus sequence at that position. Positions at which there are transversions are in bold (both in the consensus sequence and in the aligned sequences). Gaps are indicated by dashes. *C. aenea* (So) and *C. ornata* each have a single base deletion between positions 25 and 30, *C. aenea* (PK) has two single base deletions in the same region while *C. alani* has two single base insertions - one at position 60 or 61, the other between positions 367 and 370 (see text for explanation of the placements shown in this table). Columns containing gaps were omitted from many of the phylogenetic analyses presented in this study. Due to inability to place these gaps unambiguously, the remaining columns in regions 25-30 and 60-61 were also sometimes omitted. The binding positions of the 12SAR and 12SBR primers are not shown but lie immediately to the left and right, respectively, of the sequences. The binding site of the SK12SL and SK12SR primers is underlined (positions 186 and 204).

```
                        1         2         3         4         5         6
                123456789012345678901234567890123456789012345678901234567890
                GCuCaGCCGUcAACAAAGACAGUAuaAaauACAAuACUGUUCGCCaGAGAACUACAAGCu
C.aenea(So)     ..c.....................-...c....c.....................
C.aenea(PK)     .........u............--...c.........................a
C.whitakeri     ........................g.............g.............
C.alani         ........................c....c.....................
C.macgregori    ......................c...g.............................
C.oliveri(Mo)   ....g...................g......c.....................
C.oliveri(GM)   ..c.g...................g.c...........................
C.ornata        ........................g-............................

                        7         8         9        10        11        12
                123456789012345678901234567890123456789012345678901234567890
                -AAAGCUaaAAACuCCAAGGACUUGGCGGUGCUCCACAUCacCCUAGAGGAGCCUGUCCU
C.aenea(So)     -...........c.....................................................
C.aenea(PK)     -......cg........................................................
C.whitakeri     -.....................................c.....................
C.alani         u.....................................c.....................
C.macgregori    -.....................................u.....................
C.oliveri(Mo)   -................................................................
C.oliveri(GM)   -................................................................
C.ornata        -................................................................

                       13        14        15        16        17        18
                123456789012345678901234567890123456789012345678901234567890
                AUAAUCGAUACCCCcCGAUCuACCUCACCAuUUUUUGAAacUCAGCCUAUAUACCGCCGU
C.aenea(So)     ..................c.........c.........u.....................
C.aenea(PK)     ............................c.............................
C.whitakeri     ............a.....................................................
C.alani         ..................c.....................................
C.macgregori    ..................c.................c.....................
C.oliveri(Mo)   ............a.....................................................
C.oliveri(GM)   ............a.....................................................
C.ornata        .....................................................
```

..... continued next page

```
                         19        20        21        22        23        24
              123456789012345678901234567890123456789012345678901234567890
              CGuCAGCCUACCUUaUGAaAGAagCacAGUAAGCaAAAuAGUCACCAACUAaAACGUCAG
C.aenea(So)   ..............g..........gu.................................
C.aenea(PK)   ..................g...ga....................................
C.whitakeri   ..................................g.........................
C.alani       ..a.........................................................
C.macgregori  ..........................a.......g................g........
C.oliveri(Mo) ............................................................
C.oliveri(GM) ..................................g.........................
C.ornata      ..............g.......................c.....................

                         25        26        27        28        29        30
              123456789012345678901234567890123456789012345678901234567890
              GUCAAGGUGUAGCACAUaAAauGGAAGAGAUGGGCUACACUCUCUCCCcCAGAGAACACg
C.aenea(So)   ....................gc......................................
C.aenea(PK)   ....................g.......................................
C.whitakeri   ............................................................
C.alani       ................................................a...........
C.macgregori  ...........................................................a
C.oliveri(Mo) ............................................................
C.oliveri(GM) ............................................................
C.ornata      .................g..........................................

                         31        32        33        34        35        36
              123456789012345678901234567890123456789012345678901234567890
              AAcAGCAuCAAUGAAAcaCuGCUCaAAGGuGGAUUUAGUAGUAAGAUAAACAAGAGAaCU
C.aenea(So)   .......c.....................c..............................
C.aenea(PK)   ..u................g....g...................................
C.whitakeri   .............................c...........................g..
C.alani       ..g....c....................................................
C.macgregori  .......c........u............c..............................
C.oliveri(Mo) .................c..........................................
C.oliveri(GM) ............................................................
C.ornata      .......c.....................c..............................

                         37        38
              123456789012345678901234567
              UAUCUUAA-ACCAGCCCUGGAGCGCGC
C.aenea(So)   ........-..................
C.aenea(PK)   ........-..................
C.whitakeri   ........-..................
C.alani       ........a..................
C.macgregori  ........-..................
C.oliveri(Mo) ........-..................
C.oliveri(GM) ........-..................
C.ornata      ........-..................
```

Some Features and Anomalies noticed in Sequencing.

Hickson (1993) found that the 12SBR primer gave consistently fainter sequencing lanes. In this study, sequence obtained using the 12SBR or SK12SL primers was found to be slightly fainter than that obtained using the other two primers, but was always clear and readable.

Although several attempts were made to obtain sequence for *C. aenea* (PK) using the SK12SR primer, the resulting sequences were extremely faint and, when read, were not for the correct region. It would seem that for *C. aenea* (PK), the SK12SR primer binds at an alternate priming site - exactly where was not determined, but it appeared to be somewhere near the 12SAR binding site. When the SK12SR binding site is investigated in the *C. aenea* (PK) sequence, a probable explanation is found: several changes have occurred in the crucial 3' binding region - the 3' and penultimate 3' base are different from the primer sequence and another change has occurred six bases from the 3' end (see Table 3.3). Kocher *et al.* (1989) noted that primers that have several mismatches to the template can be successfully used, as long as there is exact matching between primer and template and the last few bases of the 3' end of the oligonucleotide. The matter was not further investigated in this study.

Anomalies noticed included 'shadow-banding', more than one distinct band at one position on the gel and heavy bands in all four lanes at the same position on the gel. Faint 'shadow-banding' (faint bands in all lanes and at all positions on the gel) was observed on all gels. Innis and Gelfand (1990) noted that 'shadow-banding' represents misincorporation errors (which promote chain termination and affect fidelity) and is generally due to low or unbalanced deoxynucleotide triphosphate (dNTP) levels. However, it may also be due to the addition of excess template (A. C. Cooper, pers. comm.), or possibly be the result of the thermocycling method itself. Whatever the cause, the sequence could still be clearly read as it consisted of very intense bands while the 'shadow bands' were of uniform very light intensity.

At some sites in the sequences, more than one distinct band was observed. This phenomenon can be due to the presence of another, contaminating, template (Anon., 1992), to secondary priming (Anon., 1992) or to amplification from duplicate mitochondrial sequences - which have been identified in the mitochondrial genomes of lizards (Moritz and Brown, 1986) and in the nuclear genomes of cats (Lopez *et al.*, 1994).

In general, these extra bands, while stronger than 'shadow bands', were still comparatively light in intensity and in all but a few cases, the use of a different primer or template eliminated the problem. In many cases, sequence obtained from the same template using a different primer

(particularly one binding to the complementary strand) did not display extra bands at the same positions, while the use of the same primer with a different template generally resulted in the same bands appearing. This would suggest that neither contamination nor amplification from a duplicate mitochondrial sequence were the cause of these ambiguities. In most of the remaining few cases, the use of a different template clarified the sequence at that position.

The appearance of heavy bands in all four lanes at a single site was observed in some sequences. However, in all cases, when sequence was obtained using a primer binding to the complementary strand, the problem was no longer seen. This feature may be due to disassociation of the enzyme due to secondary structure in the template strand (Anon., 1992).

Confirming the Identity of the Sequences.

The overall consistency of the partial sequences for each skink - sequences from a number of templates and primers (including templates from different extractions) - and the fact that eight unique but similar sequences were obtained suggests that these sequences are unlikely to be the result of contamination. In addition, the sequences are very similar to, but distinct from, those obtained by Hickson (1993) for New Zealand skinks of the *Leiolopisma* genus (see Chapter Four). Indeed, phylogenetic analysis suggests that some of the *Cyclodina* species are as close to *Leiolopisma* species as to other members of *Cyclodina* (see Chapter Four).

Hickson (1993) included a variety of other vertebrate 12S sequences (frog, toad, fish, xantusiid, skink, cow, rat, mouse, human, chicken and moa) in several of his analyses - the pattern of relationships inferred by phylogenetic analysis placed *Leiolopisma* closest to the xantusiid lizard, confirming that the obtained sequences were indeed reptilian (Hickson, 1993). As the sequences obtained in this study were so close to those obtained by Hickson (1993), it would seem that they do indeed represent 12S rRNA skink sequences, while the fact that they are distinct is consistent with them being from *Cyclodina*.

Chapter Four: Analysis.

In this chapter, the *Cyclodina* sequences are aligned with the sequences of Hickson (1993) and these sequences are examined in terms of G-C content. The *Cyclodina* sequences are then fitted to the vertebrate secondary structure models refined by Hickson (1993) and Hickson *et al.* (1996). This provides a further estimate of sequencing fidelity and allows the occurrence of gaps (indels) and the number and distribution of constant and variable sites in the *Cyclodina* and New Zealand *Leiolopisma* datasets to be compared in the context of secondary structure. The number of total differences, transitions and transversions in pairwise comparisons are also examined for these datasets and some approximate estimates of times of divergence are made. Phylogenetic analysis of the *Cyclodina* sequences alone and in conjunction with the sequences of Hickson (1993) is then carried out using the Hadamard conjugation, PAUP* and PHYLIP, with the Hadamard spectra for the *Cyclodina* analyses represented in spectral diagrams or Lento-plots (Lento *et al.*, 1995). The resulting phylogenies are compared to those of Vos (1988) from allozyme data. The phylogenies are also examined with respect to the geographical distribution of the New Zealand taxa, the placement of three possible outgroups is investigated and the possible origin(s) of *Cyclodina* are briefly discussed.

Alignment of Available Skink 12S rRNA Sequences.

To achieve the primary aims of this thesis - investigating the taxonomic status, relationships and origin(s) of the New Zealand *Cyclodina* using DNA sequence data - the *Cyclodina* sequences must be analysed not only separately, but also in conjunction with other available skink sequence data. Currently, this is encompassed by the dataset of Hickson (1993), which comprises 17 New Zealand *Leiolopisma* sequences, one *Cyclodina* sequence (*C. aenea*) and three non- New Zealand skink sequences - *L. telfairi* from Mauritius (the *Leiolopisma* type species), *Lampropholis guichenoti* (an Australian skink) and a partial sequence (305 bp) for *Tropidoscincus rohssii* (a New Caledonian skink). The taxon names, geographical origins and Museum of New Zealand reptile collection catalogue numbers for Hickson (1993)'s sequences are presented in Table 4.1, as are abbreviations of these names used in this study.

Before analysis can be carried out, these sequences must be aligned. Comparison of the *Cyclodina* dataset to several of the sequences of Hickson (1993) indicated that the two datasets could be easily aligned by manual means. The resulting skink 12S rRNA alignment is presented in Table 4.2. The sequences of Hickson (1993) as given in this alignment do contain one minor change to the sequences as presented by Hickson (1993): Comparison of the *Cyclodina* dataset to Hickson (1993)'s sequences revealed that the *Cyclodina* sequences (except for the *C. aenea* sequence of Hickson, 1993 - designated *C. aenea* [R] in Table 4.2) had an 'extra' base (position 358 in Tables 3.2 and 3.3; between positions 356 and 357 in Table 3.1 and Appendix One of

Hickson, 1993). The seven *Leiolopisma* sequences in Hickson *et al.* (1992) also contain this 'extra' base (position 357), as do all sequences in the vertebrate alignment of Hickson (1993 - Table 3.3) except for the skink sequence (this was the only position where the skink sequence had a unique gap). This suggested that a column might have been accidentally deleted from the dataset presented in Hickson (1993), rather than the 'extra' base representing a genuine *Cyclodina*-specific insertion. Many of the original sequencing films of Hickson (1993) were available and examination of these confirmed that a column had been accidentally omitted. The original gels, in conjunction with the sequences of Hickson *et al.* (1992) provided the missing sequence data (position 359 in Table 4.2). However, there were still some ambiguities at this position and therefore the 'extra' column has been omitted from all phylogenetic analyses except for those involving *Cyclodina* alone.

Hickson (1993)'s *Cyclodina* sequence (*C. aenea* [R]) also differs from that obtained for the same individual (CD 1962) in this study (*C. aenea* [So]) at two other positions - the *C. aenea* (So) sequence has a deletion between positions 25 and 30 and a T instead of a C at position 30 (see Table 4.2). As this individual was sequenced several times in this study (see Chapter Three, Table 3.1 in particular), the *C. aenea* (So) sequence is assumed to be correct and the *C. aenea* sequence of Hickson (1993) has been omitted from further analysis. The sequences in Table 4.2 contain one other minor change when compared to the sequences as presented in Table 3.1 of Hickson (1993): the sequence between positions 146 and 149 (positions 145-148 in Hickson, 1993) is CACC (R. E. Hickson, pers. comm.) as in Appendix One of Hickson (1993) and Hickson *et al.* (1992) rather than GAGG as given in Table 3.1 of Hickson (1993).

G-C Content.

Phylogenetic analysis can be misled when there are differing nucleotide frequencies between taxa (Lockhart *et al.*, 1992, 1993, 1994; Hasegawa and Hashimoto, 1993). Table 4.3 shows the total numbers and frequencies of nucleotides for the *Cyclodina* taxa (Table 4.3a) and for the New Zealand *Leiolopisma* (Table 4.3b) and three overseas skink sequences (Table 4.3c) of Hickson (1993). The *Cyclodina* sequences have very similar numbers and frequencies of nucleotides when compared to each other and also when compared to Hickson (1993)'s sequences. However, these tables show the total numbers and frequencies of nucleotides in the complete sequences, while tree reconstruction methods act primarily on the variable sites. In addition, for many of the phylogenetic analyses carried out in this study, columns were removed from the dataset. Given the high level of similarity between the complete sequences, problems due to unequal nucleotide frequencies are still not expected to arise. However, a new method for reconstructing trees which is not affected by irregular G-C content - the Log Det method (Lockhart *et al.*, 1994) was included as one of the approaches to phylogenetic analysis using the Hadamard conjugation carried out in this study (see later this chapter).

Table 4.1. Taxon name, geographic origin and Museum of New Zealand reptile collection catalogue numbers for the skink sequences of Hickson (1993) and species codes used in this study. Information is from Appendix 1, Hickson (1993).

Taxon name	Geographic Origin	Cat. No.	Code
Cyclodina aenea	Somes I., Wellington	CD 1962	*C. aenea (R)*
Leiolopisma stenotis	Table Hill, Stewart I.	FT 6	*StIsGreen*
L. grande	Central Otago	CD 1055	
L. notosaurus	Masons Bay, Stewart I.	CD 1089	*L.notosauru*
L. lineoocellatum/L. chloronoton	Tekapo, Otago	CD 1217, CD 1218	*L.lin\chl*
L. suteri	Aorangi I., Poor Knights Is.	CD 1027	
L. nigriplantare nigriplantare	Chatham Is.	CD 1058, CD 1060	*L.nnigri*
L. microlepis	Taihape	CD 2123	*L.microlep*
L. acrinasum	Fiordland	CD 826	
L. inconspicuum	Gorge Burn, Southland	CD 1100, CD 1101	*L.inconspic*
L. smithi	Ruamahua-iti I., Alderman Is.	FT 569	
L. zelandicum	Outer Chetwode I., Marlb.*	CD 1952	*L.zelandic*
L. fallai	Great I., Three Kings Is.	FT 598	
L. maccanni	Gorge Burn, Southland	CD 1106-CD 1108	
L. nigriplantare polychroma	Twizel, Canterbury	CD 2126	*L.npoly-Tw*
L. infrapunctatum	Stephens I., Marlb.*	CD 535	*L.infrapunc*
L. otagense	Central Otago	CD 1053	
L. moco	Poor Knights Is.	CD 848, CD 1031	
L. telfairi	Round I., Mauritius	CD 2021	
Lampropholis guichenoti	Australia	CD 536	*La.guichenot*
Tropidoscincus rohssii	New Caledonia	NR 197‡	*T. rohssii*

* Marlborough Sounds.

‡ Australian Museum catalogue number.

Table 4.2. 29 skink 12S rRNA sequences, including the partial *T. rohssii* sequence - 388 bases (including gaps), aligned and shown with the consensus sequence at top. The *Cyclodina* sequences are from this study (with the exception of *C. aenea* [R]), while the rest are from Hickson (1993; Appendix 1). The sequences of Hickson (1993) do, however, include the 'extra' column (at position 359; see explanation in text). Sequences were all obtained via PCR using the 12SAR, 12SBR, SK12SL and SK12SR primers. Conserved positions are in upper case and variable ones in lower case. Positions at which there are transversions are in bold (both in the consensus sequence and in the aligned sequences). A dot indicates that the sequence for that taxa is the same as the consensus sequence at that position. Gaps are indicated by dashes. In addition to the gaps observed in the obtained *Cyclodina* sequences (detailed in Tables 3.2 and 3.3), *L. maccanni* has a single base deletion (placed by Hickson, 1993, at position 25) while *L. telfairi* has two single base deletions (placed by Hickson, 1993, at positions 25 and 60) and a single base insertion (placed by Hickson, 1993, between 202 and 203; at position 204 in this table). Columns containing gaps were omitted from all phylogenetic analyses involving the Hadamard and from some PAUP* analyses. As some of these gaps can not be placed unambiguously, columns 23-30 and 59-61 have also been omitted from all Hadamard and some PAUP* analyses combining *Cyclodina* and non-*Cyclodina* taxa. Columns 202-208 have been omitted in Hadamard and some PAUP* analyses involving *L. telfairi*, as have the first 81 bases of the other 28 taxa in analyses including *T. rohssii*. The 'extra' column was also excluded from all combined *Cyclodina*/non-*Cyclodina* phylogenetic analyses. The *C. aenea* (R) sequence is not included in analyses in this study.

	1 2 3 4 5 6
	123456789012345678901234567890123456789012345678901234567890
	GCucaGCCGU**c**AACAAAgAcAGu**au**aAa**a**uACAauaCUgUUCGCCaGAGAAcUAcAAGc**u**
C.aenea(So)	..c......................-...c....c.........................
C.aenea(PK)	..........u.............--...c.............................**a**
C.whitakeri	...........................g.................g..............
C.alani	.............................c....c.........................
C.macgregori	........................c...g...............................
C.oliveri(Mo)	....g......................g......c.........................
C.oliveri(GM)	..c.g......................g.c..............................
C.ornata	...........................g-...............................
C.aenea(R)	..c...............................c.........................
StIsGreen	.............................c..............................
L.grande	........................c...................................
L.notosauru	......................c**c**..............................u.....
L.lin\chl	...u......**a**..................c........a....................c
L.suteri	........................c...................................
L.nnigri	...................u.........c...g....a.....................
L.microlep	........................c..........g.......................**a**
L.acrinasum	..c.......**a**.............**a**....c....c.........................
L.inconspic	.......................**c**..........c.........................
L.smithi	........................c..................................**a**
L.zelandic	...u......u......a.......g..**c**c....c........................**a**
L.fallai	..................................c................u........
L.maccanni	..........**a**......a......-........g..........................
L.npoly-Tw	...................u.............g.g..a....................**a**
L.infrapunc	..c.......**a**..................c....c.......................u.
L.otagense	..c.......**a**...........c.c.........cg........................
L.moco	..........u...........c......c....cg........................
L.telfairi	..........**a**......a......-....c.............................-
La.guichenot	..c..........................c....c.......................u.

.... continued next page

```
                         7         8         9        10        11        12
                123456789012345678901234567890123456789012345678901234567890
                -AAaaCUcaAAACuCcAAGGACUUGGCGGUGCUCCAcAuCacCCUAGAGGAGCCUGUCCU
C.aenea(So)     -...g..a.....c..............................................
C.aenea(PK)     -...g...g...................................................
C.whitakeri     -...g..a................................c...................
C.alani         u...g..a................................c...................
C.macgregori    -...g..a.................................u..................
C.oliveri(Mo)   -...g..a....................................................
C.oliveri(GM)   -...g..a....................................................
C.ornata        -...g..a....................................................
C.aenea(R)      -...g..a.....c..............................................
StIsGreen       -.....................................c..a..................
L.grande        -.....................................c..a..................
L.notosauru     -.....................................c..a..................
L.lin\chl       -..gg....................................a..................
L.suteri        -.....................................c..a..................
L.nnigri        -..g....g................................a..................
L.microlep      -...........................................................
L.acrinasum     -..gg....................................u..................
L.inconspic     -.....................................c.ga..................
L.smithi        -...........................................................
L.zelandic      -...g.......................................................
L.fallai        -...g..a.....c..............................................
L.maccanni      -.....................................c.ga..................
L.npoly-Tw      -..g...........u.........................a..................
L.infrapunc     -..c........................................................
L.otagense      -.......................................gu..................
L.moco          -...g..g....................................................
L.telfairi      -.......g...........................u....g..................
La.guichenot    -..c........................................................
T.rohssii                            ....................a..................

                         13        14        15        16        17        18
                123456789012345678901234567890123456789012345678901234567890
                AUAAUCGAUACCCCcCGAUCuACCuCACCacUUUUUGAAacUCAGcCUAUAUACCGCCGU
C.aenea(So)     ....................c...................u...................
C.aenea(PK)     ............................................................
C.whitakeri     ..............a...............u.............................
C.alani         ....................c.........u.............................
C.macgregori    ....................c.........u........c....................
C.oliveri(Mo)   ..............a...............u.............................
C.oliveri(GM)   ..............a...............u.............................
C.ornata        ..............................u.............................
C.aenea(R)      ....................c...................u...................
StIsGreen       ............................................................
L.grande        ............................................................
L.notosauru     ........................................u...................
L.lin\chl       ............................................................
L.suteri        ..............a.............................................
L.nnigri        ........................................u...................
L.microlep      ............................................................
L.acrinasum     ............................................................
L.inconspic     ........................................u...................
L.smithi        ............................................................
L.zelandic      ............................................................
L.fallai        ....................c.........u.............................
L.maccanni      ....................c.........u.........u...................
L.npoly-Tw      ..............a..............g.........gu...................
L.infrapunc     ..............................u.............................
L.otagense      ........................c...............u...................
L.moco          ..............................u........gu....u..............
L.telfairi      .............................g..............................
La.guichenot    ..............................u.............................
T.rohssii       .............................g..........u...................
```

.... continued next page

```
                          19        20        21        22        23        24
                 123456789012345678901234567890123456789012345678901234567890
                 CGuCAGCcUACCUUaUGAaAGaa-gcacAguAaGcaAAAuAgucaccaAcUAaaACGuCA
C.aenea(So)      ............g........-..gu..................................
C.aenea(PK)      ................g...g-a.....................................
C.whitakeri      ....................-..........g...........................
C.alani          ..a.................-...................................
C.macgregori     ....................-...a.....g.............g.....
C.oliveri(Mo)    ....................-...................................
C.oliveri(GM)    ....................-..........g...........................
C.ornata         ............g........-.............c.....................
C.aenea(R)       ............g........-..gu..................................
StIsGreen        ....................-.a.....................................c..
L.grande         ....................-.u.........g.....................c..
L.notosauru      ..c.................-au.a....g.....................
L.lin\chl        ....................-...a.................u.....
L.suteri         ....................-.u.u......u..........u.....
L.nnigri         ....................-a..u....g.....
L.microlep       ....................-.u.....
L.acrinasum      ....................-.u.u............a.......u.....
L.inconspic      ..c.................-au.a.....
L.smithi         ....................-.....
L.zelandic       ....................g-au.....
L.fallai         ....................-a.....
L.maccanni       ..c..........g......g.-...a.................u.....
L.npoly-Tw       .......u............-...a....g.....
L.infrapunc      ....................-aa.u.................u......g.....
L.otagense       ................g....-.u.u.................u.....
L.moco           ....................g-c..u..c........c.....
L.telfairi       ............g........acu.a.............c.c.uu.....
La.guichenot     ....................-.a.u.................u......g.....
T.rohssii        ..c....u............-a..a.c..g..........u.uu.....gu.....

                          25        26        27        28        29        30
                 123456789012345678901234567890123456789012345678901234567890
                 GGUCAAGGUGUAGCacAUaaaguGGaAGAGAUGGGCUACACUCUCUccCcCAGAGaAcAC
C.aenea(So)      ....................c.....
C.aenea(PK)      ....................
C.whitakeri      ....................a.....
C.alani          ....................a.....................a.....
C.macgregori     ....................a.....
C.oliveri(Mo)    ....................a.....
C.oliveri(GM)    ....................a.....
C.ornata         ................g..a.....
C.aenea(R)       ....................c.....
StIsGreen        ..........................................a..u.....
L.grande         ..........................................u.....
L.notosauru      ....................
L.lin\chl        ..............u.....a.....................u.....
L.suteri         ..........................................g....
L.nnigri         ....................c.....................u.....
L.microlep       ................gu..................u.....
L.acrinasum      ....................a.....................a.....
L.inconspic      ....................g....c.....................g....
L.smithi         ................gu..................u.....
L.zelandic       ....................
L.fallai         ..............u.....a.....
L.maccanni       ..............u.....a.....
L.npoly-Tw       ..........................u.....
L.infrapunc      ....................a.....u.....
L.otagense       ..........................g.....
L.moco           ....................a.....
L.telfairi       ....................c.....
La.guichenot     ....................a.....a..u.....
T.rohssii        ..............uu......c.....ua.u.....a..
```

..... continued next page

```
                        31        32        33        34        35        36
                1234567890123456789012345678901234567890123456789012345678901234567890
                gAAcaGCAucaAUGAAAcaCuGCucaAAGGuGgAUUUAGuAGUAAGAuaaaCaAGAgaac
C.aenea(So)     ........c.....................c.............................
C.aenea(PK)     ...u.................g....g.....................................
C.whitakeri     ............................c.............................g.
C.alani         ...g....c.....................................................
C.macgregori    a.......c........u...........c.....................................
C.oliveri(Mo)   .................c.....................................................
C.oliveri(GM)   .....................................................................
C.ornata        ........c.....................c.....................................
C.aenea(R)      ........c.....................c.....................................
StIsGreen       .....................................................................
L.grande        ........................u.....................................
L.notosauru     .....................................................................g.
L.lin\chl       .................c.....................................................g.
L.suteri        .................c.....................c...............a...
L.nnigri        .................c.....................................................
L.microlep      ........c.....................................c.....g.......g.
L.acrinasum     .................c.c.....................................g.......g.
L.inconspic     ........................u.....c.................g..........
L.smithi        ........c.....................c........g.g.......g.
L.zelandic      ........c.....................................g........g.
L.fallai        u.....................................................c.........g.
L.maccanni      ........................ug.....................g........g.
L.npoly-Tw      .................c.....................c.....................
L.infrapunc     ...u.....ag......uu.....................................g........g.
L.otagense      ....g.................cu.....c.a.....................g.
L.moco          ........................g....c.....................g.......
L.telfairi      ........c.....................c.....................................
La.guichenot    .................c.....................................g.......g.
T.rohssii       .................u......u.....................c.g.......ugu

                        37        38
                12345678901234567890123456
                uuaucUuAA-AccAGCcCUGGAGCGCGC
C.aenea(So)     .........-..................
C.aenea(PK)     .........-..................
C.whitakeri     .........-..................
C.alani         .........a..................
C.macgregori    .........-..................
C.oliveri(Mo)   .........-..................
C.oliveri(GM)   .........-..................
C.ornata        .........-..................
C.aenea(R)      .........-..................
StIsGreen       ...c.....-......u...........
L.grande        ....a....-......u...........
L.notosauru     .........-..................
L.lin\chl       .........-..................
L.suteri        .........-..................
L.nnigri        .........-..................
L.microlep      c........-.u................
L.acrinasum     .........-......u...........
L.inconspic     c........-..................
L.smithi        ccg......-.u................
L.zelandic      c........-..................
L.fallai        ..g...a..-..................
L.maccanni      .........-..................
L.npoly-Tw      .........-..................
L.infrapunc     .a.......-..................
L.otagense      .........-..................
L.moco          .........-..................
L.telfairi      .........-..................
La.guichenot    .........-......u...........
T.rohssii       c.....a..-..a...............
```

Table 4.3. Total number and frequency of nucleotides for *Cyclodina* taxa, New Zealand *Leiolopisma* taxa and non-New Zealand skink taxa.

a. Total number and frequency of nucleotides for each *Cyclodina* taxon (calculated using all sites). The total numbers of nucleotides differ for each taxon because some of the taxa have gaps (insertions/deletions): *C. aenea* (So) and *C. ornata* have one single base deletion each (384 columns in total), *C. aenea* (PK) has two single base deletions (383 columns total), *C. alani* has two single base insertions (387 columns total) while the remaining four taxa do not have gaps (385 columns total). The means of the total numbers of (a), (c), (g) and (t) sum to 384 while the means of the total numbers of (a+t) and (c+g) sum to 385 because of rounding and because the mean number of total columns is not a whole number. The data for the table was calculated using Prepare (Penny *et al.*, 1993a).

Taxon	Nucleotides in each taxon (top - total number; bottom - frequency)						
	a	c	g	t	a+t	c+g	Total
C. aenea (So)	131	109	75	69	200	184	384
	0.341	0.284	0.195	0.180	0.521	0.479	1
C. aenea (PK)	130	103	77	73	203	180	383
	0.339	0.269	0.201	0.191	0.530	0.470	1
C. whitakeri	131	103	76	75	206	179	385
	0.340	0.268	0.197	0.195	0.535	0.465	1
C. alani	137	105	73	72	209	178	387
	0.354	0.271	0.189	0.186	0.540	0.460	1
C. macgregori	133	104	74	74	207	178	385
	0.345	0.270	0.192	0.192	0.538	0.462	1
C. oliveri (Mo)	133	103	74	75	208	177	385
	0.345	0.268	0.192	0.195	0.540	0.460	1
C. oliveri (GM)	133	103	75	74	207	178	385
	0.345	0.268	0.195	0.192	0.538	0.462	1
C. ornata	131	105	75	73	204	180	384
	0.341	0.273	0.195	0.190	0.531	0.469	1
Mean number:	132	104	75	73	206	179	384.75
Standard Deviation:	2.2	2.1	1.2	2.0	3.0	2.2	1.16
Mean frequency:	0.344	0.271	0.195	0.190	0.534	0.466	1
Standard Deviation:	0.005	0.005	0.004	0.005	0.007	0.007	-

Table 4.3 continued.

b. Total number and frequency of nucleotides for each New Zealand *Leiolopisma* taxon (calculated using all sites - including the 'extra' column). *L. maccanni* has a different number of nucleotides from the other taxa (384 columns in total, rather than 385) - this is because *L. maccanni* has a single base deletion. The means of the total numbers of (a), (c), (g) and (t) sum to 386 rather than 385 because of rounding. The data for the table was calculated using Prepare (Penny *et al.*, 1993a).

Taxon	Nucleotides in each taxon (top - total number; bottom - frequency)						
	a	c	g	t	a+t	c+g	Total
L. stenotis	137 0.356	103 0.268	72 0.187	73 0.190	210 0.545	175 0.455	385 1
L. grande	135 0.351	101 0.262	73 0.190	76 0.197	211 0.548	174 0.452	385 1
L. notosaurus	134 0.348	103 0.268	73 0.190	75 0.195	209 0.543	176 0.457	385 1
L. lin/chl	135 0.351	100 0.260	73 0.190	77 0.200	212 0.551	173 0.449	385 1
L. suteri	135 0.351	102 0.265	72 0.187	76 0.197	211 0.548	174 0.452	385 1
L. n. nigriplantare	132 0.343	102 0.265	73 0.190	78 0.203	210 0.545	175 0.455	385 1
L. microlepis	130 0.338	105 0.273	76 0.197	74 0.192	204 0.530	181 0.470	385 1
L. acrinasum	134 0.348	102 0.265	74 0.192	75 0.195	209 0.543	176 0.457	385 1
L. inconspicuum	131 0.340	106 0.275	75 0.195	73 0.190	204 0.530	181 0.470	385 1
L. smithi	129 0.335	107 0.278	77 0.200	72 0.187	201 0.522	184 0.478	385 1
L. zelandicum	131 0.340	105 0.273	75 0.195	74 0.192	205 0.532	180 0.468	385 1
L. fallai	135 0.351	104 0.270	72 0.187	74 0.192	209 0.543	176 0.457	385 1
L. maccanni	132 0.344	99 0.258	77 0.201	76 0.198	208 0.542	176 0.458	384 1
L. n. polychroma (Tw)	132 0.343	98 0.255	77 0.200	78 0.203	210 0.545	175 0.455	385 1
L. infrapunctatum	134 0.348	98 0.255	74 0.192	79 0.205	213 0.553	172 0.447	385 1
L. otagense	131 0.340	103 0.268	77 0.200	74 0.192	205 0.532	180 0.468	385 1
L. moco	129 0.335	105 0.273	77 0.200	74 0.192	203 0.527	182 0.473	385 1
Mean number:	133	103	75	75	208	177	384.94
Standard Deviation:	2.4	2.7	2.0	2.0	3.5	3.5	0.24
Mean frequency:	0.345	0.266	0.194	0.195	0.540	0.460	1
Standard Deviation:	0.006	0.007	0.005	0.005	0.009	0.009	-

Table 4.3 continued.

c. Total number and frequency of nucleotides for each non-New Zealand skink taxon (calculated using all sites - including the 'extra' column). The total numbers of nucleotides differ for each taxon because of gaps: *L. telfairi* has two single base deletions and one single base insertion (384 columns in total), *La. guichenoti* has no gaps (385 columns total) while the *T. rohssii* sequence is only partial - lacking the first 80 bp (305 columns total). Means and standard deviations for the total numbers of nucleotides were only calculated using the values for *L. telfairi* and *La. guichenoti* because of the missing 80 bp of the *T. rohssii* sequence. The data for the table was calculated using PAUP * Version 4.0.0d34.

Taxon	Nucleotides in each taxon (top - total number; bottom - frequency)						
	a	c	g	t	a+t	c+g	Total
L. telfairi	133	105	74	72	205	179	384
	0.346	0.273	0.193	0.188	0.534	0.466	1
La. guichenoti	132	101	74	78	210	175	385
	0.343	0.262	0.192	0.203	0.545	0.455	1
T. rohssii	98	73	63	71	169	136	305
	0.321	0.239	0.207	0.233	0.554	0.446	1
L. telfairi and *La. guichenoti* only:							
Mean number:	133	103	74	75	208	177	384.5
Standard Deviation:	0.7	2.8	0	4.2	3.5	2.8	0.71
L. telfairi, *La. guichenoti* and *T. rohssii*:							
Mean frequency:	0.337	0.258	0.197	0.208	0.544	0.456	1
Standard Deviation:	0.014	0.017	0.008	0.023	0.010	0.010	-

Secondary Structure.

Different regions and positions in genes have different structural and functional constraints, and so different rates and patterns of change (see for example Kimura, 1983; Noller *et al.*, 1990). The comparison of sequences from a number of different groups and the development and refinement of secondary structure models for different genes provides information on where each molecule varies and where it is constrained, and on the patterns of variability within and between groups. This information has several important applications for evolutionary studies.

Firstly, it can be applied to developing and refining mathematical methods of adjusting for variable rates of change among sites in a sequence. This is necessary if phylogenetic reconstruction is not to be misled, and particularly important for non-protein coding sequences for which methods to correct for among-site rate variation are not well developed (see Hickson, 1993, and Sullivan *et al.*, 1995, for further discussion of these points). Knowledge of the patterns of variability within non-protein coding sequences is also important to another aspect of phylogenetic reconstruction. Just as third (fast) and first and second (slower) codon positions in protein coding sequences are used to investigate more recent and older divergences respectively, regions of relatively high or low variability in non-protein coding sequences may be more informative for examining more closely or more distantly related taxa (Kocher *et al.*, 1989; Thomas and Beckenbach, 1989; Simon, 1991; Simon *et al.*, 1991; Dixon and Hillis, 1993).

Knowledge of secondary structure is also important in relating patterns of sequence evolution and of sequence variability among groups to the structure and function of the molecule. Furthermore, investigations of the tertiary structure and interactions of the products of genes such as the mitochondrial (12S) and nuclear (18S) small subunit rRNA genes in eukaryotes and the prokaryotic (16S) small subunit rRNA gene allow studies of the evolution of these genes to be related to evolutionary studies of the appropriate binding proteins (see Hickson, 1993, for a brief discussion on using the small subunit rRNA genes as a model system for investigating molecular evolution). Finally, placing nucleotide sequences in the context of a secondary structure model can identify possible errors in sequence determination (resulting from cloning, PCR and/or sequencing artifacts or gel reading errors).

In this section, the *Cyclodina* sequences are fitted to the secondary structure models refined by Hickson (1993) and Hickson *et al.* (1996) - providing a further estimate of sequencing fidelity and showing the number and distribution of constant and variable sites and the types of changes which are observed. The occurrence of gaps among the available lizard sequences is discussed in the context of secondary structure. The *Cyclodina* and New Zealand *Leiolopisma* secondary structures are compared and the distribution of variable sites among the two datasets examined - giving an indication of the applicability of Hickson (1993)'s assumptions about the behaviour of the skink sequences he examined to the *Cyclodina* sequences.

Hickson (1993) and Hickson et al. (1996)'s Refined Secondary Structure Models for Domain III of the 12S rRNA Gene.

Structurally, the mitochondrial 12S rRNA gene is divided into four domains (I, II, III and IV), each separated by a highly conserved single-stranded region. Complete 12S rRNA sequences are available for only a few vertebrates, but the use of PCR and the universal 12S rRNA primers (12SAR and 12SBR; Kocher *et al.*, 1989; Palumbi *et al.*, 1991), which amplify the last part of domain II and all of domain III, are leading to an increasing number of 'domain III' sequences for a wide range of taxa.

Hickson (1993) used the small subunit rRNA compilation of Neefs *et al.* (1990; see also Neefs *et al.*, 1991, for the compilation in computer format) as the basis for refining a secondary structure model for 'domain III' (part of domain II plus domain III) of vertebrate 12S rRNA. He aligned sequences from all five classes of vertebrates and then refined a model using comparative sequence analysis, identification of fixed compensatory mutations and free energy predictions. The resulting model is very similar to that of Noller and Woese (1981; see also Gutell *et al.*, 1985; Noller *et al.*, 1990) but provides a more detailed view of the vertebrate 12S rRNA 'domain III'.

A version of Hickson (1993)'s secondary structure model summarising the number and distribution of conserved and non-conserved sites among the vertebrate sequences he examined is presented in Figure 4.1. Two of the sequences included in Hickson (1993)'s alignment were incomplete and not used in refining the secondary structure model; examination of the complete sequences in the alignment indicates that 122 out of approximately 385 sites are conserved (Figure 4.1). Comparison of the *Cyclodina* sequences (Table 3.3) to Figure 4.1 reveals that the *Cyclodina* dataset agrees with all of the conserved positions in the vertebrate secondary structure model. This provides a further estimate of sequencing fidelity - no errors were found in approximately 1000 positions in the *Cyclodina* dataset (122 sites x 8 taxa = 976).

Hickson *et al.* (1996) further refined the above model (for domain III only) using a larger number of 12S rRNA sequences, including many from invertebrates. The consensus sequences of the *Cyclodina* and New Zealand *Leiolopisma* datasets were each fitted to the secondary structure models of Hickson (1993; part of domain II only) and Hickson *et al.* (1996; domain III). The secondary structure model for the consensus sequence of the *Cyclodina* dataset (Table 3.3) is presented in Figure 4.2, while that for the consensus sequence of the New Zealand *Leiolopisma* dataset is presented in Figure 4.3. The New Zealand *Leiolopisma* consensus sequence is shown in Table 4.4. Secondary structure models (Hickson, 1993; Hickson *et al.*, 1996) showing variable sites and what types of changes have occurred at each site (transition, transversion or both) are also given for *Cyclodina* (Figure 4.4), the New Zealand *Leiolopisma* (Figure 4.5) and the New Zealand skinks (Figure 4.6).

Figure 4.1 Secondary structure model showing conserved (□) and non-conserved (■) sites in the region of vertebrate 12S rRNA between PCR primers 12SAR and 12SBR. The figure, which is based on comparisons of 12S rRNA sequences from fish, amphibians, reptiles, birds and mammals, is from Hickson (1993 - Fig. 3.5) - although several changes have been made. Examination of the vertebrate sequence alignment of Hickson (1993 - Table 3.3) indicates that positions 170, 176 and 244 are conserved, while position 270 is non-conserved - the figure shown here has been adjusted accordingly. In addition, the colour coding of Hickson (1993 - Fig. 3.5) has been reversed and symbols for non-conventional pairs and less certain pairings changed to make this figure consistent with others in this chapter. Finally, the base omitted in Hickson (1993) has also been included for consistency (position 357). The bold+underlined numbers (for example, **S25'**) indicate the regions Hickson (1993) divided the molecule into for secondary structure analysis, with S denoting a helical (stem or hairpin) region, L denoting a loop and ' indicating that only the distal (3') arm of a helix is present. Loops also have ' as they are distal to the helices they are numbered after. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (between guanine and uracil or adenine and cytosine), and * represents less certain pairings. Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (|), represent gaps introduced to simplify the figure between helices 34 and 36 and in helices with non-symmetrical bulges (30, 32 and 36) - these are not included in the numbering.

Figure 4.2 Secondary structure model (Hickson, 1993; Hickson *et al.*, 1996) of the region of 12S rRNA between the PCR primers 12SAR and 12SBR for the *Cyclodina* consensus sequence (Table 3.3). Capital letters indicate sites which are conserved among these taxa, while small case letters denote variable sites. The pairs comprising each helix (stem or hairpin region) and the helix numbering follow Hickson (1993) for the part of domain II included in this region (helices 25′, 26, 20′, 29 and 2′) and Hickson *et al.* (1996) for domain III (the remaining helices). The column omitted in Hickson (1993) has been included (position 357). Helix numbers are bold+underlined, with ′ indicating that only the distal (3′) arm of a helix is present. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (G•U and C•A). Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (!) represent gaps introduced into helices with non-symmetrical bulges to simplify the figure; these gaps are not included in the numbering. The sites of insertion in *C. alani* (between positions 60 and 61 and between 367 and 368) are marked by ↑ but are not included in the numbering (which is why the numbering is slightly different to that in Table 3.3), while sites at which deletions are found in *C. aenea* (So), *C. aenea* (PK) and *C. ornata* (positions 25, 26 and 29) are underlined.

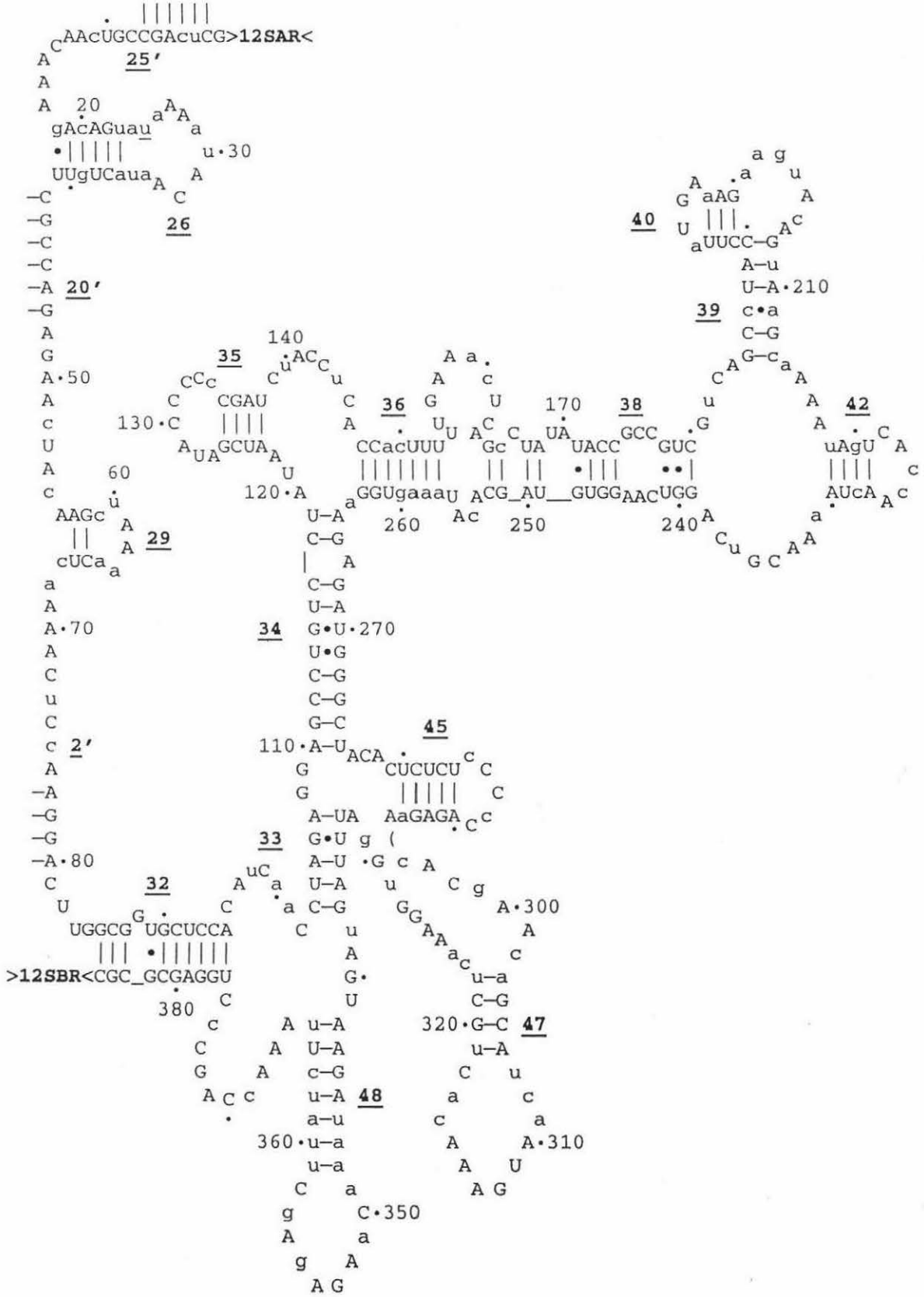


Figure 4.3 Secondary structure model (Hickson, 1993; Hickson *et al.*, 1996) of the region of 12S rRNA between the PCR primers 12SAR and 12SBR for the consensus sequence of 17 New Zealand *Leiolopisma* 12S rRNA sequences (Hickson, 1993; Table 4.4). Capital letters indicate sites which are conserved among these taxa, while small case letters denote variable sites. The pairs comprising each helix (stem or hairpin region) and the helix numbering follow Hickson (1993) for the part of domain II included in this region (helices 25′, 26, 20′, 29 and 2′) and Hickson *et al.* (1996) for domain III (the remaining helices). The column omitted in Hickson (1993) has been included (position 357). Helix numbers are bold+underlined, with ′ indicating that only the distal (3′) arm of a helix is present. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (G•U and C•A). Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (l) represent gaps introduced into helices with non-symmetrical bulges to simplify the figure; these gaps are not included in the numbering. The site of deletion in *L. maccanni* (position 25) is underlined.

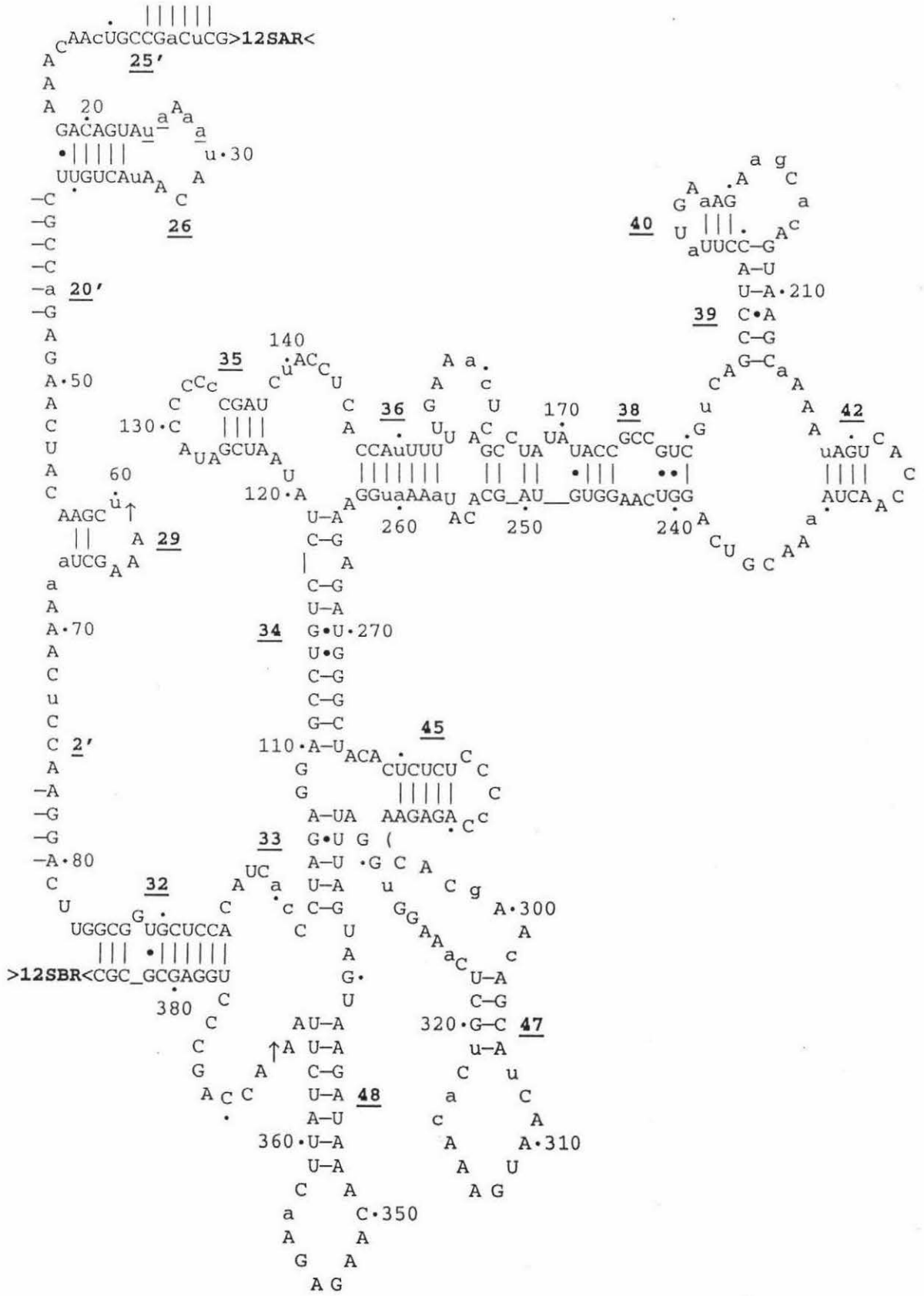


Table 4.4. 17 New Zealand *Leiolopisma* 12S rRNA sequences - 385 bases (including gaps), aligned and shown with the consensus sequence at top. These sequences are from Appendix 1, Hickson (1993) but do, however, include one extra column (see explanation in text; at position 357 in this table). Sequences were all obtained via PCR using the 12SAR, 12SBR, SK12SL and SK12SR primers. Conserved positions are in upper case and variable ones in lower case. Positions at which there are transversions are in bold (both in the consensus sequence and in the aligned sequences). A dot indicates that the sequence for that taxa is the same as the consensus sequence at that position. Gaps are indicated by dashes. *L. maccanni* has a single base deletion (placed by Hickson, 1993, at position 25).

```
                       1         2         3         4         5         6
              123456789012345678901234567890123456789012345678901234567890
              GCucAGCCGUcAACAAAgAcAGuauaAAauACAauaCUgUUCGCCAGAGAAcUAcAAGcu
StIsGreen     .............................c..............................
L.grande      ........................c...................................
L.notosauru   ......................cc..............................u.....
L.lin\chl     ...u......a..................c........a....................c
L.suteri      ........................c...................................
L.nnigri      ...................u.........c...g....a.....................
L.microlep    ........................c..........g.......................a
L.acrinasum   ..c.......a.............a....c....c.........................
L.inconspic   .......................c..........c.........................
L.smithi      ........................c..................................a
L.zelandic    ...u......u......a.......g..cc....c........................a
L.fallai      ..................................c................u........
L.maccanni    ..........a......a......-........g..........................
L.npoly-Tw    ...................u.............g.g..a....................a
L.infrapunc   ..c.......a..................c....c.......................u.
L.otagense    ..c.......a...........c.c.........cg........................
L.moco        ..........u...........c......c....cg........................

                       7         8         9        10        11        12
              123456789012345678901234567890123456789012345678901234567890
              AAaaCUcaAAACuCcAAGGACUUGGCGGUGCUCCACAuCaaCCUAGAGGAGCCUGUCCUA
StIsGreen     .....................................c......................
L.grande      .....................................c......................
L.notosauru   .....................................c......................
L.lin\chl     ..gg........................................................
L.suteri      .....................................c......................
L.nnigri      ..g....g....................................................
L.microlep    ........................................c...................
L.acrinasum   ..gg....................................u...................
L.inconspic   .....................................c.g....................
L.smithi      ........................................c...................
L.zelandic    ...g....................................c...................
L.fallai      ...g..a.....c...........................c...................
L.maccanni    .....................................c.g....................
L.npoly-Tw    ..g...........u.............................................
L.infrapunc   ..c.....................................c...................
L.otagense    .......................................gu...................
L.moco        ...g..g.................................c...................
```

.... continued next page

```
                      13        14        15        16        17        18
              123456789012345678901234567890123456789012345678901234567890
              UAAUCGAUACCCCcCGAUCuACCuCACCacUUUUUGAAacUCAGcCUAUAUACCGCCGUC
StIsGreen     ............................................................
L.grande      ............................................................
L.notosauru   .......................................u....................
L.lin\chl     ............................................................
L.suteri      .............a..............................................
L.nnigri      .......................................u....................
L.microlep    ............................................................
L.acrinasum   ............................................................
L.inconspic   .......................................u....................
L.smithi      ............................................................
L.zelandic    ............................................................
L.fallai      ...................c.........u..............................
L.maccanni    ...................c.........u.........u....................
L.npoly-Tw    .............a..............g.........gu....................
L.infrapunc   .............................u..............................
L.otagense    .......................c...............u....................
L.moco        .............................u........gu....u...............

                      19        20        21        22        23        24
              123456789012345678901234567890123456789012345678901234567890
              GuCAGCcUACCUUaUGAaAGaaguAcAGuAaGcaAAAuAgUCAccAAcUAaAACGuCAGG
StIsGreen     .......................a...............................c....
L.grande      .................................g.....................c....
L.notosauru   .c....................a..a....g.............................
L.lin\chl     .......................c.a.................u................
L.suteri      .........................u......u..........u................
L.nnigri      ......................ac.u....g.............................
L.microlep    ............................................................
L.acrinasum   .........................u.............a.......u............
L.inconspic   .c....................a..a..................................
L.smithi      .......................c....................................
L.zelandic    .....................ga.....................................
L.fallai      ......................ac....................................
L.maccanni    .c...........g......g..c.a.................u................
L.npoly-Tw    ......u................c.a....g.............................
L.infrapunc   ......................aa.u.................u......g.........
L.otagense    .................g.......u...................u..............
L.moco        .....................gcc.u..c........c......................

                      25        26        27        28        29        30
              123456789012345678901234567890123456789012345678901234567890
              UCAAGGUGUAGCAcAUaaagUGGaAGAGAUGGGCUACACUCUCUcCCcCAGAGaAcACgA
StIsGreen     ............................................a..u............
L.grande      ...............................................u............
L.notosauru   ............................................................
L.lin\chl     .............u.....a...........................u............
L.suteri      .....................................................g......
L.nnigri      ...................c...........................u............
L.microlep    ................gu.............................u............
L.acrinasum   ...................a........................a...............
L.inconspic   ..................g....c.............................g......
L.smithi      ................gu.............................u............
L.zelandic    .......................................................u....
L.fallai      .............u.....a......................................u.
L.maccanni    .............u.....a........................................
L.npoly-Tw    ...............................................u............
L.infrapunc   ...................a...........................u............
L.otagense    ............................................g...............
L.moco        ...................a........................................
```

.... continued next page

```
                    31        32        33        34        35        36
           123456789012345678901234567890123456789012345678901234567890
           AcaGCAucaAUGAAAcaCuGCucaAAGGuGgAUUUAGuAGUAAGAuaaaCaAGAgAgCuu
StIsGreen  ..........................................................a...
L.grande   ....................u.....................................a...
L.notosauru ....................................................................
L.lin\chl  ................c.....................................................
L.suteri   ................c.....................c................a.a...
L.nnigri   ................c.....................................a...
L.microlep .....c..............................c.........g.........c.
L.acrinasum ...............c.c..............................g..........
L.inconspic ...................u.....c................g........a.c.
L.smithi   .....c..............................c........g.g.........cc
L.zelandic .....c.....................................g..........c.
L.fallai   ....................................................c.............
L.maccanni ....................ug.....................g...........
L.npoly-Tw ................c.....................c...................a...
L.infrapunc .u.....ag.....uu..............................g...........a
L.otagense ..g................cu.....c.a.....................................
L.moco     ....................g....c.....................g.....a...

                    37        38
           12345678901234567890123 45
           aucUuAAAcCAGCcCUGGAGCGCGC
StIsGreen  .c..........u..........
L.grande   ..a.........u..........
L.notosauru ........................
L.lin\chl  ........................
L.suteri   ........................
L.nnigri   ........................
L.microlep ........u...............
L.acrinasum ............u..........
L.inconspic ........................
L.smithi   g.......u...............
L.zelandic ........................
L.fallai   g...a....................
L.maccanni ........................
L.npoly-Tw ........................
L.infrapunc ........................
L.otagense ........................
L.moco     ........................
```

Comparison of Gaps (Insertions/Deletions) among Available Lizard 12S rRNA Sequences.

All gaps in the *Cyclodina* and *Leiolopisma* datasets fall into loop regions (unpaired regions of four or more bases). The *Cyclodina* dataset contains more sites with insertions/deletions (indels) than the New Zealand *Leiolopisma* dataset. The *Cyclodina* dataset has a total of six gaps (four single base deletions and two single base insertions) - *C. aenea* (So), *C. aenea* (PK) and *C. ornata* all have deletions between positions 25 and 30 (in the loop of helix 26) while *C. alani* has one insertion around position 61 (in the loop of helix 29) and one around position 370 (in the loop between helices 48 and 32; see Figure 4.2). The New Zealand *Leiolopisma* dataset contains only one gap - a single base deletion around position 25 in *L. maccanni* (see Figure 4.3).

Other known lizard sequences include the three non-New Zealand skink sequences (*L. telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* - see Tables 4.1 and 4.2) of Hickson (1993) and the xantusiid lizard dataset of Hedges *et al.* (1991; data not shown). While the xantusiid dataset contains a large number of indels (located throughout the sequence), only one of the three non-New Zealand skink sequences has gaps. *L. telfairi* has three indels: two single base deletions - one at position 25 (in the loop of helix 26) and one at position 60 (in the loop of helix 29) - and a single base insertion around position 204 (in the loop between helices 40 and 39). However, it is possible that gaps occur within the first 80 bases of the *T. rohssii* sequence, which have not as yet been successfully sequenced, particularly as this is the area with the most gaps in the skink dataset.

The low number of indels in the New Zealand *Leiolopisma* dataset compared to the *Cyclodina* dataset may possibly be due to sequencing and/or reading errors. Or it could be a sampling effect - resulting from the small sample size. Alternatively, it may reflect a genuine difference in the mechanism of change between the two New Zealand skink genera and/or - considering the number of indels in the *L. telfairi* sequence - a closer relationship between *L. telfairi* and the *Cyclodina* taxa than between the two New Zealand genera.

Gaps cannot yet be handled by certain phylogenetic analysis programmes, including the Hadamard, therefore columns containing gaps have been excluded from many of the phylogenetic analyses presented later this chapter. However, with or without phylogenetic analysis by a given programme, gaps can be used as indicators for phylogenetic relationships. *C. ornata*, *C. aenea* (So) and *C. aenea* (PK) all have deletions in the same area (the loop of helix 26) which may indicate a closer phylogenetic relationship between these taxa - particularly between *C. aenea* (So) and *C. aenea* (PK). *L. maccanni* and *L. telfairi* also have a

deletion each in this area - probably at the same site as *C. aenea* (PK)'s first deletion (rather than in the same place as *C. aenea* [PK]'s second deletion and the deletions of *C. ornata* and *C. aenea* [So]). However, due to the small sample sizes of the *Cyclodina* and non-New Zealand skink datasets, such conclusions should be treated warily. *L. telfairi* also has a deletion at position 60 (in the loop of helix 29) while *C. alani* has an insertion in this region - however, as these are opposite changes, this undoubtedly does not indicate a close relationship, but rather a variable region in terms of secondary structure.

Conserved and Variable sites in the Cyclodina and New Zealand Leiolopisma Datasets and the Types of Changes that have Occurred.

Of the 387 columns in the eight taxa *Cyclodina* dataset: 341 are constant sites, 29 are singleton sites (sites at which only one code occurs more than once) - two of which also contain a single deletion, 14 are parsimony sites (sites at which two or more codes occur more than once), one is a site with two deletions - but with no other change, and two are insertion sites (see Table 3.3 and Figure 4.2). This gives a total of 46 variable sites if sites at which insertions/deletions are the only change are counted as variable sites, 44 if insertion sites are not counted and 43 if neither are counted (see Table 3.3 and Figure 4.2).

The 17 taxa New Zealand *Leiolopisma* dataset contains 385 columns of which 289 are constant sites, 44 are singleton sites and 52 are parsimony sites (one of which also contains a single deletion) - giving a total of 96 variable sites in all (see Table 4.4 and Figure 4.3).

Of the 43 sites in the *Cyclodina* dataset at which more than one nucleotide code is observed - 32 have transitions only (20 singleton sites - two with a single deletion - and 12 parsimony sites), nine have transversions only (seven singleton sites and two parsimony sites), and two have both transitions and transversions (both are singleton sites, that is, each site has one transition and one transversion only) (see Figure 4.4).

The New Zealand *Leiolopisma* dataset has 96 variable sites - of these, 74 have transitions only (36 singleton sites and 38 parsimony sites), 13 have transversions only (seven singleton sites and six parsimony sites), and nine have both transitions and transversions (one singleton site and eight parsimony sites) (see Figure 4.5). A secondary structure model summarising variable site information from the New Zealand skink dataset (combined *Cyclodina* and New Zealand *Leiolopisma* datasets) is presented in Figure 4.6.

Figure 4.4 Secondary structure model of the region of 12S rRNA between the PCR primers 12SAR and 12SBR for the *Cyclodina* consensus sequence (Table 3.3) showing conserved (□) sites, sites with transitions (▲), sites with transversions (▼), and sites with both transitions and transversions (◆). The figure is based on that of Hickson (1993 - Fig. 4.17), but the pairs comprising each helix (stem or hairpin region) and the helix numbering follow Hickson *et al.* (1996) for helices 32 to 48 (domain III). The column omitted in Hickson (1993) has been included (position 357). Helix numbers are bold+underlined, with ′ indicating that only the distal (3′) arm of a helix is present. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (G•U and C•A). Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (|) represent gaps introduced into helices with non-symmetrical bulges to simplify the figure; these gaps are not included in the numbering. The sites of insertion in *C. alani* (between positions 60 and 61 and between 367 and 368) are marked by ↑ but are not included in the numbering (which is why the numbering is slightly different to that in Table 3.3), while sites at which deletions are found in *C. aenea* (So), *C. aenea* (PK) and *C. ornata* (positions 25, 26 and 29) are underlined. Position 26 is shown as a constant site in spite of the deletion placed at this position because only sites with transitions and/or transversions are being considered as variable in this diagram.

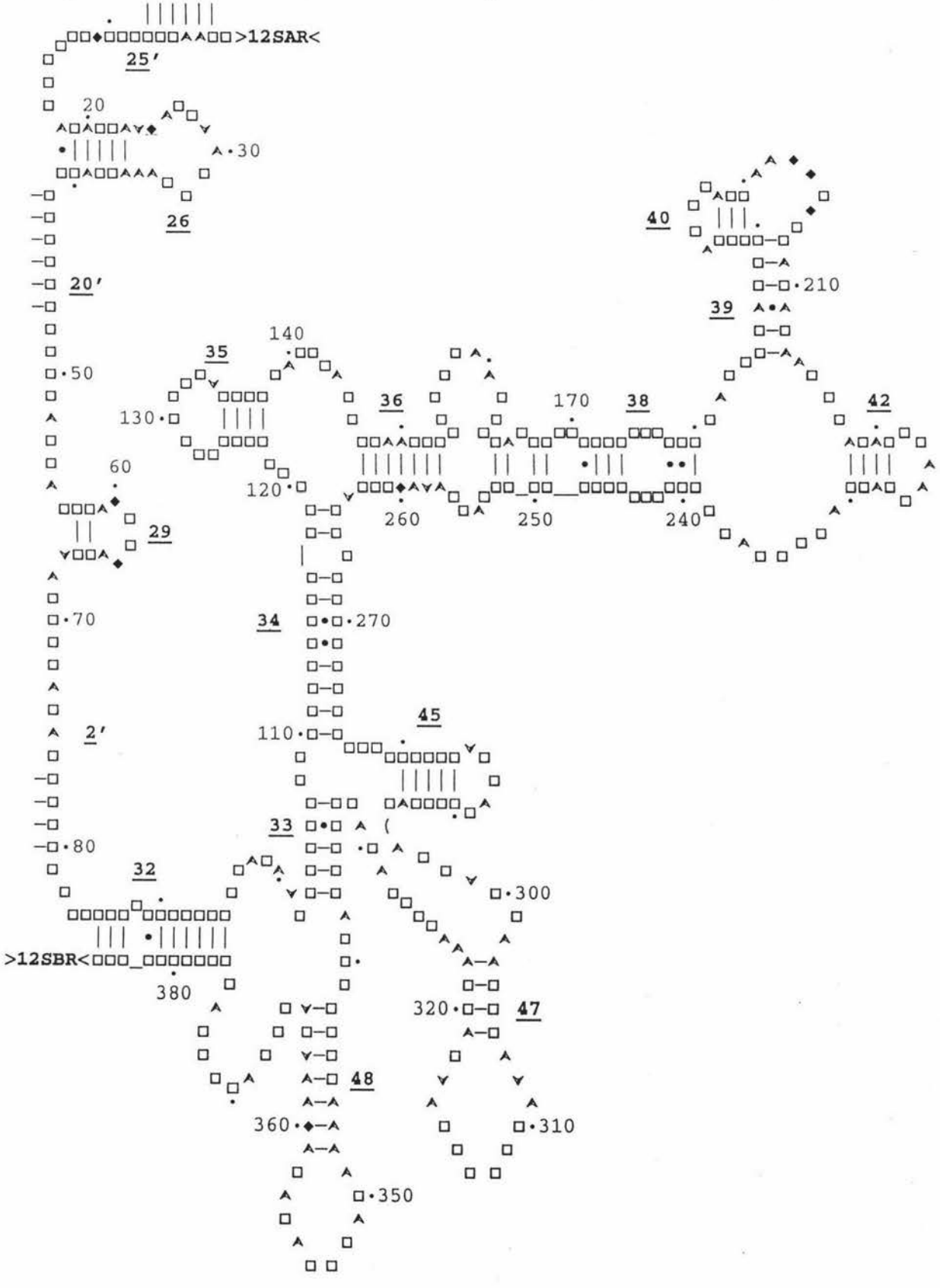


Figure 4.5 Secondary structure model of the region of 12S rRNA between the PCR primers 12SAR and 12SBR for the consensus sequence of 17 New Zealand *Leiolopisma* 12S rRNA sequences (Hickson, 1993; Table 4.4) showing conserved (□) sites, sites with transitions (^), sites with transversions (v), and sites with both transitions and transversions (◆). The figure is based on that of Hickson (1993 - Fig. 4.17), but the pairs comprising each helix (stem or hairpin region) and the helix numbering follow Hickson *et al.* (1996) for helices 32 to 48 (domain III). The column omitted in Hickson (1993) has been included (position 357). Helix numbers are bold+underlined, with ′ indicating that only the distal (3′) arm of a helix is present. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (G•U and C•A). Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (|) represent gaps introduced into helices with non-symmetrical bulges to simplify the figure; these gaps are not included in the numbering. The site of deletion in *L. maccanni* (position 25) is underlined.

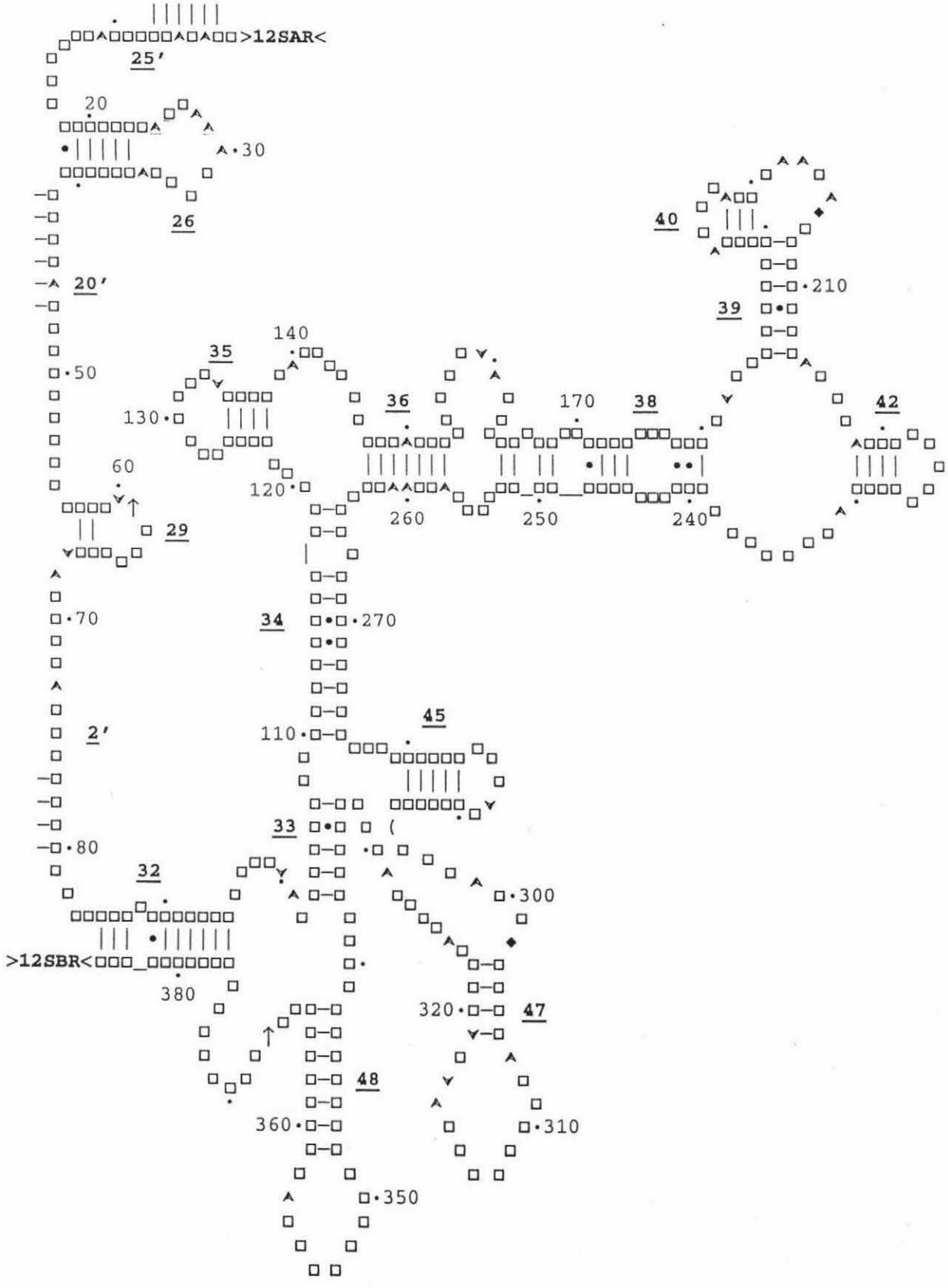
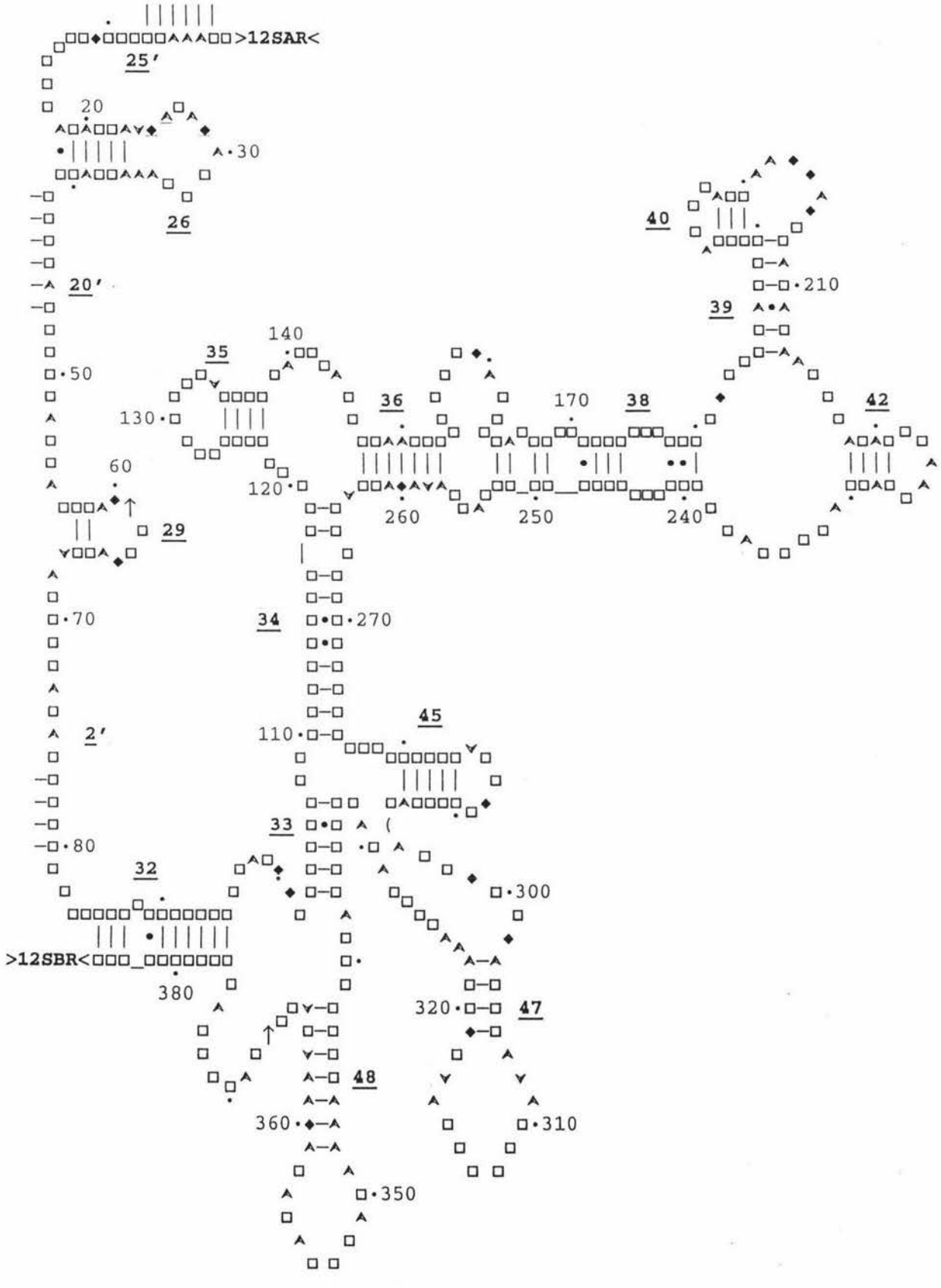


Figure 4.6 Secondary structure model of the region of 12S rRNA between the PCR primers 12SAR and 12SBR for the consensus sequence of available New Zealand skink 12S rRNA sequences (17 New Zealand *Leiolopisma* taxa and eight *Cyclodina* taxa; see sequences in Table 4.2) showing conserved (□) sites, sites with transitions (^), sites with transversions (v), and sites with both transitions and transversions (◆). The figure is based on that of Hickson (1993 - Fig. 4.17), but the pairs comprising each helix (stem or hairpin region) and the helix numbering follow Hickson *et al.* (1996) for helices 32 to 48 (domain III). The column omitted in Hickson (1993) has been included (position 357). Helix numbers are bold+underlined, with ′ indicating that only the distal (3′) arm of a helix is present. Conventional nucleotide pairing is represented by | and –, while • denotes non-conventional pair-bonds (G•U and C•A). Every tenth base is marked by ., and numbers are also given where possible. Underscores (_) and, in one case, a side score (l) represent gaps introduced into helices with non-symmetrical bulges to simplify the figure; these gaps are not included in the numbering. The sites of insertion in *C. alani* (between positions 60 and 61 and between 367 and 368) are marked by ↑ but are not included in the numbering (which is why the numbering is slightly different to that in Table 4.2), while sites at which deletions are found in *L. maccanni*, *C. aenea* (So), *C. aenea* (PK) and *C. ornata* (positions 25, 26 and 29) are underlined.



Comparison of the Secondary Structures of Cyclodina and the New Zealand Leiolopisma.

Comparison of Figures 4.2 and 4.3 indicate that the *Cyclodina* secondary structure is almost identical that of *Leiolopisma*. Only 7/385 (1.8%) of sites differ between the two consensus sequences (positions 64, 67, 101, 150, 204, 260 and 357 in Figures 4.2 and 4.3). All but two of these are unpaired (the exceptions being 150 and 260 which are paired to each other; see later) and all are variable in one or both genera. When the actual sequences are examined (Table 3.3, but note that the numbering is slightly different to that in Figures 4.2 and 4.3, and Table 4.4) it is found that at none of these positions do all eight *Cyclodina* taxa differ from all 17 *Leiolopisma* taxa - that is, subgroup/s of *Leiolopisma* taxa share the *Cyclodina* code/s, and therefore the same secondary structure, at all positions.

There are five sites (5/385 = 1.3%) at which *Cyclodina* is variable and *Leiolopisma* is not (positions 5, 28, 46, 205 and 261 in Figures 4.2 and 4.3). However, again none of these involve all eight *Cyclodina* taxa differing from all 17 *Leiolopisma* taxa. Two of these positions involve more than one taxon (5 and 28) while the other three (46, 205 and 261) involve one taxon only (but are unlikely to be the result of PCR and/or sequencing errors due to the exhaustive sequencing and checking procedure carried out in this study; see Chapter Three). Two of these positions are unpaired (position 28, which is in the loop of helix 26, and position 205, which is in the loop between helices 39 and 40), while the other three are paired (position 5 is in helix 25′, position 46 is in helix 20′ while position 261 is in helix 36).

As positions 5 and 46 are in regions of domain II for which the complementary sites are not available, it cannot be established for certain whether the changes have any effect on secondary structure. However it seems likely that both positions are A-U pairs, with the changes giving G-U pairs and therefore no significant disruption of the secondary structure. In the case of position 261, pairing is maintained provided a non-standard A-C pair is allowed (interestingly, *L. n. polychroma* has a change at the complementary site, position 149, to give a G-U pair). Thus *Cyclodina* would not appear to have a significantly different secondary structure from *Leiolopisma* due to variability at these sites.

The main difference between the two datasets is the number of sites which are variable in *Leiolopisma* and conserved in *Cyclodina* (57/385 = 15%). Again subgroups of *Leiolopisma* share the same code as the *Cyclodina* taxa, and thus the same secondary structure, at all sites. The higher level of variation in *Leiolopisma* than *Cyclodina* is probably a sampling effect, reflecting the small size of the *Cyclodina* dataset. To investigate whether the *Cyclodina* sequences are behaving in a similar way to the *Leiolopisma* sequences, and so whether Hickson (1993)'s findings on the skinks are also applicable to *Cyclodina*, the distribution of variable sites in the two datasets were examined.

Distribution of Variable Sites.

The distribution of variable sites among the regions defined by Hickson (1993; see Figure 4.1) was determined for the *Cyclodina* dataset and compared to the distribution of *Leiolopisma* variable sites (Table 4.5; see also Hickson, 1993). The patterns of variability are very similar for the two genera - with S30, S31, S32 and S36 being more conserved in both genera than would be expected by chance (S20′ and S33 are also conserved in *Leiolopisma*) while S26, L38′ and S45 are more variable (S34 is also variable in *Leiolopisma*) (see Table 4.5). The two genera do, however, have a noticeably different pattern of variability in one region - S46 - which is conserved in *Cyclodina* and variable in *Leiolopisma*. This may represent a genus-specific pattern of variability or be the result of sampling error.

These patterns also agree well with the findings of Hickson (1993) for the vertebrates - that helices 31-33 and 36 (and, to a lesser extent, helix 30) and regions S20′ and S2′ are conserved, while regions S26 (particularly), S34, SL40, S45 and S46 are variable. He also noted possible taxon-specific patterns of variability - with higher levels of variability in region SL40 in the ratites, S45 in the great apes and L43′ in the mammals, while L38′ was more variable in the skinks (which the *Cyclodina* dataset also agrees with).

Based on the extreme similarities in secondary structure between the two genera and the apparent correspondence in distribution of variable and conserved sites, it is assumed that the *Cyclodina* sequences are evolving in much the same way as the skink sequences investigated by Hickson (1993). This in turn implies that Hickson (1993)'s assumptions for these sequences - for example, that the skink sequences do not appear to be evolving in a different way from other vertebrate 12S rRNA sequences - also hold true for the *Cyclodina* sequences.

Investigation of Paired Variable Sites in the Cyclodina Dataset.

Investigation of the paired variable sites in the *Cyclodina* dataset indicates one double (paired) change at positions 150 and 260 (helix 36 in Figures 4.2 and 4.3) - with *C. aenea* (So) and *C. aenea* (PK) sharing a G-C pair and the rest of the *Cyclodina* taxa sharing an A-U pair. When the *Leiolopisma* taxa are examined, most have the same pair as *C. aenea* (So) and *C. aenea* (PK), while *L. fallai*, *L. maccanni*, *L. infrapunctatum* and *L. moco* share the same codes as the rest of the *Cyclodina* taxa. However, it is unclear what the significance of this change is with regards to the phylogenetic relationships of the *Cyclodina* taxa - with weighting schemes down-weighting such sites because of the non-independent substitutions involved (Wheeler and Honeycutt, 1988; Kraus *et al.*, 1992; Dixon and Hillis, 1993) and weighting such sites more heavily on the basis that such sites tend to evolve more slowly (Miyamoto *et al.*, 1994) both in existence (see also Van de Peer *et al.*, 1993).

Table 4.5. Distribution of variable sites in the *Cyclodina* and *Leiolopisma* datasets. Regions are those of Hickson (1993) (see Figure 4.1). Expected numbers of variable sites in each region were calculated using the formula:

$$\frac{\text{Total number of variable sites in the dataset x total number of sites in the region}}{\text{Total number of sites in the dataset (385)}}$$

For example, the expected number of variable sites in region S25′ in *Cyclodina* $= \frac{43 \times 17}{385} = 1.9$.

Region	Number of bases	Encompasses positions*	Number of variable positions (*Cyclodina*)		Number of variable positions (*Leiolopisma*)	
			Observed	Expected	Observed	Expected
S25′	17	1-17	3	1.9	3	4.2
S26	24	18-41	5	2.7	12	6.0
S20′	15	42-56	1	1.7	2	3.7
S29	10	57-66	1	1.1	4	2.5
S2′	18	67-84	3	2.0	4	4.5
S30	21	85-95 + 376-385	0	2.3	0	5.2
L30′	6	96-101	2	0.7	3	1.5
S31	15	102-108 + 333-340	0	1.7	1	3.7
S32	26	109-120 + 265-278	0	2.9	0	6.5
S33	18	121-138	1	2.0	1	4.5
L33′	8	139-146	1	0.9	2	2.0
S34	23	147-156 + 252-264	4	2.6	8	5.7
L34′	9	157-165	2	1.0	3	2.2
S36	34	166-182 + 235-251	1	3.8	2	8.5
S37+38	22	185-200 + 208-213	2	2.5	6	5.5
L38′	7	201-207	4	0.8	5	1.7
SL40	23	183,184 + 214-234	3	2.6	7	5.7
S43	16	279-294	1	1.8	3	4.0
L43′	8	295-302	2	0.9	3	2.0
S45	20	303-322	4	2.2	8	5.0
L45′	10	323-332	2	1.1	4	2.5
S46	26†	341-366	1	2.9	13	6.5
L46′	9	367-375	0	1.0	2	2.2
Total:	385		43		96	

* Hickson (1993) did not specifically define the boundaries of his regions in this way for the skinks - instead giving a general set of boundaries for the vertebrates in his vertebrate sequence alignment and labelling the regions as shown in Figure 4.1. He did however give the number of bases in each region for the skinks as shown in the second column above, and the boundaries of the regions were deduced from these and from his figures.

† Includes the 'extra' base (the region has 25 bases in Hickson, 1993).

Secondary Structure - Conclusions.

Several conclusions can be drawn from the fitting of the *Cyclodina* sequences to the secondary structure models of Hickson (1993) and Hickson *et al.* (1996). The use of secondary structure helps check sequencing accuracy, additional to that from a multiple alignment. In this case, the comparison of the *Cyclodina* sequences to the vertebrate conserved sites figure (Figure 4.1) of Hickson (1993) gave a further estimate of sequencing fidelity for the *Cyclodina* dataset of approximately less than one error in 1000.

The *Cyclodina* secondary structure is almost identical to that of the New Zealand *Leiolopisma* and is fundamentally similar to the vertebrate secondary structure model of Hickson (1993) and the refined secondary structure model for multicellular animals of Hickson *et al.* (1996). In addition, the patterns of variability in the *Cyclodina* dataset agree very well with those of the New Zealand *Leiolopisma* (Table 4.5), and with those determined by Hickson (1993) for the vertebrates. This would tend to suggest that Hickson (1993)'s findings on the behaviour of the skink sequences he investigated are also applicable to *Cyclodina* and that the *Cyclodina* sequences, too, are not evolving in a significantly different way from other vertebrate 12S rRNA sequences.

Cyclodina indels (deletions at positions 25, 26 and 29 in *C. aenea* (So), *C. aenea* (PK) and *C. ornata* and insertions between positions 60 and 61 and between 367 and 368 in *C. alani*) occur in loops rather than in stems or single-stranded conserved regions. Gaps provide some support for closer relationships between *C. aenea* (So), *C. aenea* (PK) and possibly *C. ornata* and between *C. aenea* (PK), *L. maccanni* and *L. telfairi*. However, as the gaps supporting these relationships fall in helix 26, one of the most variable regions among both the skinks and the vertebrates (see Table 4.5 and Hickson, 1993), these conclusions are tentative. The low number of indels in the New Zealand *Leiolopisma* dataset compared to the relatively higher number among *Cyclodina* taxa and in the *L. telfairi* sequence may also reflect a closer relationship between *L. telfairi* and *Cyclodina*. However, the small sample size of the *Cyclodina* and non-New Zealand skink datasets means that all such conclusions should be treated cautiously.

The use of secondary structure allows detection of double (paired) changes in stems. The two *C. aenea* sequences along with most of the *Leiolopisma* taxa share a G-C pair at positions 150 and 260 in stem 36 (Figures 4.2 and 4.3), while *L. fallai*, *L. maccanni*, *L. infrapunctatum* and *L. moco* and the rest of the *Cyclodina* taxa share an A-U pair at these positions. However, at present it is not clear whether such positions should be weighted more or less heavily, and indeed, it now appears that simply differentially weighting stem and loop sites in phylogenetic analysis of rRNA genes may be an oversimplification (see Van de Peer *et al.*, 1993; see also Hickson, 1993).

Pairwise Comparisons (Total Differences, Transitions and Transversions) between the Skink Taxa.

Information on the number of total differences, transitions and transversions in pairwise comparisons between the skink taxa can be used both in assessing the possibility of saturation of transitions within the New Zealand skink sequences and in estimating times of divergence among the skinks. Pairwise comparisons giving numbers of total differences, transitions and transversions were carried out for the *Cyclodina* and the New Zealand *Leiolopisma* datasets, the transversion/transition (Tv/Ts) ratio was calculated for each pair and the means and standard deviations of each of the different measures were calculated for each taxon. This information is presented in Tables 4.6 (*Cyclodina*) and 4.7 (*Leiolopisma*). Taxa from the two genera were then compared and the numbers of total differences, transitions and transversions are presented in Table 4.8. Pairwise comparisons between the New Zealand skinks and the three overseas skinks sequences of Hickson (1993) - *L. telfairi*, *La. guichenoti* and *T. rohssii* - were also carried out (see Table 4.9). Insertions/deletions were not scored as differences, as the emphasis here is on numbers of transitions and transversions.

C. ornata and *C. aenea* (So) have no transversions between them (see Table 4.6) - however, a close relationship can not necessarily be assumed from this because of the high variance associated with small numbers of observations (standard deviation due to sampling error).

Saturation of Transitions?

If the skink 12S rRNA sequences are saturated with substitutions, then the phylogenies obtained from them will not be reliable. One indication of saturation is that the proportion of transitions decreases as the total number of pairwise differences increases (Wilson *et al.*, 1985; DeSalle *et al.*, 1987; Miyamoto and Boyle, 1989; Irwin *et al.*, 1991), and thus the transversion to transition (Tv/Ts) ratio increases. Therefore it is useful to calculate the expected Tv/Ts ratio at saturation (R_∞), and to compare this to the observed Tv/Ts ratios. Holmquist (1983 - equation 11) gives a formula for calculating R (Tv/Ts after a very long evolutionary period) or R_∞ :

$$R_\infty = \frac{(A_\infty + G_\infty)(C_\infty + T_\infty)}{A_\infty G_\infty + C_\infty T_\infty}$$

where A_∞ is the expected equilibrium proportion of adenosine, and so on.

For the New Zealand *Cyclodina*, using the mean nucleotide frequencies from Table 4.3a,

$$R_\infty = \frac{(0.344 + 0.195)(0.271 + 0.190)}{(0.344 \times 0.195) + (0.271 \times 0.190)} = 2.10$$

Table 4.6. Pairwise comparisons between *Cyclodina* taxa (observed differences). The complete 387 column dataset (see Tables 3.2 and 3.3) was used, but these numbers do not include differences resulting from insertions/deletions.

a. Number of total differences in lower left of matrix, number of transversions (Tv) in upper right.

		1	2	3	4	5	6	7	8
1	*C.aenea*(So)	-	3	2	4	2	2	1	0
2	*C.aenea*(PK)	21	-	5	7	5	5	4	3
3	*C.whitakeri*	19	20	-	4	4	2	1	2
4	*C.alani*	14	18	13	-	6	6	5	4
5	*C.macgregori*	19	24	15	16	-	4	3	2
6	*C.oliveri*(Mo)	17	18	8	11	17	-	1	2
7	*C.oliveri*(GM)	16	17	7	12	16	5	-	1
8	*C.ornata*	14	19	9	12	13	9	10	-

b. Number of transitions (Ts) lower left, transversion/transition (Tv/Ts) ratios upper right.

		1	2	3	4	5	6	7	8
1	*C.aenea*(So)	-	0.17	0.12	0.40	0.12	0.13	0.07	0
2	*C.aenea*(PK)	18	-	0.33	0.64	0.26	0.38	0.31	0.19
3	*C.whitakeri*	17	15	-	0.44	0.36	0.33	0.17	0.29
4	*C.alani*	10	11	9	-	0.60	1.20	0.71	0.50
5	*C.macgregori*	17	19	11	10	-	0.31	0.23	0.18
6	*C.oliveri*(Mo)	15	13	6	5	13	-	0.25	0.29
7	*C.oliveri*(GM)	15	13	6	7	13	4	-	0.11
8	*C.ornata*	14	16	7	8	11	7	9	-

c. Mean number of differences ± standard deviations.

	Mean number of total differences	Mean number of transversions (Tv)	Mean number of transitions (Ts)	Mean Tv/Ts ratio
C.aenea(So)	17.1 ± 2.7	2.0 ± 1.3	15.1 ± 2.7	0.14 ± 0.13
C.aenea(PK)	19.6 ± 2.4	4.6 ± 1.4	15.0 ± 2.9	0.33 ± 0.16
C.whitakeri	13.0 ± 5.3	2.9 ± 1.5	10.1 ± 4.4	0.29 ± 0.11
C.alani	13.7 ± 2.5	5.1 ± 1.2	8.6 ± 2.1	0.64 ± 0.27
C.macgregori	17.1 ± 3.5	3.7 ± 1.5	13.4 ± 3.4	0.29 ± 0.16
C.oliveri(Mo)	12.1 ± 5.2	3.1 ± 1.9	9.0 ± 4.5	0.41 ± 0.36
C.oliveri(GM)	11.9 ± 4.7	2.3 ± 1.7	9.6 ± 4.2	0.26 ± 0.21
C.ornata	12.3 ± 3.5	2.0 ± 1.3	10.3 ± 3.5	0.22 ± 0.16

Table 4.7. Pairwise comparisons between the New Zealand *Leiolopisma* (observed differences). The complete 385 column dataset (including the 'extra' column; see Table 4.4) was used, but these numbers do not include differences resulting from insertions/deletions.

a. Number of total differences in lower left of matrix, number of transversions (Tv) in upper right.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	*L.stenotis*	-	3	4	5	4	4	5	5	5	5	5	6	4	6	7	3	5
2	*L.grande*	8	-	3	4	3	3	4	6	4	4	4	5	3	5	8	4	4
3	*L.notosaurus*	16	16	-	3	4	4	5	7	1	5	5	6	2	4	9	5	5
4	*L.lin\chl*	18	21	21	-	3	3	6	4	4	6	6	7	1	3	6	4	6
5	*L.suteri*	16	14	17	20	-	2	5	5	5	5	5	6	4	2	7	5	5
6	*L.n.nigri*	17	20	17	15	20	-	5	5	5	5	5	6	4	4	7	5	5
7	*L.microlepis*	19	17	21	23	19	25	-	6	6	0	2	5	5	5	8	4	4
8	*L.acrinasum*	20	23	25	18	23	23	25	-	8	6	6	7	4	7	6	2	6
9	*L.inconspic*	21	19	13	28	20	24	26	30	-	6	6	7	3	5	10	6	6
10	*L.smithi*	21	20	24	24	22	26	5	28	29	-	2	5	5	5	8	4	4
11	*L.zelandicum*	23	25	23	22	27	27	20	25	24	23	-	5	5	5	8	4	4
12	*L.fallai*	24	25	23	22	27	26	26	26	28	25	24	-	6	8	9	5	3
13	*L.maccanni*	25	23	20	21	24	26	30	30	21	31	30	25	-	4	7	3	5
14	*L.n.poly-Tw*	23	24	22	20	21	13	23	31	29	26	33	33	30	-	9	7	7
15	*L.infrapunc*	25	29	28	22	28	25	30	24	31	31	27	27	29	35	-	6	8
16	*L.otagense*	27	24	23	30	26	30	26	24	24	31	31	32	28	33	32	-	4
17	*L.moco*	27	30	26	27	30	26	31	28	29	34	26	25	30	31	30	28	-

b. Number of transitions (Ts) lower left, transversion/transition (Tv/Ts) ratios upper right.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	*L.stenotis*	-	0.60	0.33	0.38	0.33	0.31	0.36	0.33	0.31	0.31	0.28	0.33	0.19	0.35	0.39	0.13	0.23
2	*L.grande*	5	-	0.23	0.24	0.27	0.18	0.31	0.35	0.27	0.25	0.19	0.25	0.15	0.26	0.38	0.20	0.15
3	*L.notosaurus*	12	13	-	0.17	0.31	0.31	0.31	0.39	0.08	0.26	0.28	0.35	0.11	0.22	0.47	0.28	0.24
4	*L.lin\chl*	13	17	18	-	0.18	0.25	0.35	0.29	0.17	0.33	0.38	0.47	0.05	0.18	0.38	0.15	0.29
5	*L.suteri*	12	11	13	17	-	0.11	0.36	0.28	0.33	0.29	0.23	0.29	0.20	0.11	0.33	0.24	0.20
6	*L.n.nigri*	13	17	13	12	18	-	0.25	0.28	0.26	0.24	0.23	0.30	0.18	0.44	0.39	0.20	0.24
7	*L.microlepis*	14	13	16	17	14	20	-	0.32	0.30	0	0.11	0.24	0.20	0.28	0.36	0.18	0.15
8	*L.acrinasum*	15	17	18	14	18	18	19	-	0.36	0.27	0.32	0.37	0.15	0.29	0.33	0.09	0.27
9	*L.inconspic*	16	15	12	24	15	19	20	22	-	0.26	0.33	0.33	0.17	0.21	0.48	0.33	0.26
10	*L.smithi*	16	16	19	18	17	21	5	22	23	-	0.10	0.25	0.19	0.24	0.35	0.15	0.13
11	*L.zelandicum*	18	21	18	16	22	22	18	19	18	21	-	0.26	0.20	0.18	0.42	0.15	0.18
12	*L.fallai*	18	20	17	15	21	20	21	19	21	20	19	-	0.32	0.32	0.50	0.19	0.14
13	*L.maccanni*	21	20	18	20	20	22	25	26	18	26	25	19	-	0.15	0.32	0.12	0.20
14	*L.n.poly-Tw*	17	19	18	17	19	9	18	24	24	21	28	25	26	-	0.35	0.27	0.29
15	*L.infrapunc*	18	21	19	16	21	18	22	18	21	23	19	18	22	26	-	0.23	0.36
16	*L.otagense*	24	20	18	26	21	25	22	22	18	27	27	27	25	26	26	-	0.17
17	*L.moco*	22	26	21	21	25	21	27	22	23	30	22	22	25	24	22	24	-

c. Mean number of differences ± standard deviations.

	Mean number of total differences	Mean number of transversions (Tv)	Mean number of transitions (Ts)	Mean Tv/Ts ratio
L.stenotis	20.6 ± 5.0	4.8 ± 1.1	15.9 ± 4.6	0.32 ± 0.10
L.grande	21.1 ± 5.6	4.2 ± 1.3	16.9 ± 4.9	0.27 ± 0.11
L.notosaurus	20.9 ± 4.2	4.5 ± 1.9	16.4 ± 2.9	0.27 ± 0.10
L.lin\chl	22.0 ± 3.8	4.4 ± 1.6	17.6 ± 3.7	0.27 ± 0.11
L.suteri	22.1 ± 4.6	4.4 ± 1.4	17.8 ± 3.9	0.25 ± 0.08
L.n.nigri	22.5 ± 4.9	4.5 ± 1.2	18.0 ± 4.3	0.26 ± 0.08
L.microlepis	22.9 ± 6.3	4.7 ± 1.8	18.2 ± 5.2	0.26 ± 0.10
L.acrinasum	25.2 ± 3.6	5.6 ± 1.5	19.6 ± 3.2	0.29 ± 0.08
L.inconspic	24.8 ± 4.9	5.4 ± 2.0	19.3 ± 3.6	0.28 ± 0.09
L.smithi	25.0 ± 6.7	4.7 ± 1.8	20.3 ± 5.7	0.23 ± 0.09
L.zelandicum	25.6 ± 3.5	4.8 ± 1.5	20.8 ± 3.4	0.24 ± 0.09
L.fallai	26.1 ± 2.9	6.0 ± 1.4	20.2 ± 2.9	0.31 ± 0.09
L.maccanni	26.4 ± 3.8	4.1 ± 1.5	22.4 ± 3.0	0.18 ± 0.07
L.n.poly-Tw	26.7 ± 6.1	5.4 ± 1.9	21.3 ± 4.9	0.26 ± 0.08
L.infrapunc	28.3 ± 3.3	7.7 ± 1.2	20.6 ± 2.9	0.38 ± 0.07
L.otagense	28.1 ± 3.3	4.4 ± 1.3	23.6 ± 3.1	0.19 ± 0.06
L.moco	28.6 ± 2.4	5.1 ± 1.3	23.6 ± 2.5	0.22 ± 0.07

Table 4.8. Pairwise comparisons between *Cyclodina* and the New Zealand *Leiolopisma* (observed differences). The complete 387 column dataset (including the 'extra' column; see Table 4.2 [part]) was used, but these numbers do not include differences resulting from insertions/deletions. *Cyclodina* taxa are numbered 1-8 while the New Zealand *Leiolopisma* are numbered 9-25.

a. Number of total differences.

		1	2	3	4	5	6	7	8
9	*L.stenotis*	21	18	20	18	24	18	17	19
10	*L.grande*	24	20	19	21	21	19	18	20
11	*L.notosaurus*	23	21	21	23	25	21	22	22
12	*L.lin\chl*	24	20	20	20	24	18	19	21
13	*L.suteri*	24	22	21	24	24	17	20	22
14	*L.n.nigri*	22	19	24	22	27	20	21	23
15	*L.microlepis*	25	21	22	22	25	22	23	19
16	*L.acrinasum*	23	23	24	23	27	20	21	25
17	*L.inconspic*	24	26	25	25	28	24	27	25
18	*L.smithi*	26	22	23	23	26	23	24	20
19	*L.zelandicum*	22	15	24	21	28	22	23	22
20	*L.fallai*	19	22	16	15	20	14	17	17
21	*L.maccanni*	27	28	24	24	26	25	26	24
22	*L.n.poly-Tw*	29	27	27	29	30	23	26	28
23	*L.infrapunc*	28	26	27	24	29	24	24	28
24	*L.otagense*	25	28	27	30	30	28	29	29
25	*L.moco*	21	20	22	21	25	20	21	19

b. Number of transversions (Tv).

		9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	*C.aenea*(So)	4	3	4	5	4	4	3	5	5	3	3	2	4	6	7	3	1
2	*C.aenea*(PK)	5	4	5	6	5	5	2	5	6	2	2	5	5	5	8	4	4
3	*C.whitakeri*	6	5	6	7	4	6	5	7	7	5	5	4	6	6	9	5	3
4	*C.alani*	8	7	8	9	8	8	7	9	9	7	7	6	8	10	11	7	5
5	*C.macgregori*	6	5	4	5	6	6	5	7	5	5	5	4	4	6	9	5	3
6	*C.oliveri*(Mo)	6	5	6	5	2	4	5	5	7	5	5	4	6	4	7	5	3
7	*C.oliveri*(GM)	5	4	5	6	3	5	4	6	6	4	4	3	5	5	8	4	2
8	*C.ornata*	4	3	4	5	4	4	3	5	5	3	2	2	4	6	7	3	1

c. Number of transitions (Ts).

		1	2	3	4	5	6	7	8
9	*L.stenotis*	17	13	14	10	18	12	12	15
10	*L.grande*	21	16	14	14	16	14	14	17
11	*L.notosaurus*	19	16	15	15	21	15	17	18
12	*L.lin\chl*	19	14	13	11	19	13	13	16
13	*L.suteri*	20	17	17	16	18	15	17	18
14	*L.n.nigri*	18	14	18	14	21	16	16	19
15	*L.microlepis*	22	19	17	15	20	17	19	16
16	*L.acrinasum*	18	18	17	14	20	15	15	20
17	*L.inconspic*	19	20	18	16	23	17	21	20
18	*L.smithi*	23	20	18	16	21	18	20	17
19	*L.zelandicum*	19	13	19	14	23	17	19	20
20	*L.fallai*	17	17	12	9	16	10	14	15
21	*L.maccanni*	23	23	18	16	22	19	21	20
22	*L.n.poly-Tw*	23	22	21	19	24	19	21	22
23	*L.infrapunc*	21	18	18	13	20	17	16	21
24	*L.otagense*	22	24	22	23	25	23	25	26
25	*L.moco*	20	16	19	16	22	17	19	18

Table 4.9. Pairwise comparisons between the New Zealand skinks and three overseas skinks (observed differences). The complete 388 column dataset (including the 'extra' column; see Table 4.2) was used, but these numbers do not include differences resulting from insertions/ deletions. *Cyclodina* taxa are numbered 1-8 and the New Zealand *Leiolopisma* are numbered 9-25 (as in Table 4.8).

a. Number of total differences.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
26 *L.telfairi*	21	22	27	26	26	27	26	22	23	24	24	25	25	25
27 *L.guichenoti*	23	26	22	20	26	18	19	23	16	22	25	18	22	21
28 *T.rohssii*	39	44	44	42	45	42	41	46	39	41	35	37	44	37

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
26 *L.telfairi*	25	28	27	28	24	32	29	30	33	28	30	-		
27 *L.guichenoti*	23	14	30	25	27	24	28	29	12	28	27	31	-	
28 *T.rohssii*	42	43	36	43	39	38	40	40	32	42	47	44	32	-

b. Number of transversions (Tv).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
26 *L.telfairi*	7	7	9	11	7	9	8	7	7	6	5	4	7	7
27 *L.guichenoti*	5	6	7	9	7	5	6	5	3	6	7	6	5	5
28 *T.rohssii*	11	12	13	15	11	13	12	11	13	12	11	12	13	13

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
26 *L.telfairi*	7	7	6	7	7	9	3	6	10	6	6	-		
27 *L.guichenoti*	6	4	8	6	6	7	7	7	4	4	6	10	-	
28 *T.rohssii*	12	14	12	12	12	11	11	13	14	12	12	14	12	-

c. Number of transitions (Ts).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
26 *L.telfairi*	14	15	18	15	19	18	18	15	16	18	19	21	18	18
27 *L.guichenoti*	18	20	15	11	19	13	13	18	13	16	18	12	17	16
28 *T.rohssii*	28	32	31	27	34	29	29	35	26	29	24	25	31	24

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
26 *L.telfairi*	18	21	21	21	17	23	26	24	23	22	24	-		
27 *L.guichenoti*	17	10	22	19	21	17	21	22	8	24	21	21	-	
28 *T.rohssii*	30	29	24	31	27	27	29	27	18	30	35	30	20	-

The average Tv /Ts ratios for the *Cyclodina* taxa range from 0.14 ± 0.13 to 0.64 ± 0.27 (Table 4.6c) - well below the expected *Cyclodina* Tv/Ts ratio at saturation (2.10). For the New Zealand *Leiolopisma*, $R_\infty = 2.09$ (mean base frequencies for the New Zealand *Leiolopisma* are given in Table 4.3b). This is almost identical to the value of R_∞ for *Cyclodina*. The observed Tv /Ts ratios for the New Zealand *Leiolopisma* taxa range from 0.18 ± 0.07 to 0.38 ± 0.07 (Table 4.7c) - again well below the expected Tv/Ts ratio at saturation (R_∞ *Leiolopisma* = 2.09).

In addition to applying Holmquist (1983)'s equation to his sequences, Hickson (1993) also examined them with regards to whether the Tv/Ts ratio increased as the number of transversions increased (implying a constant number of transitions and thus saturation) and found that it did. However, as the variation associated with the Tv/Ts ratio decreased as this occurred, and as this trend also occurs in other datasets (bovid, ratite and simulated data - see summary in Hickson, 1993), he concluded that this increase in Tv/Ts ratio resulted from greater variability in the number of transitions when only a few transversions occur, rather than from saturation. Due to the small sample size of the *Cyclodina* dataset, a similar evaluation is not feasible; however, it is assumed that as the *Cyclodina* sequences are so similar to those of Hickson (1993), the same conclusion is valid. Consequently, the New Zealand skink 12S rRNA sequences are assumed to be unsaturated for transitions.

Estimating Divergence Times for the Cyclodina Taxa.

Rates of evolutionary change often vary among divergent taxonomic groups and are also lineage-dependant for specific sequences (Britten, 1986; Gillespie, 1986; Vawter and Brown, 1986; Goddard *et al.*, 1990; Palmer, 1990; Satta and Takahata, 1990). Thus a reptile 12S rRNA rate is required to accurately determine the ages of the various *Cyclodina* lineages from the sequence data obtained in this study. However, as yet no precise rate has been established for the reptiles. Recent work indicates that the poikilothermic vertebrates (fish, reptiles and amphibians) have slower rates of mitochondrial DNA evolution than both birds and mammals (Larson and Wilson, 1989; Thomas and Beckenbach, 1989; Wallis and Arntzen, 1989; Avise *et al.*, 1992; Martin *et al.*, 1992; Martin and Palumbi, 1993). In turn, rates of bird sequence evolution appear to be slower than those of mammals (Kessler and Avise, 1985; Hickson, 1993; but see also Shields and Wilson, 1987).

Hickson (1993) reviewed other vertebrate datasets for which both 12S rRNA sequences and estimates of divergence time are available and presented six 12S rRNA rate estimates. Three of these were for warm-blooded vertebrates (bovids and the great apes), and thus are unlikely to be accurate in estimating times of divergence among the skinks. Two were for reptiles (xantusiid lizards), but these may not be reliable (Hickson, 1993; see also Hedges *et al.*, 1991). Therefore

the sixth rate estimate of 0.5 change per million years - derived from comparisons between the ostrich and *Rhea americana* - was chosen as the best available for estimating *Cyclodina* divergence times (see also Steel *et al.*, 1996 and Penny, Hasegawa and Cooper, in prep.).

The above rate was applied to the total numbers of differences given in Tables 4.6, 4.7, 4.8 and 4.9 and Table 4.10 shows the estimated divergence times for the closest and most distant pairs among the *Cyclodina* taxa (the intra-*Cyclodina* divergence range), the New Zealand *Leiolopisma* (the intra-*Leiolopisma* divergence range), between the two genera (the intergeneric range) and between the New Zealand skinks and three possible outgroups. Estimates of divergence time are also given for the two pairs which involve the proposed *Cyclodina* cryptic species (*C. aenea* [PK]-*C. aenea* [So] and *C. oliveri* [Mo]-*C. oliveri* [GM]) and for several other pairs of interest (such as *L. fallai* and *C. alani*; see later this chapter).

Pairwise Comparisons between the Skink Taxa - Conclusions.

The estimated *Cyclodina* divergence times indicate that the *Cyclodina* taxa, like the New Zealand *Leiolopisma* (see Hickson, 1993), are likely to be much older than previously suggested (see Chapter One). Furthermore, although these divergence times were calculated using a non-reptilian rate estimate, if anything it would seem likely for two reasons that they are underestimates. Firstly, the figures given in Tables 4.6-9 do not include differences resulting from insertions/deletions, which, if taken into account, would increase the divergence time estimates for some pairs of taxa, and secondly, if the skink mitochondrial DNA is evolving at a slower rate than that of both birds and mammals, then the ratite rate will be an overestimate.

These older divergence times provide some support for the hypothesis that *Cyclodina* is of Gondwanan origin, with one or more lineages from Gondwana, and subsequent speciation in New Zealand. Nevertheless, a reptile 12S rRNA rate is required to confirm these estimates, necessitating DNA sequence investigation of reptile groups with a more substantial fossil record such as crocodilians and iguanids (see Carroll, 1969; Estes and Price, 1973; Pregill, 1989).

Another interesting feature of the divergence time estimates presented in Table 4.10 is the fact that the closest *L. telfairi* pair involves *C. aenea* (So) and that the estimated divergence time for this pair is similar to that proposed for a divergence between *C. aenea* (So) and *C. aenea* (PK). However, phylogenetic analysis of the sequence data and an understanding of the relationships of these taxa is required before these findings, among others, can be accurately interpreted. Comparing the observed transversion/transition ratios to the transversion/transition ratio expected at saturation indicates that saturation of transitions is unlikely in either the *Cyclodina* or the *Leiolopisma* sequences (see also Hickson, 1993). Phylogenetic analysis is therefore not expected to be affected by this problem.

Table 4.10. Estimates of divergence times among *Cyclodina*, among the New Zealand *Leiolopisma*, between the two genera, and between the New Zealand skinks and three possible outgroups. These estimates were obtained by dividing the appropriate total number of changes (see Tables 4.6, 4.7, 4.8 and 4.9) by the ostrich-*Rhea americana* rate estimate of 0.5 changes/million years (Hickson, 1993).

a. Estimated New Zealand *Cyclodina* divergence times.

	Estimated divergence time (million years ago)
Most distant pair (*C. macgregori-C. aenea* [PK])	48
Closest pair (*C. oliveri* [GM]-*C. oliveri* [Mo])	10
C. aenea (So)-*C. aenea* (PK)	42
Most distant non-*C. aenea* (PK) pair (*C. macgregori-C. aenea* [So] and *C. whitakeri-C. aenea* [So])	38
Most distant non-*C. aenea* (So) pair (*C. macgregori-C. oliveri* [Mo])	34

b. Estimated New Zealand *Leiolopisma* divergence times.

	Estimated divergence time (million years ago)
Most distant pair (*L. moco-L. smithi*)	68
Closest pair (*L. microlepis-L. smithi*)	10
Most distant pair not involving *L. moco*, *L. fallai* or *L. zelandicum* (*L. otagense-L. n. polychroma*)	66
Most distant pair not involving *L. moco*, *L. fallai*, *L. zelandicum*, *L. otagense* or *L. infrapunctatum* (*L. acrinasum-L. n. polychroma* and *L. maccanni-L. smithi*)	62

c. Estimated *Cyclodina*-New Zealand *Leiolopisma* divergence times.

	Estimated divergence time (million years ago)
Most distant pair (*L. otagense-C. alani*, *L. otagense-C. macgregori* and *L. n. polychroma-C. macgregori*)	60
Closest pair (*L. fallai-C. oliveri* [Mo])	28
Next closest pair (*L. fallai-C. alani* and *L. zelandicum-C. aenea* [PK])	30
L. moco-C. ornata	38
Closest pair not involving *L. moco*, *L. fallai*, *L. zelandicum*, *C. alani*, *C. aenea* (PK), *C. ornata* or *C. aenea* (So) (*L. suteri-C. oliveri* [Mo] and *L. stenotis-C. oliveri* [GM])	34

d. Estimated divergence times from three possible outgroups - *L. telfairi*, *La. guichenoti* and *T. rohssii*.

		Estimated divergence time (million years ago)
L. telfairi	Most distant pair (*L. telfairi-L. infrapunctatum*)	66
	Closest pair (*L. telfairi-C. aenea* [So])	42
La. guichenoti	Most distant pair (*La. guichenoti-L. inconspicuum*)	60
	Closest pair (*La. guichenoti-L. infrapunctatum*)	24
T. rohssii	Most distant pair (*T. rohssii-L. moco*)	94
	Closest pair (*T. rohssii-L. infrapunctatum*)	64

Phylogenetic Analysis.

As mentioned in Chapter One, aligned DNA sequences contain signals which can be used to infer evolutionary relationships. The process of phylogeny reconstruction has two aspects - any transformations to the data (such as conversion to distance information, correction for multiple changes) and tree selection. Methods for tree selection can be divided into two general classes - those which use an optimality criterion (that is, select a tree based on optimisation of an objective function) and those which do not (see Penny *et al.*, 1993b). There are two classes of search strategies for optimality criterion methods - exact and heuristic (see Penny *et al.*, 1993b). Exact methods consider all possible trees and so guarantee to find the best tree/s for the optimality criterion used; however, they run relatively slowly. Heuristic methods do not examine all possible trees and so cannot guarantee optimality (although they are expected to find a tree close to 'optimal'), but have the advantage of running more quickly. Branch and bound methods (see Hendy and Penny, 1982; Penny and Hendy, 1987), which are used in this chapter, combine aspects of the two - they guarantee to find the optimal tree/s but also run quickly.

Another aspect of phylogeny reconstruction is the assumption of a model of evolution. Models of evolution have three parts (Penny *et al.*, 1992, 1993a and b):

1. A tree (the putative ancestral relationships of the taxa).
2. A mechanism for changes to the sequences (often the i.i.d or standard model - a simple mechanism which assumes changes occurring in the sequences are approximately independent and identically distributed).
3. Probabilities of change on edges of the tree (edge lengths).

The assumption of a model of evolution, with a reasonable mechanism for change, is essential if phylogenetic analysis methods are to give trees close to historical reality. For example, a method might join taxa based solely on the alphabetical order of their names - however, it is extremely unlikely that the resulting tree would be the correct (historical) one, produced as it was without reference to any realistic mechanism of evolution (Penny *et al.*, 1993b). Models of evolution and knowledge of mechanisms of evolution are also required in the investigation of the evolutionary histories of taxa - when they diverged and why and how (see Penny *et al.*, 1993b). A further application of models of evolution lies in using them to evaluate the reliability and accuracy of a selected phylogeny.

Phylogenetic analyses can be complicated or misled by the presence of different types of signals, from a variety of sources, in biological datasets. The signal of interest in phylogenetic reconstruction is the historical one - the patterns in the data arising from shared common ancestry and which indicate the correct historical tree. Conflicting or misleading signals can arise through convergence (resulting from multiple changes), natural selection (leading to convergence or apparent divergence), a complex mechanism of change (as mentioned above, most methods assume a simple mechanism), evolution in a non-tree-like fashion (most methods

assume sequences evolved in a binary or tree-like fashion), computer error (problems, limitations and/or inadequacies in either hardware or programs) and/or data collection errors (sampling, sequencing or alignment errors) (Penny *et al.*, 1993b). The use of spectral analysis (and the Hadamard conjugation; see Hendy and Penny, 1989, 1993; Penny *et al.*, 1992, 1993a; Lento *et al.*, 1995) allows the number of conflicting signals in the data and the degree of support for each to be directly assessed - thus indicating the degree of resolution of a selected phylogeny (and the level of confidence that can be placed in it) and possible alternative relationships.

Phylogenetic Analysis of Cyclodina.

The relationships of the *Cyclodina* taxa could not be clearly resolved with Hadamard analysis, undoubtedly due to the very small number of signals in the *Cyclodina* dataset. There are only 14 parsimony sites within the 382 column version of this dataset and 12 within the 378 column version (see below). This is a small number for any type of phylogenetic analysis. Low bootstrap values are expected. It therefore makes for an interesting study on how well the different tree selection methods handle datasets with small numbers of signals.

Three versions of the eight taxa *Cyclodina* dataset were analysed - the complete 387 column dataset shown in Tables 3.2 and 3.3, a smaller (382 column) dataset lacking columns containing gaps (columns 25, 26, 29, 61 and 369) and a further reduced (378 column) dataset without both gaps and regions within which gaps could not be accurately placed (columns 25-30, 60-61 and 369). Several optimality criteria were used - closest tree (CT; phylogeny program: Hadtree), maximum parsimony (MP; phylogeny program: PAUP*), minimum evolution (ME; PAUP*) and maximum likelihood (ML; PAUP*).

Spectral analyses (closest tree optimality criterion) were carried out on the 382 and 378 column datasets (the gap-containing 387 column dataset cannot be analysed using the Hadamard conjugation), while maximum parsimony, minimum evolution and maximum likelihood analyses were performed on both these and the 387 column dataset. Details of the methods and corrections used are given in Swofford *et al.* (1996). Spectral analyses were carried out using four colour information (4frq) - corrected using Kimura's three-parameter model (K3), two colour information (2frq; sum of seven option used) - corrected using Cavender's model (Cavend), and pairwise distances (dis) - Jukes-Cantor corrected (J-C) or Log Det corrected (correct for nucleotide frequencies F_ij option). Maximum parsimony analyses were carried out on uncorrected data, while maximum likelihood analyses were performed on data corrected using the Hasegawa-Kishino-Yano (HKY85) model (two substitution parameters, all sites assumed to evolve at the same rate) with a transition/transversion ratio of 2. Minimum evolution analyses (on pairwise distances) were carried out on uncorrected (uncorr.) data, on Jukes-Cantor corrected data and on data corrected using Kimura's two-parameter model (K2).

The search strategies used include branch and bound for the closest tree, maximum parsimony and minimum evolution analyses and an exhaustive search for maximum likelihood. In addition to finding the optimal trees, bootstrap consensus trees (1000 replicates) were also generated using maximum parsimony and minimum evolution (this was also done for maximum likelihood, but the results are not shown here due to problems with this option of the PAUP* test version). PHYLIP (Version 3.51c) analyses (uncorrected parsimony and neighbor-joining on Jukes-Cantor corrected data) were also run for the 378 and 387 column datasets.

The Hadamard spectra are represented in spectral diagrams or Lento-plots (Lento *et al.*, 1995) (see Appendix One for details on the construction of spectral diagrams). Figure 4.7 allows the unique mathematical labels (bipartition numbers) given to each grouping of taxa represented in these spectral diagrams to be translated back to their component taxa (see Penny *et al.*, 1993b, for an explanation of calculating bipartition numbers). The distance phylogenies are better resolved than those based on four or two frequency information and spectral diagrams for these are given in Figures 4.8 (Jukes-Cantor corrected distances from the 382 column dataset), 4.9 (Log Det corrected distances from the 382 column dataset), 4.10 (Jukes-Cantor corrected distances from the 378 column dataset) and 4.11 (Log Det corrected distances from the 378 column dataset). The less well resolved four frequency and two frequency spectral diagrams are given in Appendix Two.

Because of the number of PAUP* analyses which were carried out, only one tree (the best resolved based on bootstrap values) is shown for each of the three datasets (387, 382 and 378 column). In the case of the 387 and 378 column datasets, this was the minimum evolution tree from uncorrected distances; for the 382 column dataset, the minimum evolution tree (uncorrected distances) and the maximum parsimony tree were the same and equivalent on bootstrap values, and both sets of bootstrap values are given (see Figure 4.12).

A summary of the *Cyclodina* phylogenetic analyses - showing the bipartitions selected for inclusion in each optimal or bootstrap tree, along with the ranking of the bipartitions representing {*C. aenea* (So) and *C. aenea* (PK)} and {*C. oliveri* (GM) and *C. oliveri* (Mo)} (bipartitions 3 and 96 respectively) - is given in Table 4.11. Bipartitions which have 50% or greater bootstrap support or which are ranked among the first five in the Hadamard spectra are highlighted. The PHYLIP (Version 3.51c) analyses are not included in this table, as they gave the same trees as the equivalent PAUP* minimum evolution and maximum parsimony analyses. The use of different corrections in the minimum evolution analyses - uncorrected (as for maximum parsimony), one-parameter, or two-parameter (as for maximum likelihood) - did not appear to have any significant effect. Like the spectral analyses, the maximum parsimony, minimum evolution and maximum likelihood analyses did not produce a resolved *Cyclodina* phylogeny, and as expected, bootstrap values were low.

Figure 4.7 Bipartition numbers used to describe groupings of *Cyclodina* taxa in this study. This figure allows each bipartition number in the *Cyclodina* spectral diagrams (Figures 4.8-4.11 and Appendix Two Figures 2.1-2.2) and Table 4.11 to be translated back to its component taxa. Each bipartition 'splits' the taxa into two disjoint subsets (subsets which have no taxa in common). For example, bipartition 3 is {taxa 1+2} versus {taxa 3-8}, which in this case is {*C. aenea* (So) and *C. aenea* (PK)} versus {*C. whitakeri*, *C. alani*, *C. macgregori*, *C. oliveri* (Mo), *C. oliveri* (GM) and *C. ornata*}; bipartition 100 is {taxa 3+6+7} versus {taxa 1+2+4+5+8}, or {*C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM)} versus {*C. aenea* (So), *C. aenea* (PK), *C. alani*, *C. macgregori* and *C. ornata*}.

Taxa:

1 *C. aenea* (So) 2 *C. aenea* (PK) 3 *C. whitakeri* 4 *C. alani*
5 *C. macgregori* 6 *C. oliveri* (Mo) 7 *C. oliveri* (GM) 8 *C. ornata*

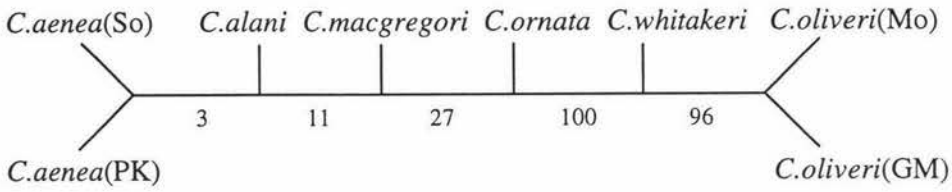
Taxa included in each tree bipartition:

	1 3 5 7		1 3 5 7		1 3 5 7		1 3 5 7		1 3 5 7
3	**.....	9	*..*....	11	**.*....	17	*...*...	25	*..**...
27	**.**...	68	..*...*.	96	.....**.	98	.*...**.	100	..*..**.
102	.**..**.	110	*...*..*	111	....*..*	116	..*.***.	119	...*...*
126	*.....*								

Taxa included in each non-tree bipartition (those bipartitions for which there is support in the spectral analyses but which are not selected as tree bipartitions - see spectral diagrams):

	1 3 5 7		1 3 5 7		1 3 5 7		1 3 5 7		1 3 5 7
10	.*.*....	12	..**....	13	*.**....	14	.***....	15	****....
19	**..*...	20	..*.*...	22	.**.*...	23	***.*...	24	...**...
28	..***...	30	.****...	35	**...*..	36	..*..*..	37	*.*..*..
40	...*.*..	41	*..*.*..	42	.*.*.*..	43	**.*.*..	44	..**.*..
45	*.**.*..	46	.***.*..	47	....*.**	49	*...**..	50	.*..**..
54	.**.**..	55	...*..**	56	...***..	58	.*.***..	59	..*...**
60	..****..	61	.*....**	65	*.....*.	66	.*....*.	70	.**...*.
71	***...*.	75	**.*..*.	76	..**..*.	78	.***..*.	79	....**.*
80	....*.*.	81	*...*.*.	82	.*..*.*.	83	**..*.*.	84	..*.*.*.
85	*.*.*.*.	87	...*.*.*	89	*..**.*.	90	.*.**.*.	91	..*..*.*
92	..***.*.	93	.*...*.*	95	.....*.*	99	**...**.	106	.*.*.**.
107	..*.*..*	108	..**.**.	109	.*..*..*	113	*...***.	114	.*..***.
117	.*.*...*	118	*..*...*	120	...****.	121	.**....*	123	..*....*
124	**.....*								

Figure 4.8 Spectral diagram for pairwise distance (Jukes-Cantor corrected) information from the 382 column *Cyclodina* dataset. The format for each of the *Cyclodina* Lento-plots is the same: the unrooted tree is shown diagrammatically at the top with the appropriate bipartition number next to each of the internal edges. The main (lower) figure shows bipartitions considered for inclusion in the optimal tree (a tree being a mutually compatible set of t-3 bipartitions, where t is the number of taxa). The support for each bipartition (after any correction for multiple changes) is shown above the x axis, the sum of the contradictions for each bipartition is below the x axis. The latter values are normalised so that they sum to the same value as the support. Bipartitions in black are those selected as being in the optimal tree by the tree selection criterion (closest tree), these correspond to the numbers shown in the unrooted tree at the top of the figure.



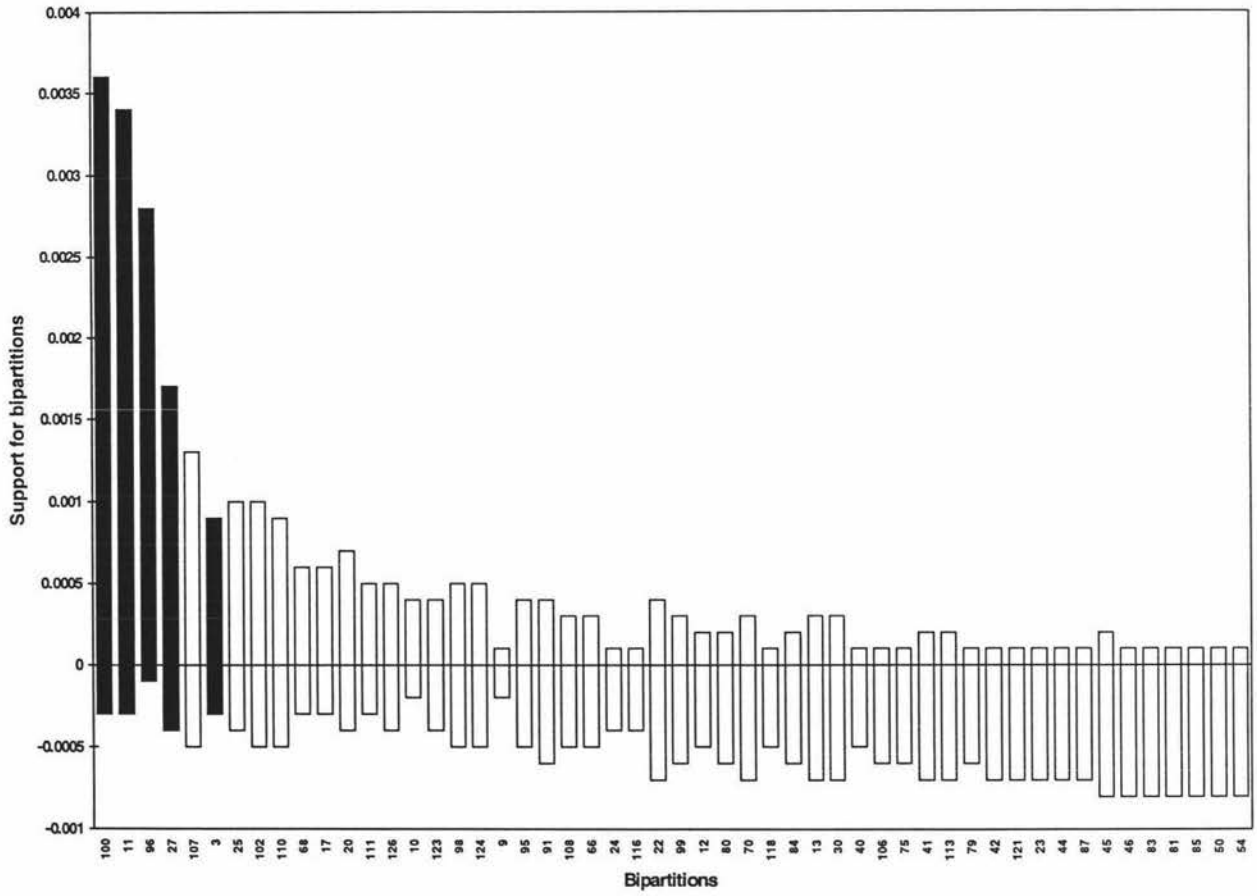
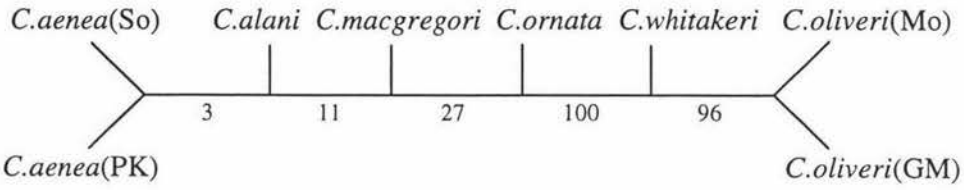


Figure 4.9 Spectral diagram for pairwise distance information (Log Det corrected using the default option = correct for nucleotide frequencies F_ij) from the 382 column *Cyclodina* dataset.



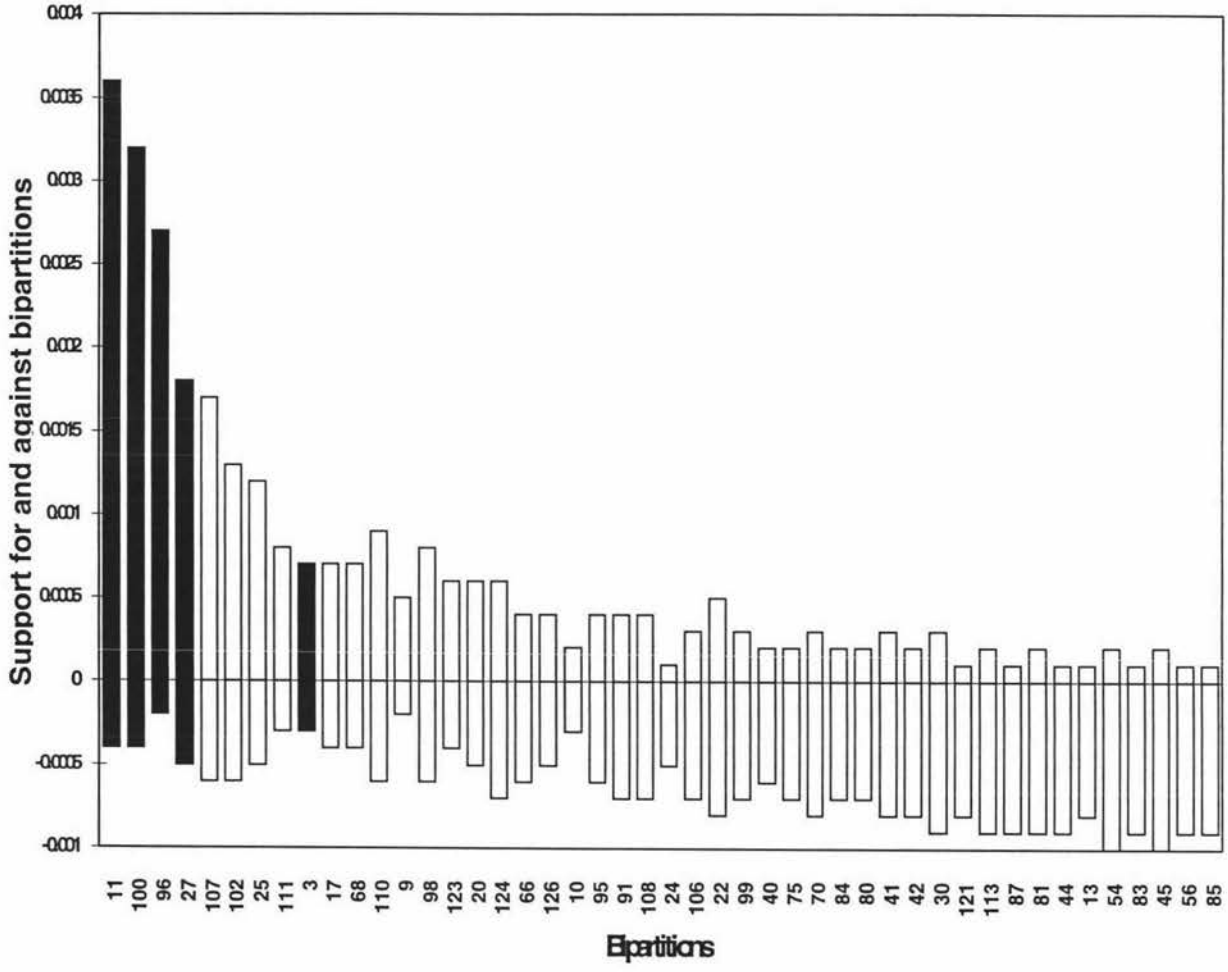
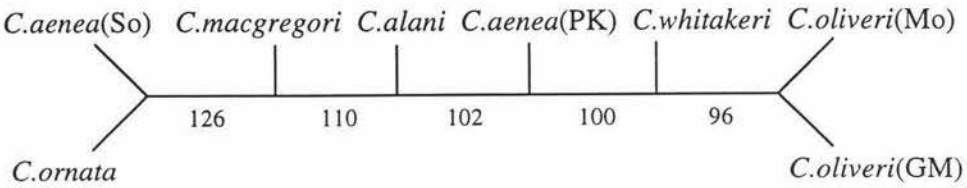


Figure 4.10 Spectral diagram for pairwise distance (Jukes-Cantor corrected) information from the 378 column *Cyclodina* dataset.



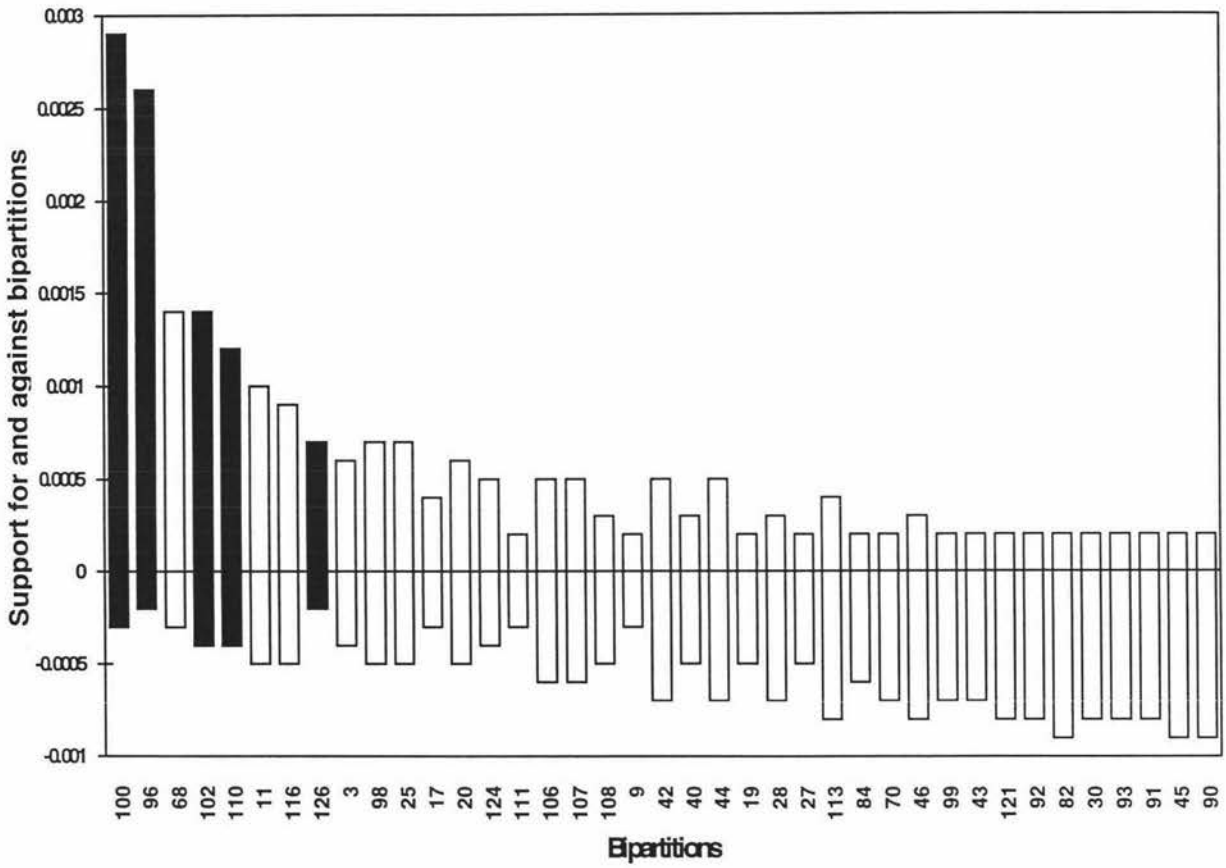
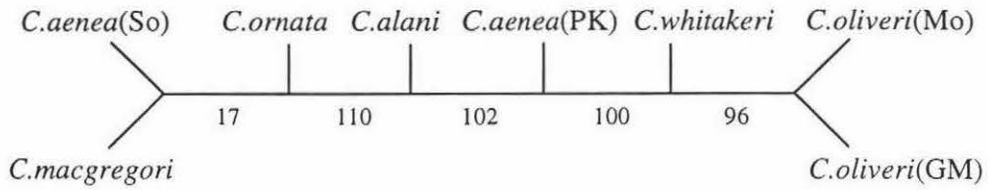


Figure 4.11 Spectral diagram for pairwise distance information (Log Det corrected using the default option = correct for nucleotide frequencies F_ij) from the 378 column *Cyclodina* dataset.



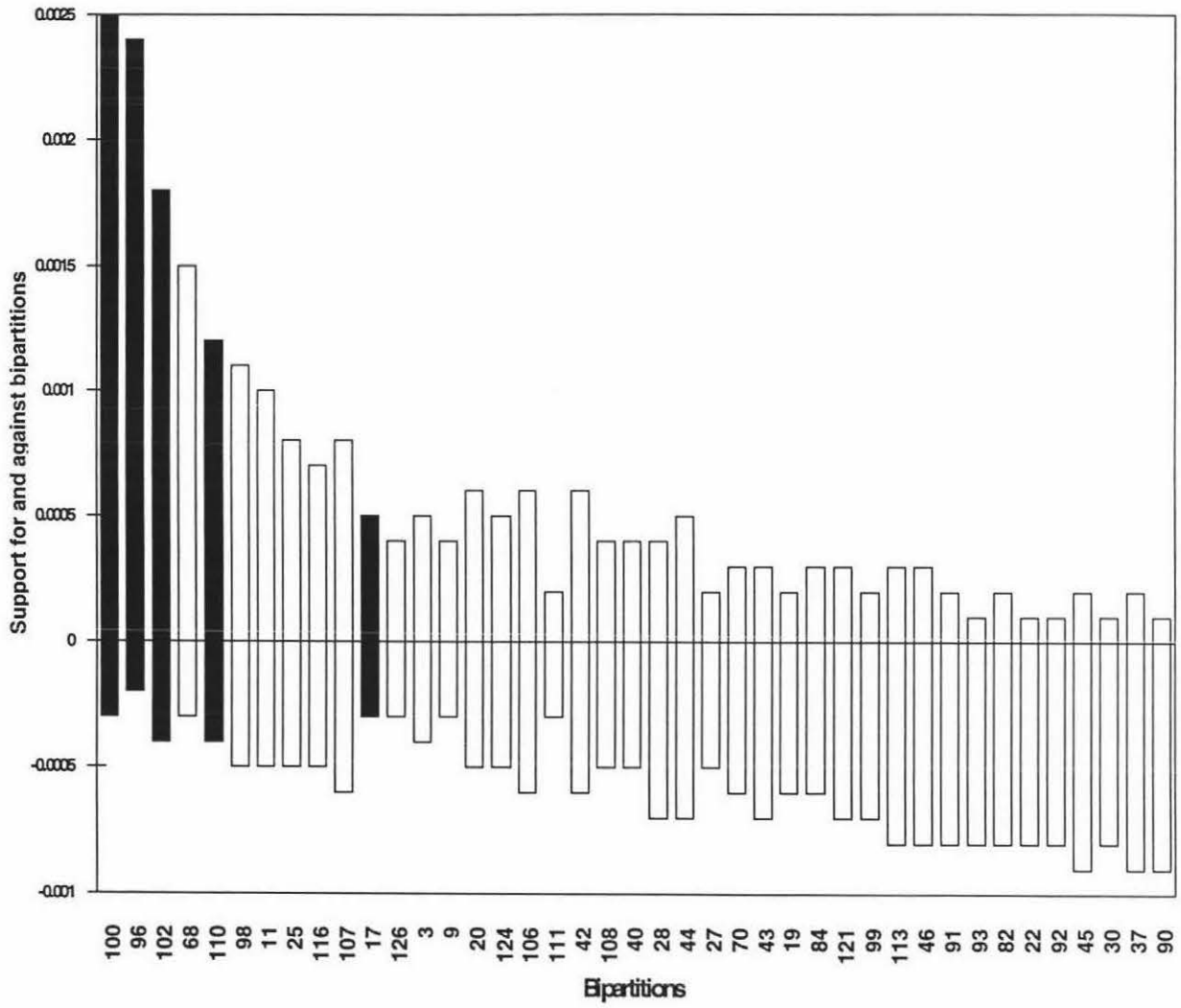
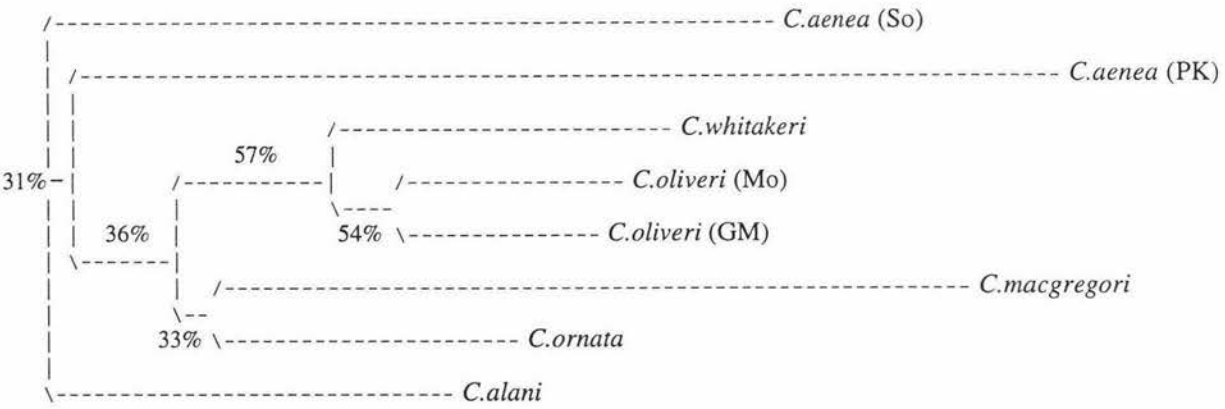
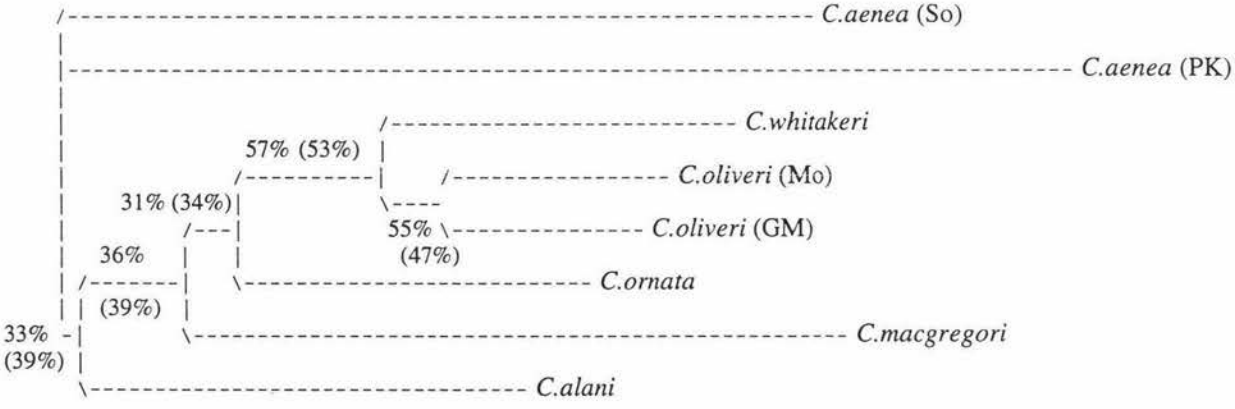


Figure 4.12 PAUP* phylogenies for the 387, 382 and 378 column versions of the eight taxa *Cyclodina* dataset. Each tree is a bootstrap 50% majority-rule consensus tree, with other groups compatible with the majority-rule groupings included. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The trees are unrooted. PAUP* Version 4.0.0d38 was used for these analyses.

a. Minimum evolution tree derived from uncorrected distances ("p") from the 387 column *Cyclodina* dataset. The minimum evolution score is 0.13911.



b. Minimum evolution tree derived from uncorrected distances ("p") and maximum parsimony tree (uncorrected) from the 382 column *Cyclodina* dataset. The minimum evolution bootstrap percentages are unbracketed, the maximum parsimony percentages are in brackets. The minimum evolution score is 0.13593; the parsimony tree length is 55.



c. Minimum evolution tree derived from uncorrected distances ("p") from the 378 column *Cyclodina* dataset. The minimum evolution score is 0.12497.

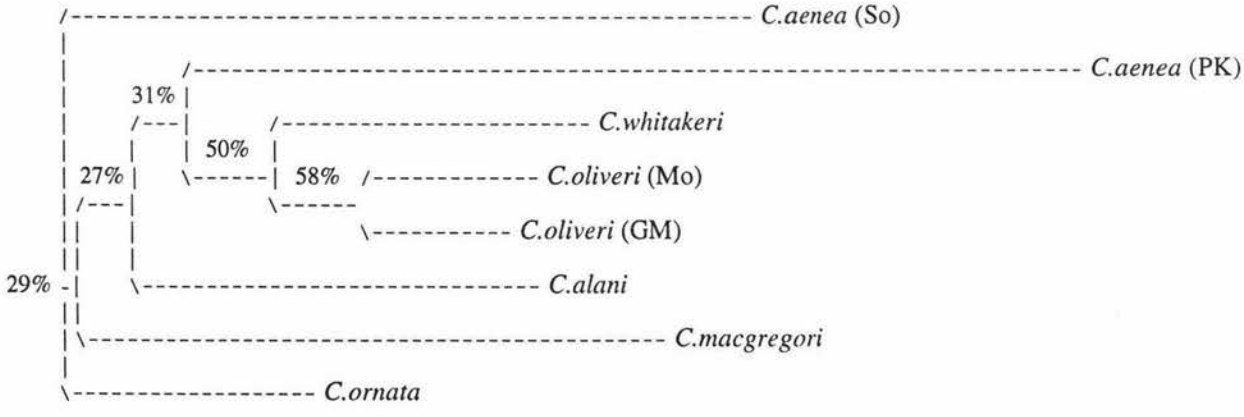
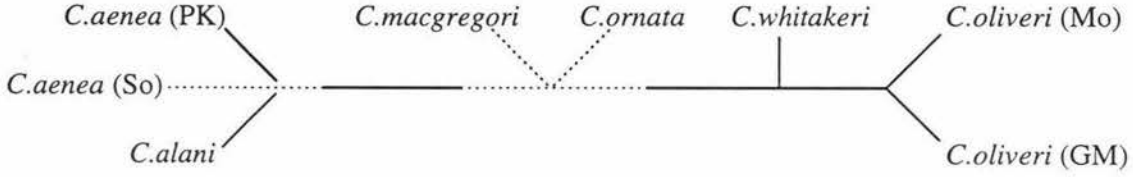


Table 4.11. Summary of the bipartitions selected in phylogenetic analyses of the eight *Cyclodina* taxa. An x marks each bipartition included in the optimal or bootstrap tree. Bipartitions which have 50% or greater bootstrap support in the bootstrap trees or which are in the first five bipartitions in the spectral diagrams are indicated with bold+underlined x's (**x**). The abbreviations used in detailing the analyses are defined in the text. The ranks of two specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]) and 96 (*C. oliveri* [Mo] and *C. oliveri* [GM]) - are also shown so that the support for these two bipartitions can be investigated. In the closest tree phylogenies, rank is supplied by the Hadtree program and is based on the strength of the signals and clashes (this ranking system includes the external edges). In the maximum likelihood, maximum parsimony and minimum evolution optimal trees, rank is based on the length of the internal edge supporting the bipartition (the bipartition is left unranked if not in the tree). In the bootstrap trees, rank is based on bootstrap value (for example, the bipartition with the highest bootstrap value is ranked as number one and so on). Where the optimal and bootstrap trees are the same, the optimal tree rank is given first and then the bootstrap rank (for example, -;6). There are 16384 possible bipartitions for the four colour datasets, a similar number for the maximum parsimony and maximum likelihood datasets and 128 for the two colour and distance datasets.

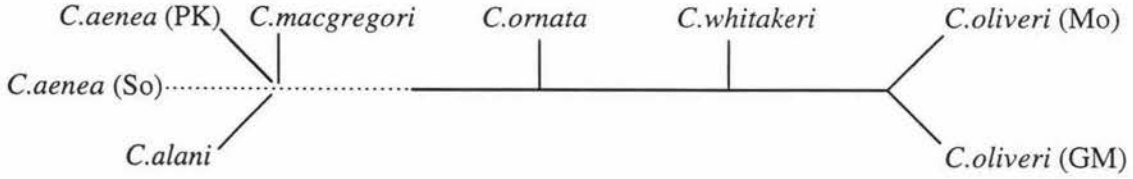
Data set	Optimality criterion	Transformation (conversion, correction)	Optimal or Bootstrap tree	Bipartitions															Rank of:	
				3	9	11	17	25	27	68	96	98	100	102	110	111	116	119	126	3	96
387	MP	uncorrected	optimal	x		x			x		x		x							2	2
			bootstrap		x	x			x		x		**x**							5	2
	ME	dis (uncorr.)	both		x	x					**x**		**x**			x				-;6	3;2
		dis (J-C)	both		x	x					**x**		**x**			x				-;6	3;2
		dis (K2)	both		x	x					**x**		**x**			x				-;6	3;2
	ML	(HKY85)	optimal	x		x			x		x		x							4	2
382	CT	4frq (K3)	optimal	**x**							**x**		x				x	x		6	8
		2frq (Cavend)	optimal				x	x	**x**		**x**		**x**							6	13
		dis (J-C)	optimal	x		**x**			**x**		**x**		**x**							16	10
		dis (Log Det)	optimal	x		**x**			**x**		**x**		**x**							19	10
	MP	uncorrected	both	x		x			x		x		**x**							2;3	2;2
	ME	dis (uncorr.)	both	x		x			x		**x**		**x**							5;4	3;2
			optimal #2		x	x			x		x		x							-	3
		dis (J-C)	optimal		x	x			x		x		x							-	3
			bootstrap	x		x			x		**x**		**x**							4	2
		dis (K2)	optimal		x	x			x		x		x							-	3
			bootstrap	x		x			x		**x**		**x**							4	2
	ML	(HKY85)	optimal	x		x			x		x		x							4	2
378	CT	4frq (K3)	optimal	**x**		x					**x**		x				x			6	10
		2frq (Cavend)	optimal				x	x			**x**		**x**	x						6	12
		dis (J-C)	optimal								**x**		**x**	**x**	**x**				x	19	9
		dis (Log Det)	optimal				x				**x**		**x**	**x**	**x**					21	9
	MP	uncorrected	both								**x**		x	x	x				x	-;8	2;1
			optimal #2							x			x	x	x				x	-	-
			optimal #3		x			x			x	x		x						-	2
			optimal #4		x			x		x			x	x						-	-
			optimal #5		x			x			x		x	x						-	2
	ME	dis (uncorr.)	both								**x**		**x**	x	x				x	-;6	2;1
		dis (J-C)	both								**x**		x	x	x				x	-;7	2;1
		dis (K2)	both								**x**		x	x	x				x	-;7	2;1
	ML	(HKY85)	optimal		x			x			x		x	x						-	1

Figure 4.13 Consensus trees for the 387, 382 and 378 column *Cyclodina* datasets. Each tree shows the underlying groupings of taxa common to all trees for that dataset. Edges which are in agreement in all of the trees for the dataset are shown as solid lines. Areas for which the exact relationships of the taxa are not clear (that is, for which there are conflicting edges amongst the set of trees) are indicated by dotted lines; the taxa groupings given for these areas result from collapsing the conflicting edges. For example, analysis of the 387 column dataset gave three trees (see Table 4.11) - two containing a {*C. aenea* (So) and *C. alani*} pair, with *C. aenea* (PK) the next most closely related taxon, and one containing a {*C. aenea* (So) and *C. aenea* (PK)} pair, with *C. alani* the next most closely related taxon. Collapsing these conflicting edges and thus removing an internal edge from the tree gives the same three-way grouping of taxa {*C. aenea* (So)/*C. alani*/*C. aenea* (PK)}, which summarises the patterns seen in all three trees (see **a.**). Similarly, there are two possible placements for *C. macgregori* and *C. ornata* among the 387 column dataset trees - paired with each other or separate (see Table 4.11). Collapsing these edges again removes an internal edge from the tree and gives a 'V-shaped' pattern summarising the relationships of these taxa (see **a.**). In addition, the external edges of taxa which fall within the dotted regions are also differentiated into solid or dotted lines, to indicate which of the taxa in these regions are observed to pair and which are not (pairing is not observed between taxa with solid external edges). For example, the pattern of external edges for *C. aenea* (So), *C. alani* and *C. aenea* (PK) in tree **a.** indicates that both {*C. aenea* (So) and *C. alani*} and {*C. aenea* (So) and *C. aenea* (PK)} can pair, but that a {*C. alani* and *C. aenea* (PK)} pair is not observed. The trees are unrooted and not drawn to scale.

a. Consensus tree for the 387 column *Cyclodina* dataset, summarising the results of maximum parsimony, minimum evolution and maximum likelihood analyses (see Table 4.11).



b. Consensus tree for the 382 column *Cyclodina* dataset, summarising the results of closest tree, maximum parsimony, minimum evolution and maximum likelihood analyses (see Table 4.11). Two of the edges in the 4frq tree (bipartitions 116 and 119) were not taken into consideration when constructing this tree due to the lack of support for these edges (see Appendix Figure 2.1).



c. Consensus tree for the 378 column *Cyclodina* dataset, summarising the results of closest tree, maximum parsimony, minimum evolution and maximum likelihood analyses (see Table 4.11). Two of the edges in the 4frq tree (bipartitions 11 and 116) were not taken into consideration when constructing this tree due to the lack of support for these edges (see Appendix Figure 2.3). The taxa are divided into two groups of four as shown here in all but one of the trees. The exception is the 4frq tree, in which *C. aenea* (PK) pairs with *C. aenea* (So) and forms part of the left-hand group (see Table 4.11 and Appendix Figure 2.3).

In addition, a consensus tree is given for each of the three datasets (see Figure 4.13). These were constructed by comparing the closest tree, maximum parsimony, minimum evolution and maximum likelihood trees for each dataset and collapsing edges which were in conflict between these trees. Edges which are in agreement in all of the trees are shown as solid lines while areas for which conflicting edges are observed are represented by dotted lines (see Figure 4.13).

Phylogenetic Analysis of *Cyclodina* - Conclusions.

Although none of these forms of analysis produce a fully resolved *Cyclodina* taxonomy, certain conclusions can still be drawn. The eight *Cyclodina* taxa form genetically distinct lineages, including the two proposed cryptic species - *C. aenea* (PK) and *C. oliveri* (Mo), and clearly merit separate taxonomic status. The exact classification status of *C. oliveri* (Mo) remains unclear. On the basis of both allozyme (Vos, 1988) and sequence data analysis, it is clearly distinct from *C. oliveri* (GM). However, as there is reasonable support for a close relationship between the two (see below and rankings of bipartition 96 in Table 4.11), further investigation is required to determine whether subspecies or separate species status is warranted. *C. aenea* (PK) definitely appears to be a separate species. Not only is it genetically quite distinct from *C. aenea* (So), but other species may lie between them (see below and rankings of bipartition 3 in Table 4.11).

Two of the edges in the *Cyclodina* trees are fairly well resolved - {*C. oliveri* (Mo) and *C. oliveri* (GM)} (bipartition 96) and {*C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM)} (bipartition 100). These bipartitions occur in all but two and all but one respectively of the trees summarised in Table 4.11, have 50% or greater bootstrap support in most of the bootstrap trees and are among the five highest ranked bipartitions in all or most of the Hadamard trees (see also Figure 4.13). Even so, the level of support is not particularly strong - the level of bootstrap support for these edges does not go over 60%, and because of the small number of signals in the *Cyclodina* dataset generally, even being among the first five bipartitions does not represent a high level of support.

The remaining relationships are unclear. In phylogenies from the 387 and 382 column datasets, *C. aenea* (So), *C. alani* and *C. aenea* (PK) appear to be closely related, but the exact relationships cannot be resolved, while *C. macgregori* appears to be closely related both to this group and to *C. ornata* (see Figure 4.13). The 378 column dataset phylogenies provide even less resolution - with *C. aenea* (So), *C. alani*, *C. macgregori* and *C. ornata* forming one group and *C. aenea* (PK), *C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM) forming a second (see Figure 4.13), although *C. aenea* (PK) does fall inside the first group in the 4frq tree, pairing with *C. aenea* (So). Certain pairings do not occur within the first group (for example, *C.*

macgregori and *C. alani*), but the exact relationships of these taxa cannot be determined. In addition, the generally well-supported {*C. oliveri* (Mo) and *C. oliveri* (GM)}and {*C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM)} groupings are disrupted in three of the trees (see Table 4.11 and Figure 4.13). The larger number of trees for the 378 column dataset (eight in total compared to a total of three and four respectively for the 387 and 382 column datasets; see Table 4.11) and the lower level of agreement between these trees may indicate that the 387 and 382 column dataset phylogenies are more reliable.

Examination of the edge lengths of these trees (actual data not shown, but see Figure 4.12 in which the edge lengths of the trees are proportional to the number of changes along each edge) indicates that internal edge lengths are small compared to external edge lengths. The relative shortness of the internal edges helps to explain why the *Cyclodina* taxonomy is so hard to resolve, while the comparative length of the external edges indicates that the *Cyclodina* lineages are relatively old. The same patterns of external edge lengths are seen in all of the trees - *C. aenea* (PK) has the longest external edge length, followed by *C. aenea* (So) and *C. macgregori*, *C. alani*, *C. whitakeri* and *C. ornata*, *C. oliveri* (Mo) and finally *C. oliveri* (GM). This indicates that the two *C. aenea* taxa and *C. macgregori* are the oldest of the *Cyclodina* lineages, and further supports the separate taxonomic status of *C. aenea* (PK).

Despite the lack of phylogenetic resolution (short internal edges and inability to determine the branching order of the taxa) from these analyses, there is phylogenetic information in the 12S rRNA dataset. For eight taxa, there are 10395 possible unrooted trees. However, only 13 trees were observed in total from sequence data analyses of the eight *Cyclodina* taxa (see Table 4.11).

Allozyme analyses (Vos, 1988) of the *Cyclodina* taxa plus three outgroups gave two trees (see Chapter One). When outgroups are excluded, the two Vos (1988) trees are identical. However this tree is not the same as the trees presented here from 12S rRNA sequence data. *C. ornata* moves within the {*C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM)} grouping in Vos (1988)'s tree and the 12S rRNA data does not support a {*C. macgregori* and *C. alani*} pair. Despite these differences, the two datasets both suggest that *C. aenea* (So), *C. aenea* (PK), *C. alani* and *C. macgregori* are the oldest *Cyclodina* lineages. The placement of the outgroups in Vos (1988)'s trees also suggests that *Cyclodina* is monophyletic with regards to *Leiolopisma*. To examine this, and to investigate the possible origin/s of the *Cyclodina* taxa, phylogenetic analyses of the *Cyclodina* sequences in conjunction with the sequences of Hickson (1993) were also carried out.

Phylogenetic Analysis of the New Zealand Skinks, Leiolopisma telfairi, Lampropholis guichenoti and Tropidoscincus rohssii.

Several versions of the skink dataset presented in Table 4.2 were analysed. The New Zealand skinks (excluding the *C. aenea* [R] sequence) were examined alone (25 taxa), in conjunction with *L. telfairi* and *La. guichenoti* (27 taxa) and with *L. telfairi*, *La. guichenoti* and *T. rohssii* (28 taxa). In addition, two versions of each of these datasets were analysed - one containing all sites (except for the 'extra' site - column 359 in Table 4.2) and one lacking both indel sites and regions in which these gaps could not be accurately placed (as well as the 'extra' column). The two 25 taxa datasets contained 386 and 374 columns respectively (the 386 column dataset also lacks the *L. telfairi* insertion site - column 204 in Table 4.2 - while columns 23-30, 59-61, 204, 359 and 369 were omitted in the 374 column dataset), the two 27 taxa datasets had 387 and 368 columns respectively (columns 23-30, 59-61, 202-208, 359 and 369 were excluded in the 368 column dataset) while the two 28 taxa datasets had 387 and 298 columns respectively (columns 1-81, 202-208, 359 and 369 were absent in the 298 column dataset).

Three optimality criteria were used - closest tree (CT; phylogeny program: Hadtree), maximum parsimony (MP; phylogeny program: PAUP*) and minimum evolution (ME; PAUP*). Maximum likelihood analyses were not carried out, as the large numbers of taxa involved and the low level of resolution for the skink dataset (see later and Hickson, 1993) made this approach too time-consuming. Again, details of the methods and corrections used are given in Swofford *et al.* (1996). Spectral analysis (closest tree optimality criterion) was applied to the 374, 368 and 298 column datasets, using pairwise distances (dis) - both Jukes-Cantor corrected (J-C) and Log Det corrected (correct for nucleotide frequencies F_ij option) - and a branch and bound search strategy. Four frequency information cannot be analysed for datasets of more than 12 taxa, while the use of two frequency information was also too time-consuming. As spectral analysis can only be applied to datasets of 20 taxa or less at present, several overlapping subsets of taxa were analysed for each dataset using this approach - thus giving a set of subtrees for each of the three datasets.

Maximum parsimony and minimum evolution bootstrap analyses (1000 replicates) were carried out for both the 'reduced' (374, 368 and 298 column) and 'complete' (386 and 387 column) datasets. Maximum parsimony was applied to uncorrected (uncorr.) data using a heuristic search strategy ('fast stepwise addition' option), while minimum evolution analyses were carried out using pairwise distances (Jukes-Cantor corrected) and a neighbor-joining search. Exact searches could not be used due to time constraints, again because of the large numbers of taxa and low level of resolution involved; similarly, a bootstrap approach only was used because of the very large number of optimal trees found under these conditions. A smaller number of parsimony (uncorrected data) and neighbor-joining (Jukes-Cantor corrected data) analyses were run using PHYLIP Version 3.51c and 100 bootstrap replicates. The resulting trees were almost identical to those obtained from the PAUP* analyses, and so are not presented here.

Two phylogenies are given for each of the 'reduced' (374, 368 and 298 column) datasets - a subtree agreement tree showing the results of all spectral analyses for that dataset (see later) and the better resolved of the two PAUP* trees, based on bootstrap values (in each case, this was the minimum evolution tree). In addition, a table summarising which bipartitions were selected in each of the different analyses is presented for each of the three datasets.

Spectral diagrams were not constructed, due to the large number of spectral analyses which were carried out. Instead, the order in which the bipartitions would appear in a spectral diagram is shown in the summary tables. The bipartition which would come first in a spectral diagram (that is, which has the largest [signal-clash] value) is listed with the rank of '1', and so on. Non-tree bipartitions are included in this ranking - thus some ranks are not shown in the tables. The edgelengths of the tree bipartitions are also given in the tables, providing an indication of the strength of the signal supporting each bipartition. Similarly, ranks (based on bootstrap value) and edgelengths are supplied for the PAUP* trees (again, non-tree bipartitions are included in the ranking), and bipartitions which have 50% or greater bootstrap support are highlighted. The ranks of three specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]), 96 (*C. oliveri* [GM] and *C. oliveri* [Mo]) and 255 (all eight *Cyclodina* taxa) - are also summarised in these tables, to allow the level of support for these taxa groupings to be examined. The ranks of these bipartitions were supplied by the Hadtree programme for the closest tree analyses and by bootstrap value (as above) for the PAUP* analyses.

The better resolved of the two PAUP* trees (based on bootstrap values) is also presented for each of the 'complete' (386 and 387 column) datasets (again, the minimum evolution tree was the better resolved in each case), along with a table summarising the results of all analyses for these datasets. As above, the rank (based on bootstrap value) and edgelength of each of the tree bipartitions are included in the table, as are the ranks of bipartitions 3, 96 and 255 (also determined on bootstrap value). Again, bipartitions which have 50% or greater bootstrap support are highlighted.

The taxa groupings described by the bipartition numbers in the summary tables are shown in Figure 4.14. Information from the 374 column dataset is presented in Figure 4.15 (subtree agreement tree), Figure 4.16 (minimum evolution tree) and Table 4.12 (summary of bipartitions selected in phylogenetic analysis of this dataset). The 368 column dataset information is given in Figure 4.17 (subtree agreement tree), Figure 4.18 (minimum evolution tree) and Table 4.13 (summary of bipartitions). The 298 column dataset information is presented in Figure 4.19 (subtree agreement tree), Figure 4.20 (PAUP* tree) and Table 4.14 (summary of bipartitions). The minimum evolution trees for the 386 and 387 column datasets are given in Figures 4.21 (386 column), 4.22 (387 column, 27 taxa) and 4.23 (387 column, 28 taxa), while Table 4.15 summarises the bipartitions selected in phylogenetic analyses of these datasets.

Subtree Agreement Trees.

The subtree agreement trees presented for these datasets (Figures 4.15, 4.17 and 4.19) were constructed using a new method for displaying the results of several phylogenetic analyses in the form of a single tree. Here it is used to summarise the results of spectral analysis of several different but overlapping subsets of taxa from a particular dataset. However, it can also be used to show the results of different types of analysis (such as maximum parsimony, minimum evolution and maximum likelihood for example) on the same dataset or of the same type of analysis on several different datasets - providing the resulting trees are not exceedingly dissimilar. Each tree is constructed from the edges for which there are no conflicts among the set of trees being compared, along with selected edges from the set of conflicting edges. The taxa groupings denoted by the remaining conflicting edges are then indicated with arrows (which can be given different colours, patterns and/or numbers to distinguish between the different trees being compared).

Selection of the conflicting edges to be included in the tree can be on a majority rules basis - for example, if three trees were compared and two of them had the grouping (taxa 1+2) while the remaining tree had the grouping (taxa 1+3), then the (1+2) edge would be included in the tree, with the (1+3) grouping indicated by an arrow. Alternately, it can be based on the type of analysis or the dataset used (for example, all edges from a maximum parsimony tree might be included in the tree, with any conflicting minimum evolution edges indicated by arrows).

Each of the subtree agreement trees presented here was constructed by comparing the set of subtrees resulting from spectral analyses of Log Det corrected distances. Selection of the conflicting edges included in the tree was made on the basis of the patterns of taxa presence/absence among the different subtrees and the corresponding occurrence or non-occurrence of particular conflicting edges - in order to give a tree as close to possible (based on these comparisons) to that which would be produced if all of the taxa could be analysed together. To give a hypothetical example - if taxa A and B paired in three out of five subtrees but were separate in the remaining two subtrees, and an examination of the taxa included in each of the subtrees revealed that the only unique pattern of taxa absence/presence between the two groups of subtrees was the absence of taxa C, D and E from the first three trees and the presence of these taxa in the remaining two trees, then taxa A and B would be shown as separate in the subtree agreement tree (the assumption being that with taxa C, D and E present, as they would be if all of the taxa could be analysed together, taxa A and B would not pair). The rest of the conflicting edges, along with any conflicting edges from the subtrees derived from Jukes-Cantor corrected distances, are indicated with arrows.

Figure 4.14 Bipartition numbers used to describe groupings of taxa in trees from phylogenetic analyses of the New Zealand skinks alone (25 taxa), in conjunction with *L. telfairi* and *Lampropholis guichenoti* (27 taxa) and with *L. telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (28 taxa). This figure allows each bipartition number in Tables 4.12 - 4.15 to be translated back to its component taxa. As these bipartition numbers are used to describe groupings of taxa in both subset trees and complete trees (the PAUP* phylogenies), the normal pattern of disjoint subsets cannot be shown. For example, for a dataset containing n taxa, bipartition 3 normally represents both the grouping (taxa 1+2) and the complementary grouping (taxa 3+4+.....+n) and is shown as:

```
123456789...n
**.........
```

However, the complementary grouping will be different in a dataset lacking, for example, taxa 4, 5 and 6, and different again in a dataset lacking, for example, taxa 6, 7 and 8. Therefore only one taxa grouping is given for each bipartition number. Which of the taxa groupings is given depends on how often each grouping occurs among the set of trees obtained in this study. For example, the taxa grouping (taxa 1+2) occurs more often than any of its complementary groupings, so this is the grouping which is given - in the form: bipartition 3

```
123456789...n.
**
```

If the groupings occur an equal number of times, then the smaller of the groupings (in terms of bipartition number - see Penny *et al.*, 1993b, for an explanation of calculating bipartition numbers) is given.

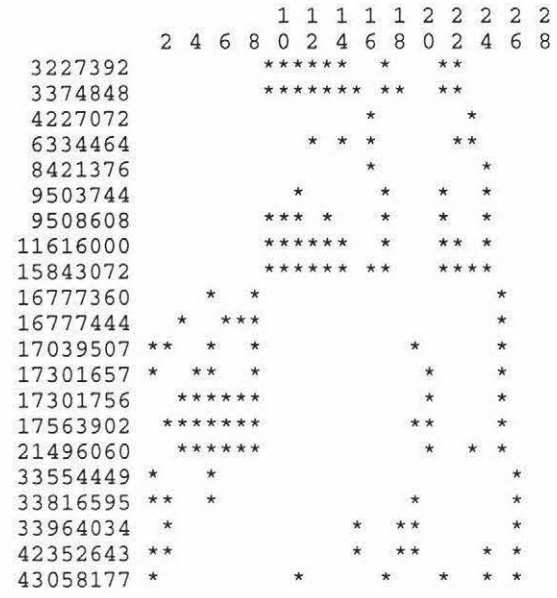
Taxa:

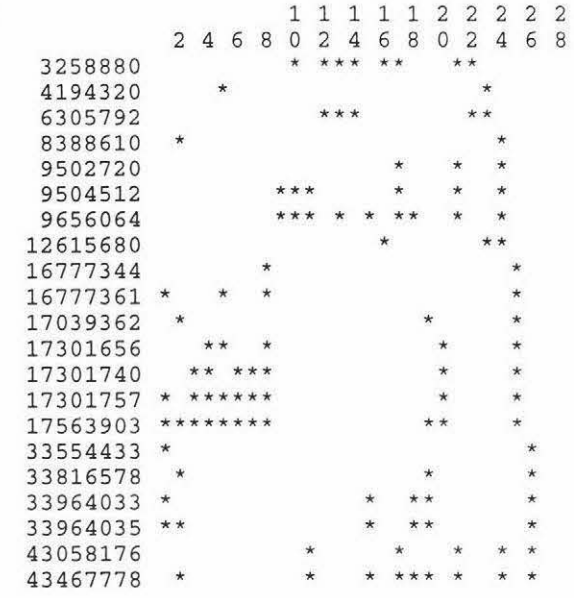
1 *C. aenea* (So)	2 *C. aenea* (PK)	3 *C. whitakeri*	4 *C. alani*
5 *C. macgregori*	6 *C. oliveri* (Mo)	7 *C.oliveri* (GM)	8 *C. ornata*
9 *L. stenotis*	10 *L. grande*	11 *L. notosaurus*	12 *L. lineoocellatum\L. chloronoton*
13 *L. suteri*	14 *L. nigriplantare nigriplantare*		15 *L. microlepis*
16 *L. acrinasum*	17 *L. inconspicuum*	18 *L. smithi*	19 *L. zelandicum*
20 *L. fallai*	21 *L. maccanni*	22 *L. nigriplantare polychroma*	
23 *L. infrapunctatum*	24 *L. otagense*	25 *L. moco*	26 *L. telfairi*
27 *Lampropholis guichenoti*		28 *Tropidoscincus rohssii*	

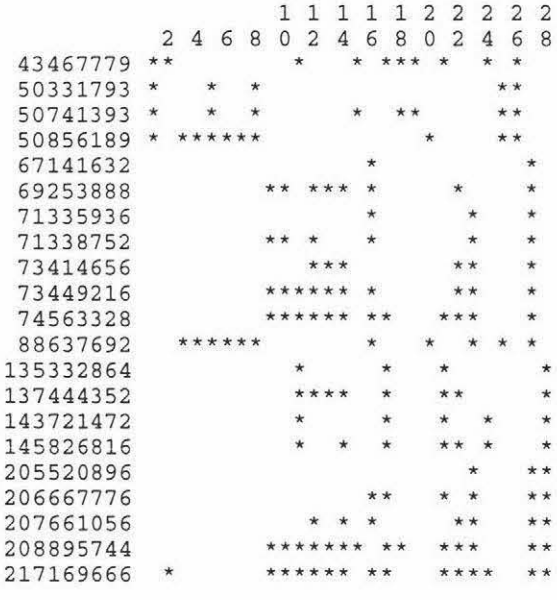
Taxa included in each bipartition:

```
                1 1 1 1 1 2 2 2 2 2
      2 4 6 8 0 2 4 6 8 0 2 4 6 8
   17 *   *
   96       **
  116   * ***
  145 *   *  *
  244   * ****
 1792          ***
 4864          **  *
 5888          *** *
66560            *       *
148224          **     *   *
152320          **   * *   *
409600                 *   **
524296    *                    *
524396   ** **                 *
524540   ******               *
1114112                    *    *
1118720           *   *    *    *
1267456           *** * * **   *
2106368             *  *         *
2110464             * **         *
2111488               ***        *
2140160               * * *      *
2259712           ** ****  *     *
```

```
                1 1 1 1 1 2 2 2 2 2
      2 4 6 8 0 2 4 6 8 0 2 4 6 8
   68   *   *
  100   *  **
  144      *  *
  228   *  ***
  768         **
 4608           *  *
 5120            * *
 6144             **
147456                   *   *
151552                  * *  *
262146 *                      *
409602 *                 *   **
524312    **                    *
524412   *****                  *
524541 * ******                 *
1115136             *       *    *
1120000          *** *    *    *
2105344                 *         *
2107392              * *          *
2111232           *** **          *
2113280           ******          *
2258944              ****   *     *
3226624              ****   *    **
```


1 1 1 1 1 2 2 2 2 2
2 4 6 8 0 2 4 6 8 0 2 4 6 8
3227392 * * * * * * * * * *
3374848 * * * * * * * * * *
4227072 * * *
6334464 * * * * * *
8421376 * * *
9503744 * * * * *
9508608 * * * * * * *
11616000 * * * * * * * *
15843072 * * * * * * * * *
16777360 * * *
16777444 * * * *
17039507 * * * * * *
17301657 * * * * * *
17301756 * * * * * * *
17563902 * * * * * * * *
21496060 * * * * * * *
33554449 * * *
33816595 * * * * *
33964034 * * * * * *
42352643 * * * * * * *
43058177 * * * * * * * *


1 1 1 1 1 2 2 2 2 2
2 4 6 8 0 2 4 6 8 0 2 4 6 8
3258880 * * * * * * *
4194320 * * *
6305792 * * * * *
8388610 * * *
9502720 * * * *
9504512 * * * * * *
9656064 * * * * * * * *
12615680 * * * *
16777344 * * *
16777361 * * * * *
17039362 * * * *
17301656 * * * * *
17301740 * * * * * *
17301757 * * * * * * *
17563903 * * * * * * * *
33554433 * * *
33816578 * * * *
33964033 * * * * *
33964035 * * * * * *
43058176 * * * * * * *
43467778 * * * * * * * * *


1 1 1 1 1 2 2 2 2 2
2 4 6 8 0 2 4 6 8 0 2 4 6 8
43467779 * * * * * * * * * *
50331793 * * * * * * *
50741393 * * * * * * * *
50856189 * * * * * * * * *
67141632 * * *
69253888 * * * * * * * *
71335936 * * * *
71338752 * * * * * * *
73414656 * * * * * *
73449216 * * * * * * * *
74563328 * * * * * * * * *
88637692 * * * * * * * * * *
135332864 * * * * * *
137444352 * * * * * * *
143721472 * * * * * * * *
145826816 * * * * * * * * *
205520896 * * * *
206667776 * * * * * * *
207661056 * * * * * * *
208895744 * * * * * * * * *
217169666 * * * * * * * * *

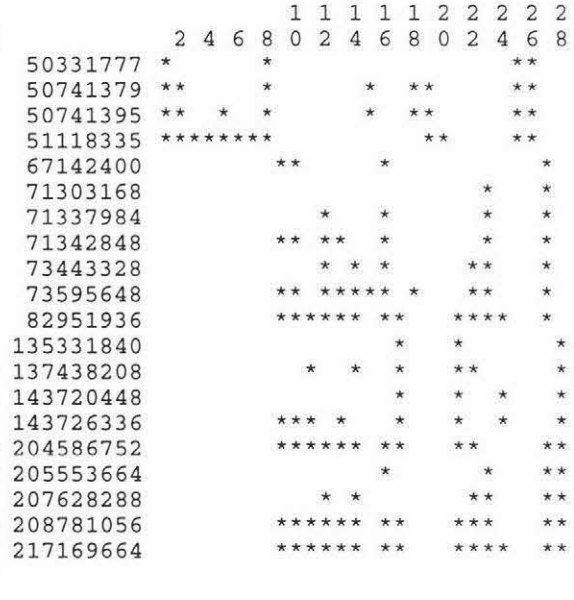
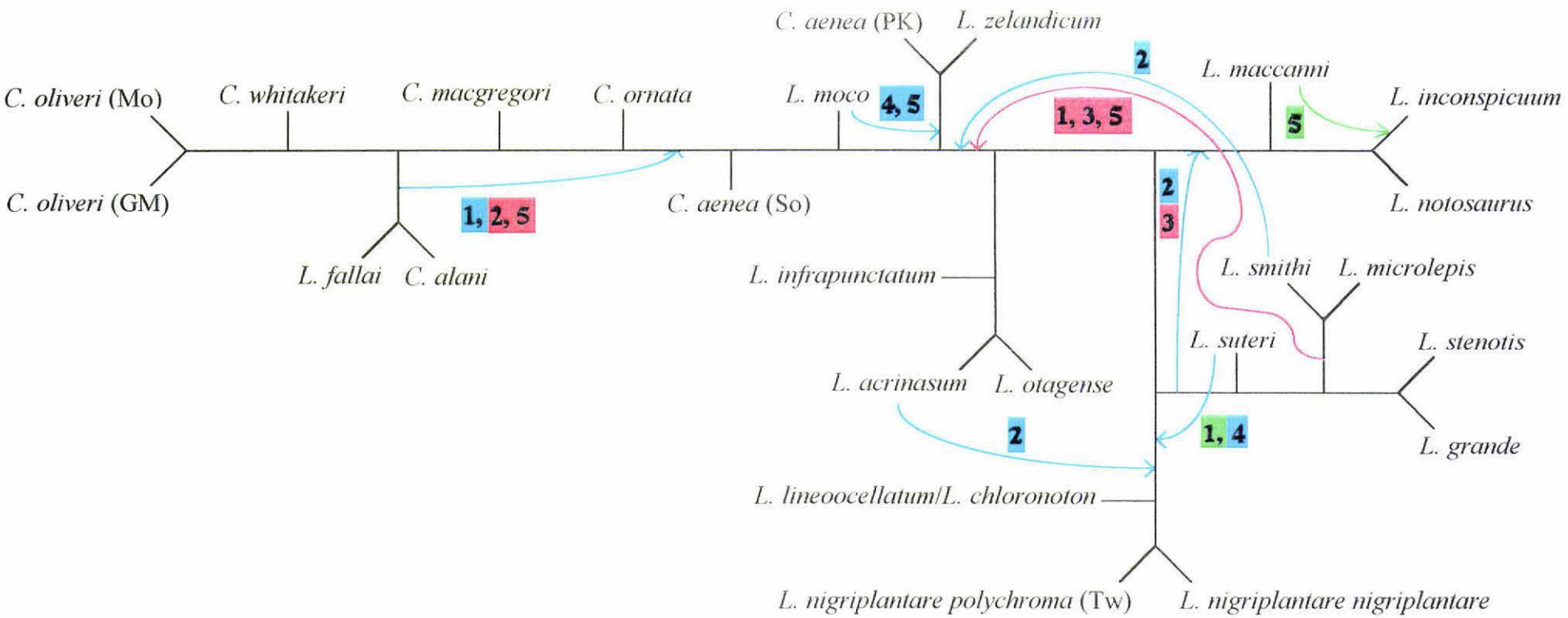

1 1 1 1 1 2 2 2 2 2
2 4 6 8 0 2 4 6 8 0 2 4 6 8
50331777 * * * * * *
50741379 * * * * * * * *
50741395 * * * * * * * *
51118335 * * * * * * * * *
67142400 * * * * * *
71303168 * * * * *
71337984 * * * * * *
71342848 * * * * * * *
73443328 * * * * * * *
73595648 * * * * * * * *
82951936 * * * * * * * *
135331840 * * * * * *
137438208 * * * * * * *
143720448 * * * * * * * *
143726336 * * * * * * * * *
204586752 * * * * * * * * *
205553664 * * * * * *
207628288 * * * * * * *
208781056 * * * * * * * *
217169664 * * * * * * * *

Figure 4.15 Subtree agreement tree for spectral analyses of the 25 New Zealand skink taxa (374 column dataset). The tree was constructed through comparison of the five subtrees resulting from spectral analysis of pairwise distances (Log Det corrected) from five subsets of the 25 taxa dataset (see text). The taxa omitted in each of the five subsets are listed below the tree. Green arrows indicate the differences between each Log Det subtree (numbered 1-5) and the agreement tree. The five subsets were also analysed using Jukes-Cantor (J-C) corrected distances and spectral analysis, and red arrows indicate the differences between the J-C subtrees (also numbered 1-5) and the agreement tree. Rearrangements which occur in both the Log Det and J-C subtrees are shown in blue. The numbers indicate which subtree the rearrangement occurred in and are colour-coded as above. For example - in Log Det subtree #5, *L. maccanni* is more closely related to *L. inconspicuum* than is *L. notosaurus*. The tree is unrooted and is not drawn to scale.



Subset tree #1 lacks *L. acrinasum*, *L. otagense*, *L. grande*, *L. stenotis*, *L. maccanni*, *L. inconspicuum* and *L. notosaurus*.
Subset tree #2 lacks *L. microlepis*, *L. otagense*, *L. infrapunctatum*, *L. stenotis* and *L. notosaurus*.
Subset tree #3 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *C. alani* and *L. fallai*.
Subset tree #4 lacks *L. grande*, *L. stenotis*, *L. maccanni*, *L. inconspicuum* and *L. notosaurus*.
Subset tree #5 lacks *L. acrinasum*, *L. infrapunctatum*, *L. lineoocellatum/L. chloronoton*, *L. nigriplantare nigriplantare* and *L. nigriplantare polychroma* (Tw).

Figure 4.16 Minimum evolution phylogeny for the 25 New Zealand skink taxa (374 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.52863. The tree is unrooted. PAUP* Version 4.0.0d38 was used for the analysis.

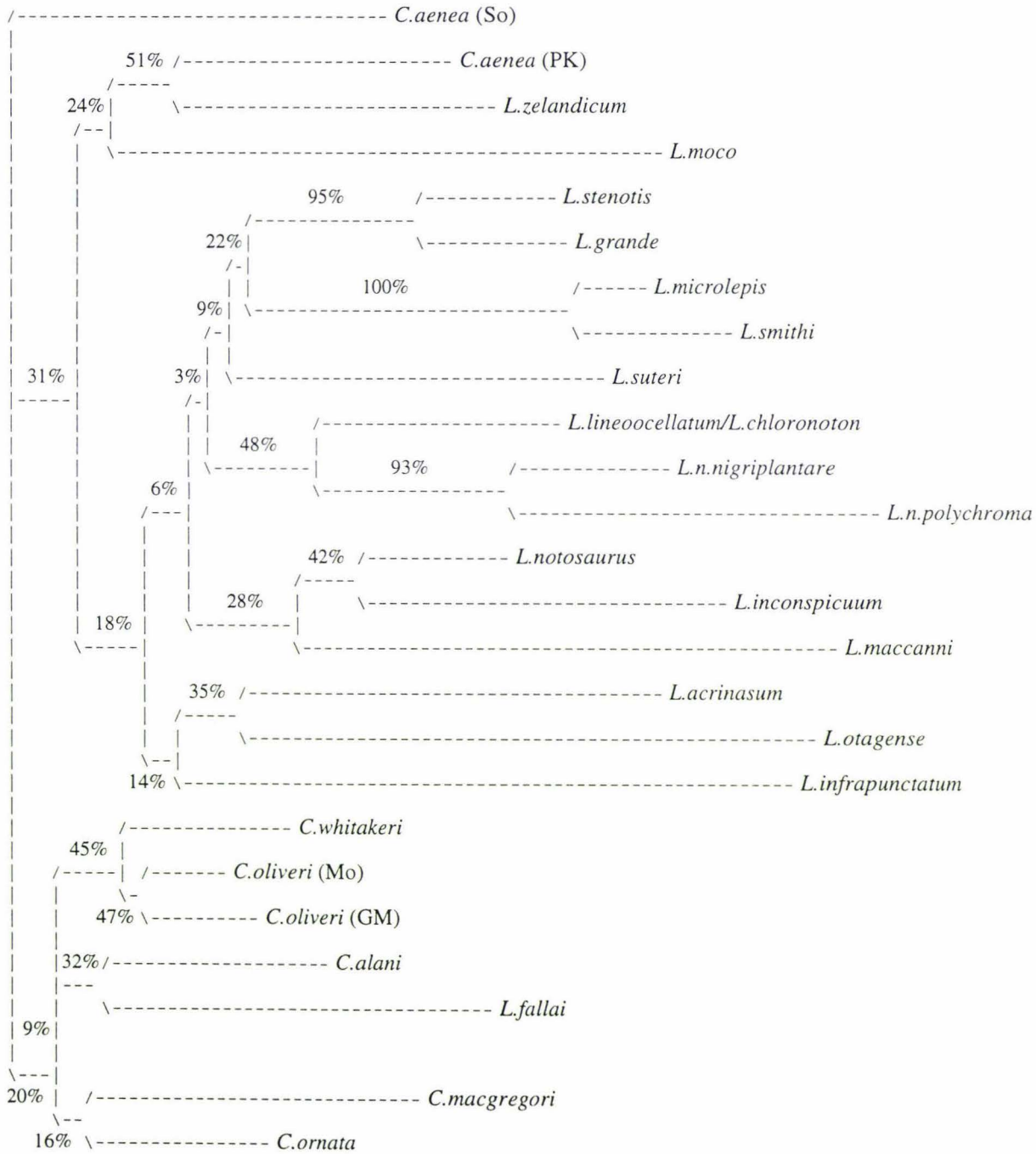
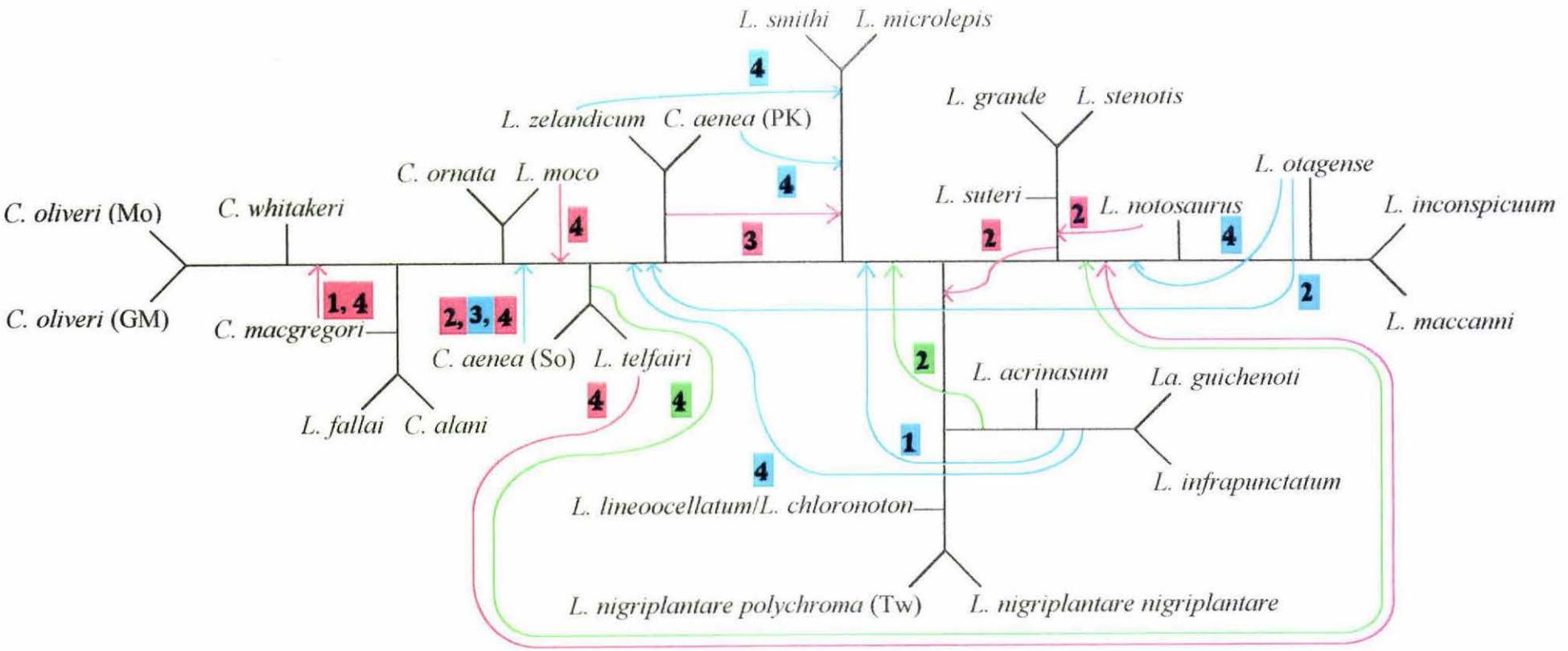


Table 4.12. Summary of the bipartitions selected in phylogenetic analysis of the 25 New Zealand skink taxa (374 column dataset). Five subsets of this dataset were investigated using spectral analysis and the resulting subtrees are labeled closest tree #1-5 (the taxa omitted from each subset are indicated in Figure 4.15). Details of the analyses and of the abbreviations used are given in the text. Shaded squares indicate bipartitions which cannot occur in the closest tree subtrees (because they include taxa which have been omitted from those analyses). Each bipartition included in the optimal (closest tree) or bootstrap consensus tree is indicated by showing the rank (first number) and edge length (second number) of that bipartition. In the closest tree phylogenies, these ranks are based on the strength of the signals and clashes - the largest (signal minus clash) value is ranked first, and so on - and indicate the order of the bipartitions in a spectral diagram. In the maximum parsimony and minimum evolution trees, rank is based on bootstrap value (for example, the bipartition with the highest bootstrap value is ranked as number one and so on). The edge lengths shown for the closest tree analyses are the q (corrected) values. The ranks of three specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]), 96 (*C. oliveri* [Mo] and *C. oliveri* [GM]) and 255 (the eight *Cyclodina* taxa) - are also shown so that the support for these three bipartitions can be investigated. The ranks of these bipartitions were determined as for Table 4.11. Bipartitions which have 50% or greater bootstrap support in the bootstrap trees are indicated with bold+underlined rank and edge length values.

Optimality criterion:	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		Closest Tree #5		MP	ME
Number of taxa:	18		20		20		20		20		25	
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
Bipartitions												
96	6; 0.0030	7; 0.0030	6; 0.0030	6; 0.0031			6; 0.0031	7; 0.0031	5; 0.0033	5; 0.0034	4; 2	6; 0.0013
100	5; 0.0034	5; 0.0031	4; 0.0034	5; 0.0031			5; 0.0034	6; 0.0031	4; 0.0036	4; 0.0034	8; 2	7; 0.0041
116	28; 0.0010	39; 0.0009	29; 0.0010						35; 0.0010			
144					6; 0.0033	5; 0.0044					20; 0	25; 0.0022
145					5; 0.0039	7; 0.0037						
244	26; 0.0011	26; 0.0013	22; 0.0012						29; 0.0014		73=; 3	
768					3; 0.0090	3; 0.0104			2; 0.0085	2; 0.0100	**3; 2**	**2; 0.0106**
4608			18; 0.0015	19; 0.0016								
4864					18; 0.0015				21; 0.0017			
66560					10; 0.0026	11; 0.0027			16; 0.0019			8; 0.0041
147456	1; 0.0184	1; 0.0199			1; 0.0181	1; 0.0199	1; 0.0186	1; 0.0202	1; 0.0180	1; 0.0200	**1; 9**	**1; 0.0194**
148224						22; 0.0016				30; 0.0018	28; 1	18; 0.0004
152320						18; 0.0018				8; 0.0028		45; 0.0019
262146	3; 0.0037	3; 0.0044	3; 0.0037	3; 0.0042	4; 0.0043	4; 0.0048	3; 0.0041	3; 0.0047	3; 0.0045	3; 0.0050	6; 1	**4; 0.0038**
524296	15; 0.0017	15; 0.0018	13; 0.0020	10; 0.0021			12; 0.0019	13; 0.0022	15; 0.0018	20; 0.0020	12; 1	11; 0.0028
524396				40; 0.0010			33; 0.0010	40; 0.0010		38; 0.0011		47; -0.0002
524412				22; 0.0013			28; 0.0012	24; 0.0015		31; 0.0016		
524540	31; 0.0009	27; 0.0012	19; 0.0012	18; 0.0016			11; 0.0020	8; 0.0026	8; 0.0024	7; 0.0030	30; 1	20; 0.0030

Optimality criterion:	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		Closest Tree #5		MP	ME
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
524541	8; 0.0029	9; 0.0029	10; 0.0024	11; 0.0021			8; 0.0026	11; 0.0024	7; 0.0028	11; 0.0026	18; 4	12; 0.0042
1114112			2; 0.0071	2; 0.0078						14; 0.0023	9; 3	
1115136					9; 0.0028	9; 0.0032			6; 0.0030	6; 0.0032	10; 3	13; 0.0068
1118720			17; 0.0018	14; 0.0022								
1120000					40; 0.0011				11; 0.0022			
1267456										17; 0.0024		
2105344	2; 0.0104	2; 0.0117	1; 0.0120	1; 0.0131	2; 0.0110	2; 0.0122	2; 0.0107	2; 0.0120			**2; 5**	**3; 0.0118**
2107392	4; 0.0037	4; 0.0041	5; 0.0033	4; 0.0038	7; 0.0033	6; 0.0038	4; 0.0038	4; 0.0044			5; 5	5; 0.0070
2111488	7; 0.0032	8; 0.0031					9; 0.0027	10; 0.0026			48; 2	
2140160			11; 0.0024	13; 0.0022								
2258944		10; 0.0031					14; 0.0020	12; 0.0025				
2259712						57; 0.0010						89=; 0.0007
3226624											74; 1	
3227392					35; 0.0012							
3258880			14; 0.0021	17; 0.0020								
3374848						28; 0.0015					69=; 4	65=; 0.0026
6305792	10; 0.0028											
8421376					11; 0.0026	8; 0.0032	7; 0.0031	5; 0.0038			13=; 2	10; 0.0042
9508608									12; 0.0023			
12615680					27; 0.0013	30; 0.0014	27; 0.0014	31; 0.0014			29=; 3	33; 0.0022
15843072					17; 0.0017							
16777361					8; 0.0029	10; 0.0031						
17039362							13; 0.0020	15; 0.0021	14; 0.0021	15; 0.0024	13=; 4	14; 0.0022
17039507					16; 0.0018	19; 0.0018						
17301757	9; 0.0028	6; 0.0032	8; 0.0024	7; 0.0027								
17563903	13; 0.0022	13; 0.0024	16; 0.0018	16; 0.0020			21; 0.0016	23; 0.0017	13; 0.0023	16; 0.0024	37; 2	22=; 0.0037
Rank of:												
3	131032 | 131044 (out of 131072 possible)		524212 | 524227 (out of 524288 possible)		523975 | 524145 (out of 524288 possible)		524023 | 524051 (out of 524288 possible)		523896 | 523298 (out of 524288 possible)		>77= (<2.5%)	>94= (<2.5%)
96	24 | 26 (out of 131072)		25 | 25 (out of 524288)				25 | 26 (out of 524288)		24 | 23 (out of 524288)		4 (38%)	6 (47%)
255	129423 | 130241 (out of 131072)		519006 | 521702 (out of 524288)				520062 | 521994 (out of 524288)		522512 | 523123 (out of 524288)		>77= (<2.5%)	>94= (<2.5%)

Figure 4.17 Subtree agreement tree for spectral analyses of the 25 New Zealand skink taxa, *Leiolopisma telfairi* and *Lampropholis guichenoti* (368 column dataset). The tree was constructed through comparison of the four subtrees resulting from spectral analysis of pairwise distances (Log Det corrected) from four subsets of the 27 taxa dataset (see text). The taxa omitted in each of the four subsets are listed below the tree. Green arrows indicate the differences between each Log Det subtree (numbered 1-4) and the agreement tree. The four subsets were also analysed using Jukes-Cantor (J-C) corrected distances and spectral analysis, and red arrows indicate the differences between the J-C subtrees (also numbered 1-4) and the agreement tree. Rearrangements which occur in both the Log Det and J-C subtrees are shown in blue. The numbers indicate which subtree the rearrangement occurred in and are colour-coded as above. The tree is unrooted and is not drawn to scale.



Subset tree #1 lacks *L. acrinasum*, *L. otagense*, *L. grande*, *L. stenotis*, *L. maccanni*, *L. inconspicuum* and *L. notosaurus*.
Subset tree #2 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *L. microlepis*, *L. smithi*, *C. aenea* (PK) and *L. zelandicum*.
Subset tree #3 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *C. alani*, *L. fallai*, *C. ornata* and *L. moco*.
Subset tree #4 lacks *L. acrinasum*, *L. lineoocellatum/L. chloronoton*, *L. nigriplantare nigriplantare*, *L. nigriplantare polychroma* (Tw), *L. grande*, *L. stenotis* and *L. suteri*.

Figure 4.18 Minimum evolution phylogeny for the 25 New Zealand skink taxa, *Leiolopisma telfairi* and *Lampropholis guichenoti* (368 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.51748. The tree is unrooted. PAUP* Version 4.0.0d38 was used for the analysis.

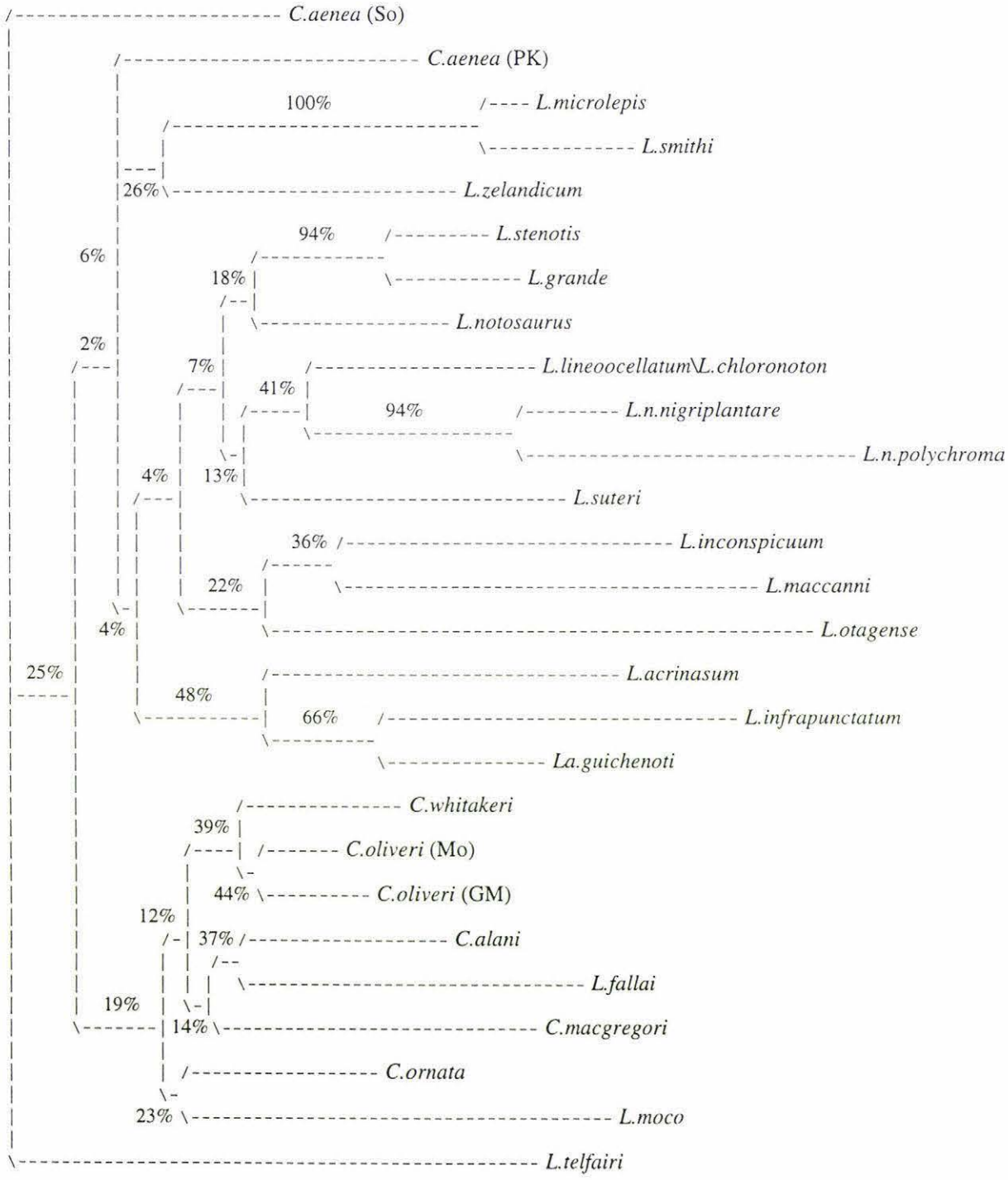
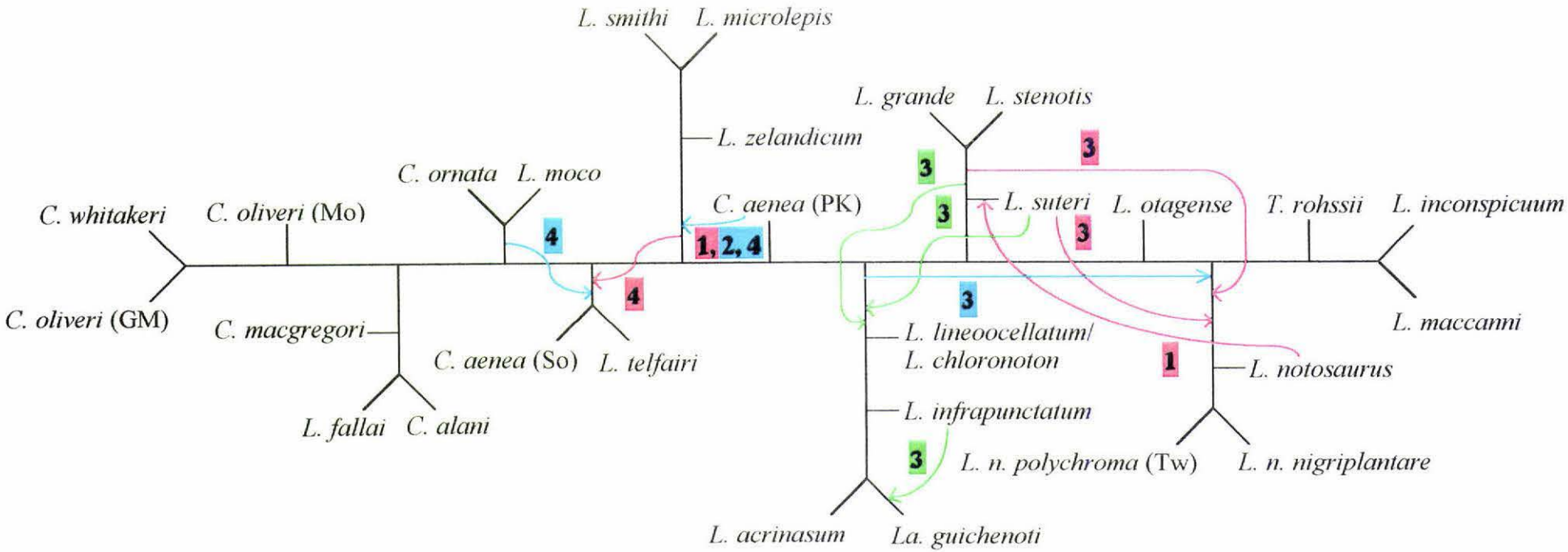


Table 4.13. Summary of the bipartitions selected in phylogenetic analysis of 27 skink taxa - the 25 New Zealand skinks, *L. telfairi* and *La. guichenoti* (368 column dataset). Four subsets of this dataset were investigated using spectral analysis and the resulting subtrees are labeled closest tree #1-4 (the taxa omitted from each subset are indicated in Figure 4.17). Details of the analyses and of the abbreviations used are given in the text. Shaded squares indicate bipartitions which cannot occur in the closest tree subtrees (because they include taxa which have been omitted from those analyses). Each bipartition included in the optimal (closest tree) or bootstrap consensus tree is indicated by showing the rank (first number) and edge length (second number) of that bipartition (see Table 4.12 for an explanation of these ranks). The edge lengths shown for the closest tree analyses are the q (corrected) values. The ranks of three specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]), 96 (*C. oliveri* [Mo] and *C. oliveri* [GM]) and 255 (the eight *Cyclodina* taxa) - are also shown so that the support for these three bipartitions can be investigated. The ranks of these bipartitions were determined as for Table 4.11. Bipartitions which have 50% or greater bootstrap support in the bootstrap trees are indicated with bold+underlined rank and edge length values.

Optimality criterion:	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		MP	ME
Number of taxa:	20		20		20		20		27	
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
Bipartitions										
17					6; 0.0045	6; 0.0041				
96	5; 0.0028	6; 0.0027					6; 0.0028	5; 0.0027	6; 2	6; 0.0013
100	6; 0.0028	9; 0.0025					5; 0.0029	7; 0.0026	11; 1	9; 0.0035
116	31; 0.0010						27; 0.0011			
768			2; 0.0086	2; 0.0099	3; 0.0082	3; 0.0096			**3; 3**	**3; 0.0082**
1792										24; 0.0020
4864			31; 0.0013	27; 0.0016	41; 0.0008	39; 0.0008				
5888			37; 0.0013							
147456	1; 0.0183	1; 0.0199			1; 0.0189	1; 0.0206	1; 0.0195	1; 0.0218	**1; 7**	**1; 0.0192**
262146	26; 0.0011	24; 0.0014			12; 0.0018	14; 0.0021				
409600							10; 0.0018	11; 0.0022	18; 1	14; 0.0026
409602					22; 0.0014		20; 0.0013	15; 0.0016		71; 0.0004
524296	11; 0.0020	10; 0.0022	10; 0.0024	11; 0.0026			9; 0.0020	9; 0.0022	14; 3	10; 0.0017
524312		34; 0.0010	15; 0.0019	15; 0.0023				35; 0.0010	45=; 2	30; 0.0015
524412	22; 0.0013	26; 0.0014					18; 0.0013	20; 0.0014	78=; 1	34=; 0.0013
524540							23; 0.0012			
524541							21; 0.0013			
1114112			7; 0.0031	6; 0.0038	8; 0.0034	8; 0.0039	11; 0.0017	10; 0.0022	5; 1	11; 0.0045
1115136				19; 0.0021			3; 0.0047	3; 0.0054		
1120000				31; 0.0017						
2105344	3; 0.0116	3; 0.0129	1; 0.0118	1; 0.0129	2; 0.0116	2; 0.0125			**2; 5**	**2; 0.0131**
2107392	4; 0.0033	4; 0.0038	8; 0.0029	7; 0.0032	9; 0.0029	9; 0.0032			7; 4	7; 0.0034
2111488	7; 0.0028	7; 0.0027							42=; 2	33; 0.0015
2113280										58; 0.0024

Optimality criterion:	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		MP	ME
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
3227392				24; 0.0020						
9502720					11; 0.0021	12; 0.0025			13= 5	17; 0.0050
9503744					26; 0.0013	20; 0.0017	4; 0.0030	4; 0.0038	42=; 2	
9504512									76=; 0	
9508608					74; 0.0008	56; 0.0010				
11616000									74=; 3	81=; 0.0024
16777344	24; 0.0013	20; 0.0017	21; 0.0016	17; 0.0023				17; 0.0015	15; 1	16=; 0.0019
17301656			11; 0.0025	12; 0.0027						
17301657			6; 0.0032							
17301756	19; 0.0015	21; 0.0018						18; 0.0015	33; 2	22; 0.0049
17301757							14; 0.0015			
33554433	10; 0.0023	5; 0.0031		8; 0.0030				12; 0.0019	13=; 3	15; 0.0041
33554449					14; 0.0018	11; 0.0024				
33816595						21; 0.0015				
33964033									72=; 1	
33964035									>89=; 0	
43058176							33; 0.0010			
43058177								31; 0.0013		
43467778							29; 0.0012			
43467779								25; 0.0014		
50856189	25; 0.0013	22; 0.0017								102=; 0.0024
51118335	13; 0.0022	11; 0.0024								
71303168	2; 0.0119	2; 0.0132	3; 0.0073	3; 0.0085	4; 0.0072	4; 0.0085	2; 0.0143	2; 0.0154	**4; 4**	**4; 0.0070**
71335936			4; 0.0050	4; 0.0051	5; 0.0047	5; 0.0048			8; 8	5; 0.0074
73414656	9; 0.0025	13; 0.0022								
73443328			24; 0.0018		20; 0.0015	25; 0.0015				
73449216			18; 0.0020							
74563328			9; 0.0026	10; 0.0028						
82951936			13; 0.0025	9; 0.0029	24; 0.0014	26; 0.0014				81=; 0.0012
88637692									70=; 2	
Rank of:										
3	524261 | 524264 (out of 524288 possible)				524129 | 524138 (out of 524288 possible)		524240 | 524250 (out of 524288 possible)		>89= (<1%)	>107= (<1.5%)
96	25 | 26 (out of 524288)						25 | 25 (out of 524288)		6 (37%)	6 (44%)
255	1218 | 1578 (out of 524288)						1726 | 2361 (out of 524288)		>89= (<1%)	>107= (<1.5%)

Figure 4.19 Subtree agreement tree for spectral analyses of the 25 New Zealand skink taxa, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (298 column dataset). The tree was constructed through comparison of the four subtrees resulting from spectral analysis of pairwise distances (Log Det corrected) from four subsets of the 28 taxa dataset (see text). The taxa omitted in each of the four subsets are listed below the tree. Green arrows indicate the differences between each Log Det subtree (numbered 1-4) and the agreement tree. The four subsets were also analysed using Jukes-Cantor (J-C) corrected distances and spectral analysis, and red arrows indicate the differences between the J-C subtrees (also numbered 1-4) and the agreement tree. Rearrangements which occur in both the Log Det and J-C subtrees are shown in blue. The numbers indicate which subtree the rearrangement occurred in and are colour-coded as above. The tree is unrooted and is not drawn to scale.



Subset tree #1 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *C. alani*, *L. fallai*, *L. lineoocellatum/L. chloronoton*, *L. n. nigriplantare* and *L. n. polychroma* (Tw).
Subset tree #2 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *C. alani*, *L. fallai*, *L. grande*, *L. stenotis* and *L. suteri*.
Subset tree #3 lacks *C. oliveri* (Mo), *C. oliveri* (GM), *C. whitakeri*, *C. alani*, *L. fallai*, *C. macgregori*, *L. microlepis* and *L. smithi*.
Subset tree #4 lacks *L. lineoocellatum/L. chloronoton*, *L. grande*, *L. stenotis*, *L. suteri*, *L. otagense*, *L. notosaurus*, *L. n. nigriplantare* and *L. n. polychroma* (Tw).

Figure 4.20 Minimum evolution phylogeny for the 25 New Zealand skink taxa, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (298 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.55158. The tree is unrooted. PAUP* Version 4.0.0d39 was used for the analysis.

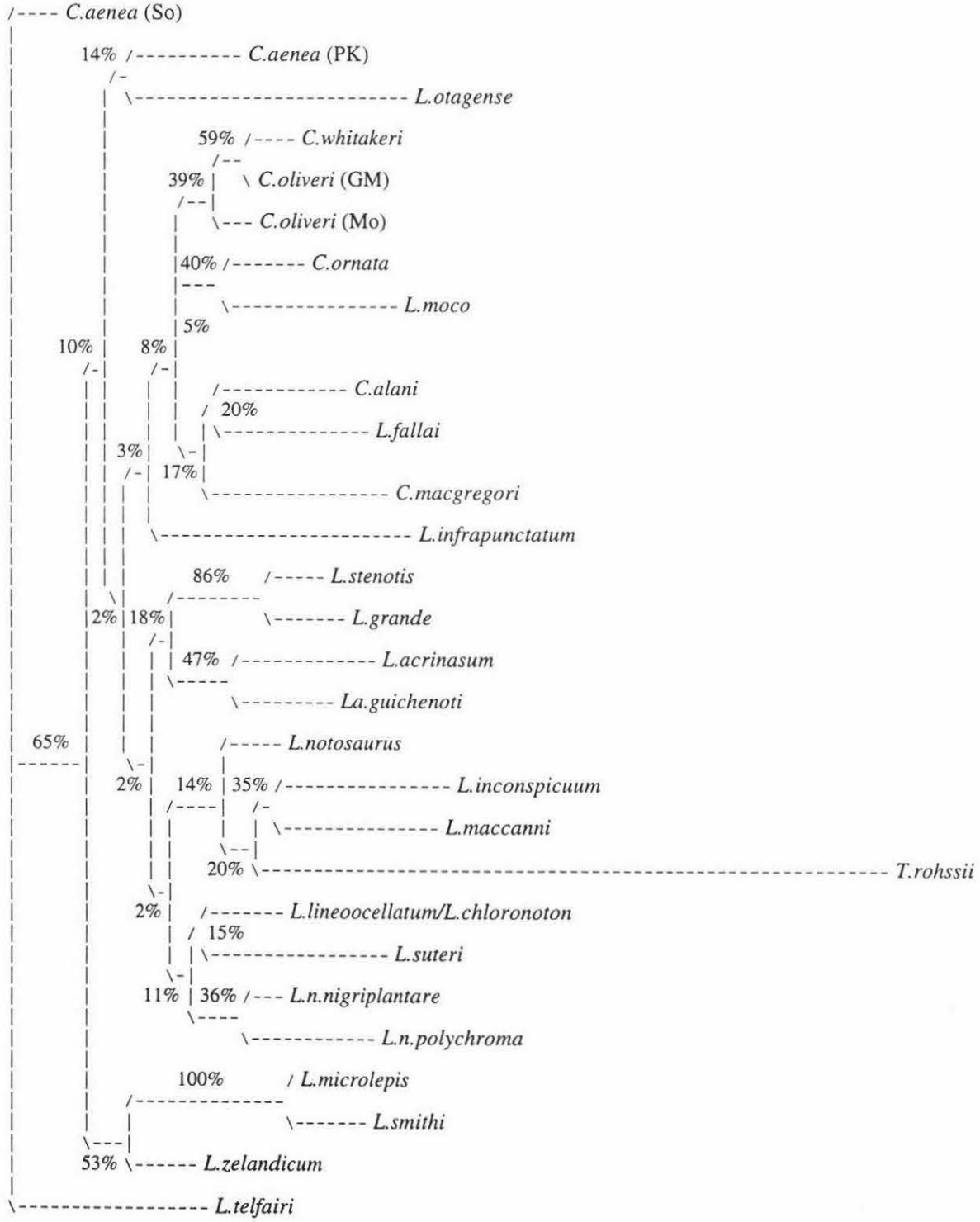


Table 4.14. Summary of the bipartitions selected in phylogenetic analysis of 28 skink taxa - the 25 New Zealand skinks, *L. telfairi*, *La. guichenoti* and *T. rohssii* (298 column dataset). Four subsets of this dataset were investigated using spectral analysis and the resulting subtrees are labeled closest tree #1-4 (the taxa omitted from each subset are indicated in Figure 4.19). Details of the analyses and of the abbreviations used are given in the text. Shaded squares indicate bipartitions which cannot occur in the closest tree subtrees (because they include taxa which have been omitted from those analyses). Each bipartition included in the optimal (closest tree) or bootstrap consensus tree is indicated by showing the rank (first number) and edge length (second number) of that bipartition (see Table 4.12 for an explanation of these ranks). The edge lengths shown for the closest tree analyses are the q (corrected) values. The ranks of three specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]), 96 (*C. oliveri* [Mo] and *C. oliveri* [GM]) and 255 (the eight *Cyclodina* taxa) - are also shown so that the support for these three bipartitions can be investigated. The ranks of these bipartitions were determined as for Table 4.11. Bipartitions which have 50% or greater bootstrap support in the bootstrap trees are indicated with bold+underlined rank and edge length values.

Optimality criterion:	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		MP	ME
Number of taxa:	20		20		20		20		28	
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
Bipartitions										
68							8; 0.0033	9; 0.0032	6; 1	**4; 0.0034**
100							10; 0.0029	11; 0.0025	11; 1	9; 0.0034
768	2; 0.0091	2; 0.0103			1; 0.0096	1; 0.0108			**2; 3**	**2; 0.0112**
4864		33; 0.0009								
5120	31; 0.0009									
5888	15; 0.0018									
6144									26; 1	24; 0.0001
147456	1; 0.0192	1; 0.0208	1; 0.0194	1; 0.0209			1; 0.0200	1; 0.0219	**1; 6**	**1; 0.0197**
409600	5; 0.0048	4; 0.0055	3; 0.0051	3; 0.0058			6; 0.0038	5; 0.0045	5; 1	**5; 0.0051**
409602	25; 0.0012		22; 0.0015	27; 0.0015			16; 0.0014	19; 0.0014		
524296							19; 0.0011	21; 0.0011	17; 1	17; -0.0001
524312							23; 0.0009	20; 0.0011		21; 0.0022
524396									83=; 0	
524412							37; 0.0007	34; 0.0008		
1114112	11; 0.0024	13; 0.0023	6; 0.0033	7; 0.0034	7; 0.0031	8; 0.0031	5; 0.0047	6; 0.0043	10; 2	11; 0.0025
2105344			5; 0.0038	5; 0.0039	6; 0.0034	6; 0.0037			8; 2	10; 0.0060
2106368			11; 0.0025	11; 0.0026	25; 0.0014	28; 0.0014				
2110464					51; 0.0010					
2111232					49; 0.0010					
2111488									52=; 0	32; 0.0024
4194320									31; 2	
8388610									18; 1	25=; 0.0023
16777344	8; 0.0028	7; 0.0035	8; 0.0027	8; 0.0033	2; 0.0054	2; 0.0063	11; 0.0026	10; 0.0030	9; 2	7; 0.0046
16777360	9; 0.0025	10; 0.0027	9; 0.0026	9; 0.0028						
16777444										67; 0.0001
17301740									>83=; 0	

	Closest Tree #1		Closest Tree #2		Closest Tree #3		Closest Tree #4		MP	ME
Correction:	J-C	Log Det	J-C	Log Det	J-C	Log Det	J-C	Log Det	uncorr.	J-C
17301756										42; 0.0026
21496060									77=; 2	89=; 0.0030
33554433	3; 0.0057	3; 0.0063	2; 0.0056	2; 0.0063	3; 0.0052	3; 0.0058	3; 0.0070	3; 0.0080	**3; 3**	**3; 0.0096**
33964033									36; 1	35; 0.0015
42352643									>83=; 0	94=; 0.0002
50331777					8; 0.0030	7; 0.0032	18; 0.0014	17; 0.0016		
50331793	12; 0.0022	11; 0.0025	13; 0.0021	13; 0.0024						
50741379							49; 0.0008			
50741393		36; 0.0012								
50741395	42; 0.0010	59; 0.0010	40; 0.0012	49; 0.0011						
50856189								64; 0.0008		
67141632	4; 0.0055	5; 0.0051	4; 0.0050	4; 0.0048	4; 0.0046		4; 0.0060	4; 0.0057	4; 3	6; 0.0081
67142400									23=; 2	20; 0.0018
69253888									>83=; 2	
71303168						4; 0.0045				
71335936	7; 0.0035	8; 0.0034	14; 0.0021	15; 0.0021	9; 0.0030	9; 0.0030	9; 0.0034	8; 0.0035		
71337984			10; 0.0025	12; 0.0025	14; 0.0018	16; 0.0018				
71338752						41; 0.0011				
71342848						40; 0.0011				
73449216					36; 0.0012	36; 0.0012				
135331840	10; 0.0025	9; 0.0031	12; 0.0022	10; 0.0027	12; 0.0021	10; 0.0026	2; 0.0073	2; 0.0089	15; 5	18; 0.0042
135332864		12; 0.0024							23=; 4	25=; 0.0055
137438208			24; 0.0016	23 0.0018						
137444352										98=; 0.0021
143720448	17; 0.0016									
143721472		19; 0.0018								
143726336	38; 0.0011	30; 0.0012								
145826816			33; 0.0013	32; 0.0015						
204586752									>83=; 1	99=; 0.0024
206667776							31; 0.0011	48; 0.0010		
208781056					13; 0.0019	15; 0.0019				
217169664					17; 0.0016	21; 0.0016				
217169666					19; 0.0016	18; 0.0017				
Rank of:										
3	522276 | 518231 (out of 524288 possible)		522749 | 519851 (out of 524288 possible)		524166 | 524124 (out of 524288 possible)		524137 | 524052 (out of 524288 possible)		>83= (<1%)	>100= (<1.5%)
96							35 | 38 (out of 524288)		28= (10%)	15 (21%)
255							486944 | 504710 (out of 524288)		>83= (<1%)	>100= (<1.5%)

Figure 4.21 Minimum evolution phylogeny for the 25 New Zealand skink taxa (386 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.57440. The tree is unrooted. PAUP* Version 4.0.0d38 was used for the analysis.

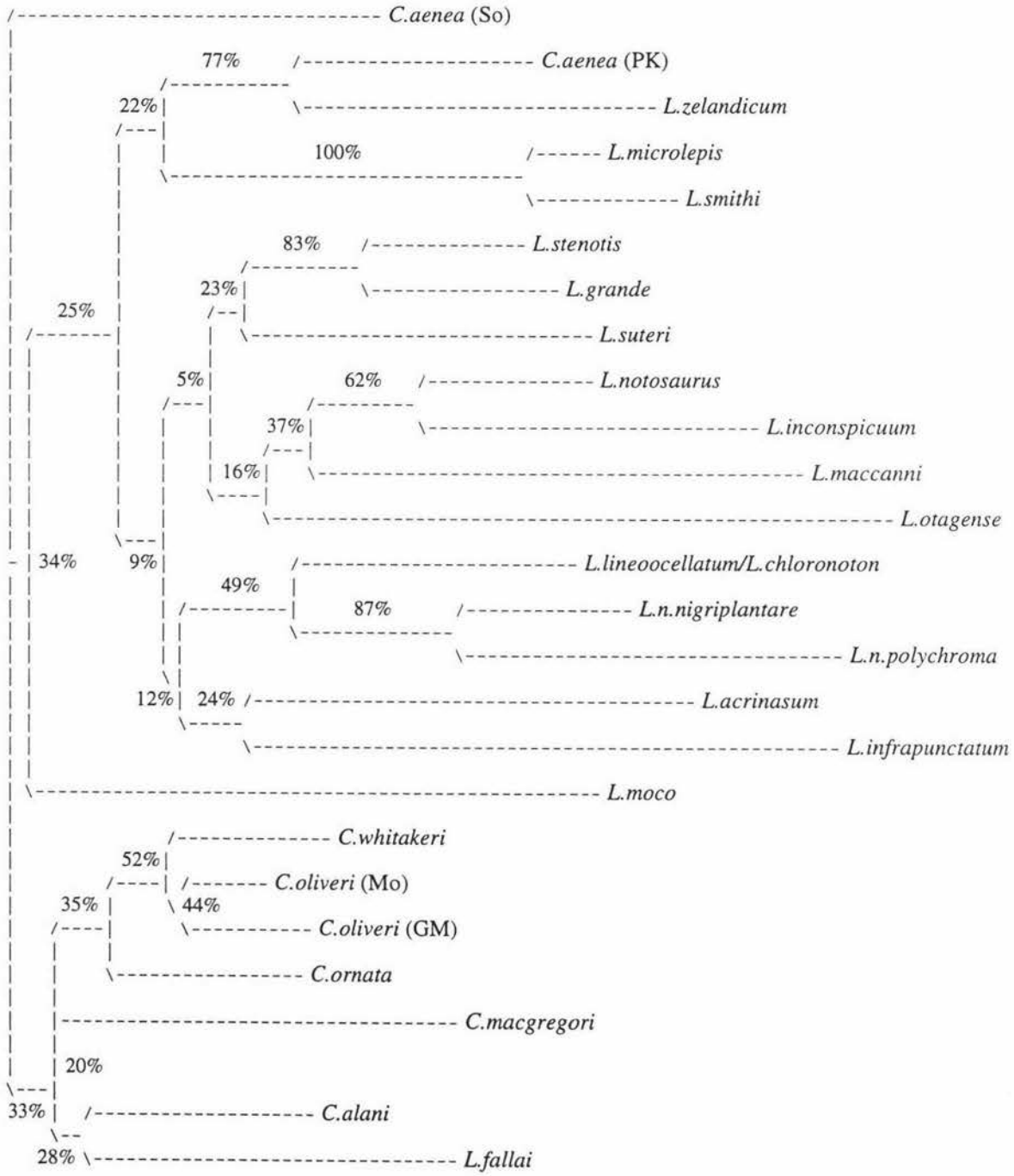


Figure 4.22 Minimum evolution phylogeny for the 25 New Zealand skink taxa, *Leiolopisma telfairi* and *Lampropholis guichenoti* (387 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.61775. The tree is unrooted. PAUP* Version 4.0.0d38 was used for the analysis.

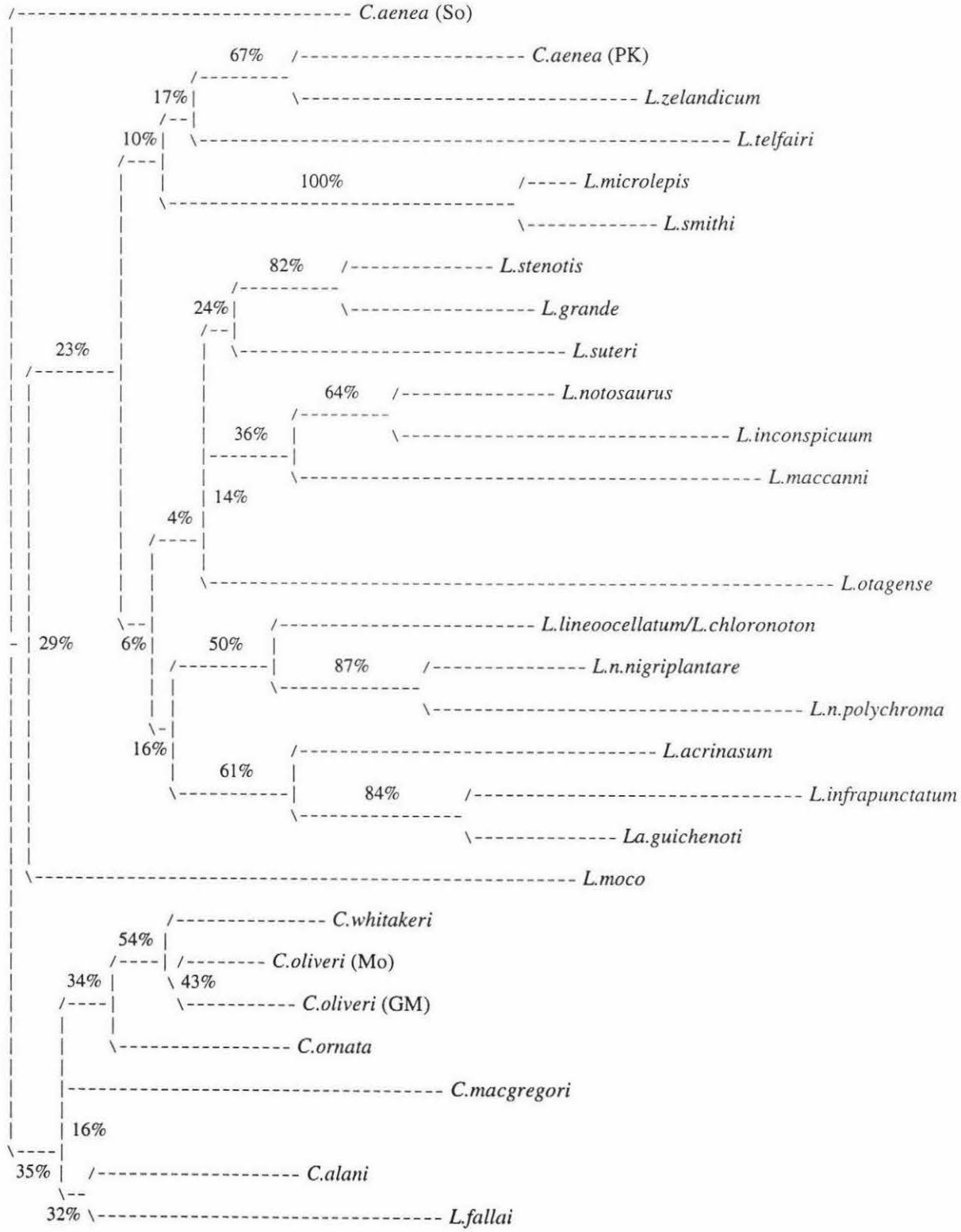


Figure 4.23 Minimum evolution phylogeny for the 25 New Zealand skink taxa, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (387 column dataset). This is a bootstrap 50% majority-rule consensus tree (but with other groups compatible with the majority-rule groupings included) which was derived from Jukes-Cantor corrected distances using a neighbor-joining search. Percentage values on the internal edges indicate how many times each edge occurred in 1000 bootstrap replicates. The edge lengths in the trees are proportional to the number of changes along each edge. The minimum evolution score is 0.69123. The tree is unrooted. PAUP* Version 4.0.0d39 was used for the analysis.

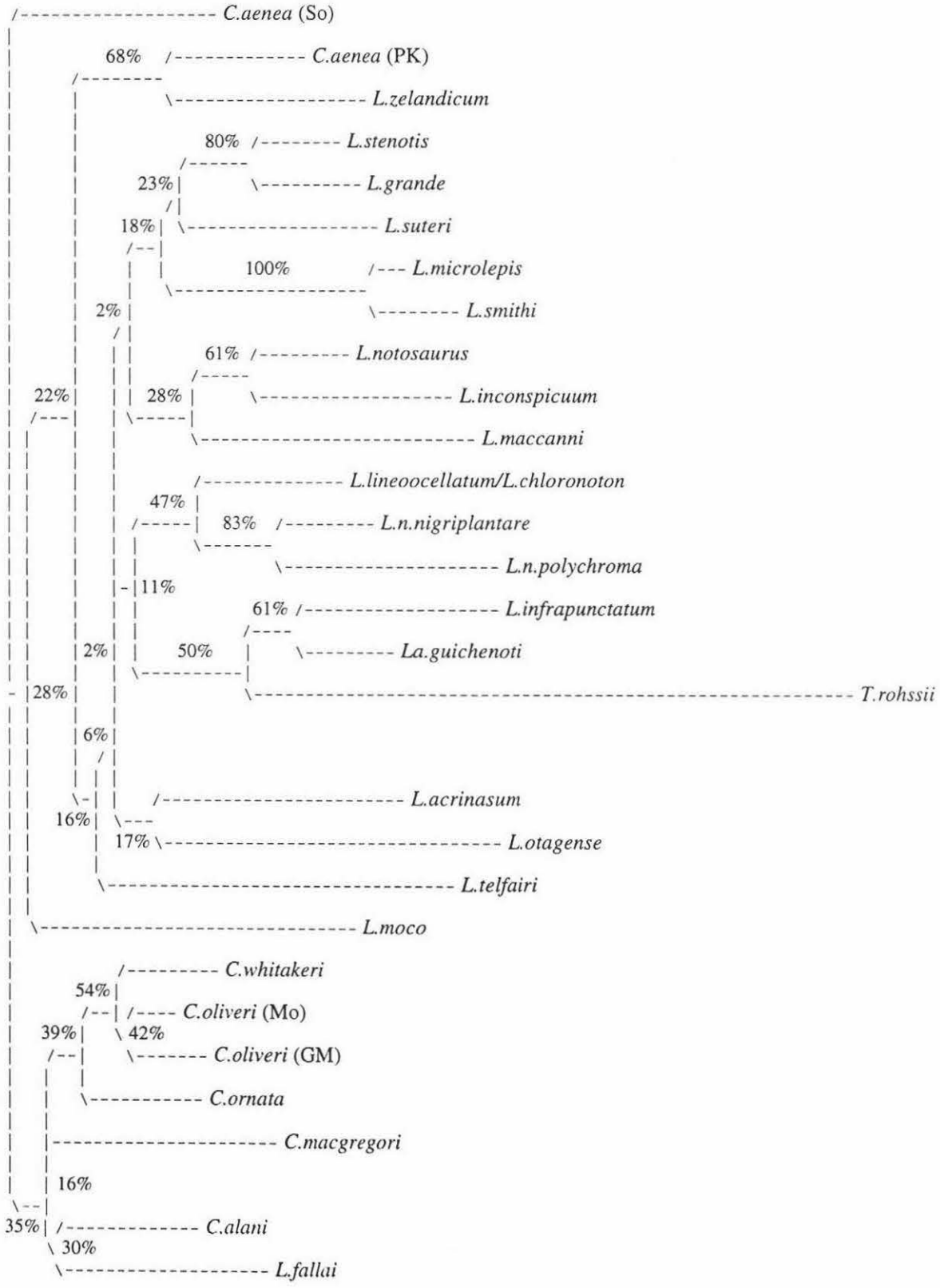


Table 4.15. Summary of the bipartitions selected in phylogenetic analysis of three datasets - the New Zealand skinks (25 taxa/386 columns), the New Zealand skinks, *L. telfairi* and *La. guichenoti* (27 taxa/387 columns) and the New Zealand skinks, *L. telfairi*, *La. guichenoti* and *T. rohssii* (28 taxa/387 columns). Details of the analyses and of the abbreviations used are given in the text. Shaded squares indicate bipartitions which cannot occur for a particular dataset (because they include taxa which have been omitted from that dataset). Each bipartition included in the bootstrap consensus tree is indicated by showing the rank (first number) and edgelength (second number) of that bipartition. The ranks of three specific bipartitions - 3 (*C. aenea* [So] and *C. aenea* [PK]), 96 (*C. oliveri* [Mo] and *C. oliveri* [GM]) and 255 (the eight *Cyclodina* taxa) - are also shown so that the support for these three bipartitions can be investigated. The ranks of these bipartitions are based on bootstrap value (as for Table 4.11). Bipartitions which have 50% or greater bootstrap support in the bootstrap trees are indicated with bold+underlined rank and edge length values.

Dataset:	New Zealand skinks (25 taxa/386 columns)		New Zealand skinks plus *L. telfairi* and *La. guichenoti* (27 taxa/387 columns)		New Zealand skinks plus *L. telfairi*, *La. guichenoti* and *T. rohssii* (28 taxa/387 columns)	
Analysis:	MP (uncorr.)	ME (J-C)	MP (uncorr.)	ME (J-C)	MP (uncorr.)	ME (J-C)
Bipartitions						
96	6; 2	8; 0.0007	7; 2	10; 0.0006	7; 2	9; 0.0004
100	7; 2	**6; 0.0035**	8; 2	**8; 0.0037**	10; 2	**6; 0.0036**
228	12; 2	11; 0.0030	14; 2	14; 0.0031	13; 2	11; 0.0033
244	28; 2	23; 0.0005	38; 3	24; 0.0003	29; 3	27; 0.0006
768	**3; 2**	**3; 0.0077**	**4; 2**	**4; 0.0076**	**3; 2**	**3; 0.0074**
4864		20; 0.0023		18; 0.0018		18; 0.0015
66560	5; 3	**5; 0.0071**	6; 3	**6; 0.0071**	8; 3	**5=; 0.0071**
147456	**1; 9**	**1; 0.0233**	**1; 9**	**1; 0.0237**	**1; 9**	**1; 0.0228**
151552	35; 1		36; 1		34; 1	
152320	46; 2		56; 4		65=; 1	23; 0.0034
262146	**4; 4**	**4; 0.0082**	5; 4	**5; 0.0070**	5; 4	**4; 0.0099**
409602		21; 0.0026				
524296	18; 1	14; 0.0022	18; 0	15; 0.0023	19; 0	13; 0.0190
524540	20; 3	13; 0.0030	19; 3	13; 0.0034	22; 4	12; 0.0037
524541	25; 5	12; 0.0009		16; 0.0006		16; 0.0002
1115136	9; 4	10; 0.0026	11; 3	12; 0.0056	15; 3	15; 0.0071
1120000				32; 0		
1267456						99=; 0.0008
2105344	**2; 6**	**2; 0.0100**	**2; 6**	**2; 0.0102**	**2; 5**	**2; 0.0100**
2107392	8; 3	7; 0.0075	10; 2	**9; 0.0071**	11; 1	8; 0.0061
4227072	14; 3	17; 0.0042				
6334464	51=; 2	36; 0.0008				
8421376						24=; 0.0052
9503744	21=; 4	28; 0.0035	30; 3		53; 5	
9508608		70=; 0.0032		71; 0.0038		

Dataset:	New Zealand skinks (25 taxa/386 columns)		New Zealand skinks plus *L. telfairi* and *La. guichenoti* (27 taxa/387 columns)		New Zealand skinks plus *L. telfairi*, *La. guichenoti* and *T. rohssii* (28 taxa/387 columns)	
Analysis:	MP (uncorr.)	ME (J-C)	MP (uncorr.)	ME (J-C)	MP (uncorr.)	ME (J-C)
9656064	76=; 4				>97=; 5	
15843072		46; 0.0023				
17039362	13; 4		17; 5		18; 4	
17301757		16; 0.0060		19; 0.0061		20; 0.0051
17563902			86=; 5		89=; 3	
17563903	42; 5					28; 0.0024
33554433			16; 5		21; 5	
33816578				22; 0.0019		
33964034				45; 0.0024		
51118335			78=; 3		90=; 3	60=; 0.0009
71303168			**3; 6**	**3; 0.0115**	4; 2	**5=; 0.0059**
71335936			9; 7	**7; 0.0081**		
73443328			42=; 4	25; 0.0014		
73595648			92=; 1			
82951936				63=; 0.0017		
205520896					6; 7	**7; 0.0123**
205553664					30; 5	
207628288						40; 0.0024
207661056					63=; 2	
208895744						99=; 0.0002
Rank of:						
3	>77= (<2.5%)	>71 (<4.5%)	>92= (1%)	>81= (<3.5%)	>97= (<1%)	>101= (<1.5%)
96	6 (38%)	8 (44%)	7 (40%)	10 (43%)	7 (39%)	9 (42%)
255	>77= (<2.5%)	>71 (<4.5%)	>92= (<1%)	>81= (<3.5%)	>97= (<1%)	>101= (<1.5%)

Consensus Trees.

A consensus tree summarising the results of all phylogenetic analyses of the 374 and 368 column datasets is presented in Figure 4.24. A second consensus tree, summarising the results of all analyses of the 386 column, 387 column (27 taxa) and 387 column (28 taxa) datasets, is given in Figure 4.25. The results of the 298 column dataset analyses were used to place *T. rohssii* in Figure 4.24, but were otherwise disregarded when constructing this tree. This was because of the large number of conflicts between trees from the 374 and 368 column datasets and those from the 298 column dataset (presumably due to the large number of sites omitted in the 298 column dataset). These included the disruption of relationships such as {*C. oliveri* [Mo] and *C. oliveri* [GM]} and {*L. lineoocellatum*/ *L. chloronoton*, *L. nigriplantare nigriplantare* and *L. n. polychroma*}, which were always selected in analyses of the 374 and 368 column datasets, and of {*L. infrapunctatum* and *La. guichenoti*}, which was not only selected in all analyses of the 368 column dataset but was also well supported in these analyses on the basis of both bootstrap and spectral analysis (signal-clash) values (some data not shown but see Figures 4.15 - 4.20 and Tables 4.12 - 4.14).

These consensus trees were constructed in a similar fashion to those for the *Cyclodina* taxa (Figure 4.13), with solid internal edges indicating groupings of taxa which were always selected in analyses of these taxa and these datasets and areas with dotted lines showing which relationships could not be resolved. In addition, edges which had 50% or greater bootstrap support in one or more of the trees being compared are indicated with asterisks (see also Tables 4.12-4.15).

Not all of the conflicting edges could be collapsed, however, without both obscuring relationships which occurred in most of the trees and implying relationships which were not actually observed. Therefore, taxa involved in local rearrangements (such as pairing or changing places with nearby taxa or moving onto neighbouring branches) for which the conflicting edges could not be unambiguously collapsed were placed as close to equidistant as possible to all of the positions in which they were observed, while major rearrangements (those which would necessitate the collapse of one or more edges for which there are otherwise no conflicts and which separate large groups of taxa) are indicated with arrows (see Figures 4.24 and 4.25 for examples). In addition, where possible, taxa in the dotted line regions were placed to indicate which local rearrangements are most likely (again, see Figures 4.24 and 4.25 for examples), but this was not always feasible. Moreover, as with the *Cyclodina* consensus trees (Figure 4.13), the external edges of taxa within the dotted line regions were also differentiated into solid and dotted lines to indicate which pairs are observed and which are not (see Figures 4.24 and 4.25).

Figure 4.24 Consensus tree for the 25 New Zealand skink taxa, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (374, 368 and 298 column datasets). The results of the 298 column dataset analyses were only used in the placement of *T. rohssii* (see text). Internal edges for which there are no conflicts amongst the set of trees for these datasets are indicated with solid lines. Areas for which the exact relationships of the taxa are not clear (that is, for which there are conflicting edges amongst the set of trees) are indicated with dotted lines. The taxa groupings given for these areas result from collapsing conflicting edges wherever possible (for example, *L. suteri* can be on either the *L. grande*/ *L. stenotis* branch or the *L. lineoocellatum*/ *L. chloronoton*/ *L. n. nigriplantare*/ *L. n. polychroma* branch), while taxa for which conflicting edges cannot be unambiguously collapsed and which are involved in local rearrangements only are placed as close to equidistant as possible to all of the positions in which they are observed. For example, *C. ornata* can pair with *C. macgregori* or *L. moco* or be separate and can be to the left or right of the *L. fallai*/ *C. alani* branch. Similarly, *L. moco* can pair with *C. ornata* or be separate and can be to the left of or on the *L. zelandicum*/ *C. aenea* (PK) branch. Major rearrangements are indicated with arrows (see text). In addition, where possible the taxa are placed to indicate their range of movement. For example, *C. aenea* (So) and *L. telfairi* can pair or be separate and can be to the left, to the right or between *L. zelandicum* and *C. aenea* (PK). Similarly, *L. zelandicum* and *C. aenea* (PK) can pair or be separate, but do not pair with either *C. aenea* (So) or *L. telfairi*. Therefore *L. zelandicum* and *C. aenea* (PK) are placed on one side of the tree and *C. aenea* (So) and *L. telfairi* on the other. Furthermore, the external edges of taxa which fall within the dotted regions are also differentiated into solid or dotted lines, to indicate which of the taxa in these regions are observed to pair and which are not. Taxa which fall within these regions and which have dotted external edges can pair with other taxa with dotted external edges in the same region; taxa which fall within these regions and which have solid external edges do not form pairs with other taxa in the region (unless they are shown as part of a pair). For example, *L. inconspicuum* and *L. maccanni* can pair, *T. rohssii* does not pair with *L. notosaurus*, *L. inconspicuum* or *L. maccanni*, while *L. notosaurus* can pair with *L. inconspicuum*. Edges with 50% or greater bootstrap support in one or more of the set of trees for these datasets are marked with asterisks. The symbol § indicates that this edge is constant apart from the major rearrangements indicated by the arrows. The tree is unrooted and is not drawn to scale.

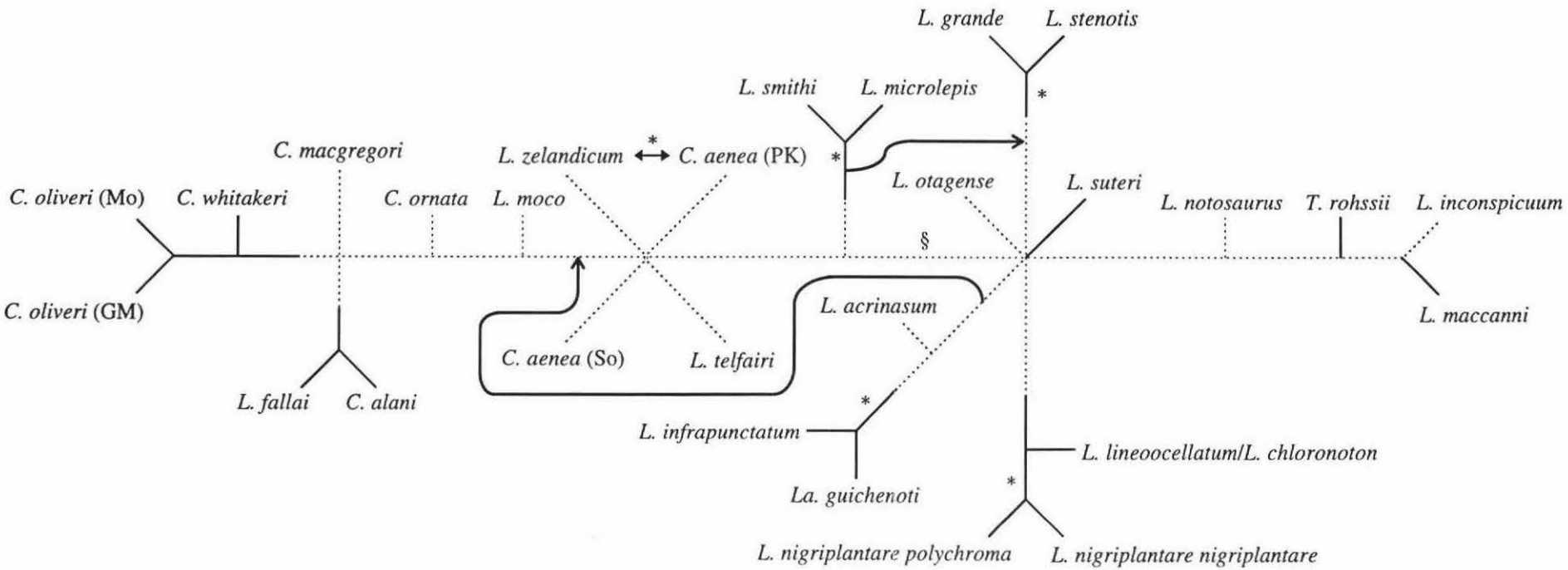
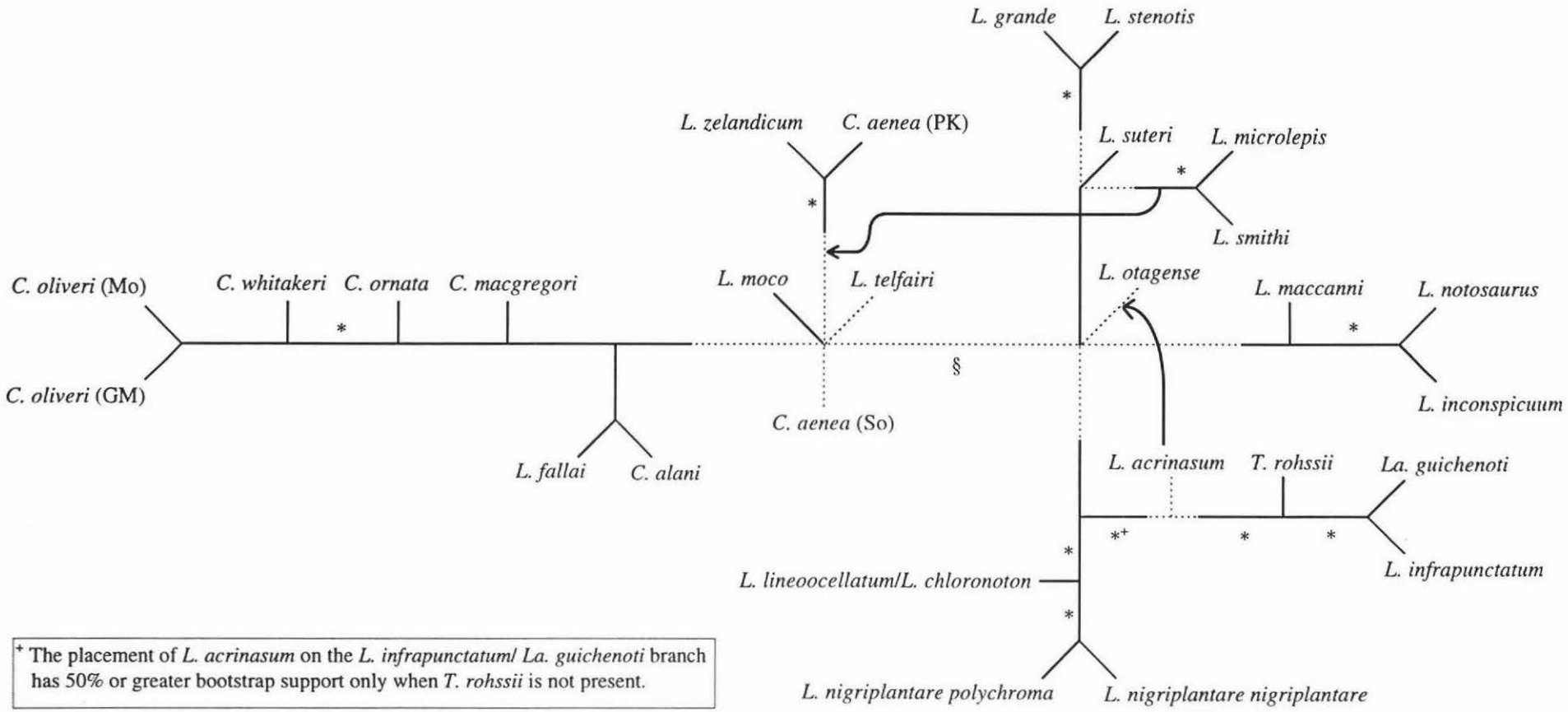


Figure 4.25 Consensus tree for the 25 New Zealand skink taxa, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* (386 and 387 column datasets). Internal edges for which there are no conflicts amongst the set of trees for these datasets are indicated with solid lines. Areas for which the exact relationships of the taxa are not clear are indicated with dotted lines; the taxa groupings given for these areas result from collapsing all conflicting edges except those involving major rearrangements (marked with arrows; see text). For example, *L. suteri* can be found on the *L. grande*/ *L. stenotis* branch or on the *L. microlepis*/ *L. smithi* branch. In addition, where possible the taxa in these regions are placed to indicate their range of movement. For example, *C. aenea* (So) can move to the left or to the right of *L. zelandicum*/ *C. aenea* (PK), *L. telfairi* can move to the right of or onto the *L. zelandicum*/ *C. aenea* (PK) branch, while *L. moco* can move to the left of or onto the *L. zelandicum*/ *C. aenea* (PK) branch. Similarly, *L. otagense* can move to the left or to the right of the *L. lineoocellatum*/ *L. chloronoton*/ *L. n. nigriplantare*/ *L. n. polychroma* branch, but does not move down onto this branch. Furthermore, the external edges of taxa which fall within the dotted regions are also differentiated into solid or dotted lines, to indicate which of the taxa in these regions are observed to pair and which are not. Single taxa which fall within these regions and which have solid external edges do not form pairs with other single taxa in the region; taxa with dotted external edges are observed to form pairs with neighbouring taxa with dotted external edges. For example, *L. moco* does not pair with either *C. aenea* (So) or *L. telfairi*, while *C. aenea* (So) and *L. telfairi* can pair. Edges with 50% or greater bootstrap support in one or more of the set of trees for these datasets are marked with asterisks. The symbol § indicates that this edge is constant apart from the rearrangement of *L. microlepis*/ *L. smithi* (as indicated by the arrow). The tree is unrooted and is not drawn to scale.



[+] The placement of *L. acrinasum* on the *L. infrapunctatum*/ *La. guichenoti* branch has 50% or greater bootstrap support only when *T. rohssii* is not present.

Phylogenetic Analysis of the New Zealand Skinks, *Leiolopisma telfairi*, *Lampropholis guichenoti* and *Tropidoscincus rohssii* - Conclusions.

As with the *Cyclodina* taxa, phylogenetic analysis of the New Zealand skinks (both with and without outgroups) did not produce a resolved phylogeny. However, partial resolution was achieved (see solid and asterisked internal edges in Figures 4.24 and 4.25). The groupings {*L. stenotis* and *L. grande*} (bipartition 768), {*L. microlepis* and *L. smithi*} (bipartition 147456), {*L. nigriplantare polychroma* and *L. n. nigriplantare*} (bipartition 2105344) and {*L. infrapunctatum* and *La. guichenoti*} (bipartition 71303168) were selected in all phylogenetic analyses of these taxa for both the 374 and 368 column and 386 and 387 column datasets. Moreover, these groupings had a reasonable degree of both bootstrap and spectral analysis (signal-clash value) support (spectral analysis values are not shown, but see Tables 4.12 and 4.13 for the ranks of these bipartitions; see also Figures 4.16, 4.18 and 4.21-4.23 and Tables 4.12, 4.13 and 4.15 for bootstrap values and ranks). These findings agree with those of Hickson (1993) and are also supported by allozyme analyses (see Hickson, 1993).

The grouping {*L. zelandicum* and *C. aenea* [PK]} (bipartition 262146) was selected in all analyses of these taxa in the 386 and 387 column datasets and in most analyses of the 374 and 368 column datasets, with a slightly lower level of both bootstrap and spectral analysis support than for the relationships already mentioned. Other groupings, such as {*C. oliveri* [Mo] and *C. oliveri* [GM]} (bipartition 96), {*C. oliveri* [Mo], *C. oliveri* [GM] and *C. whitakeri*} (bipartition 100), {*L. fallai* and *C. alani*} (bipartition 524296) and {*L. nigriplantare polychroma*, *L. n. nigriplantare* and *L. lineoocellatum*/ *L. chloronoton*} (bipartition 2107392) were also selected in all analyses of these taxa and these datasets, but with a relatively low level of bootstrap and spectral analysis support (again, see Figures 4.16, 4.18 and 4.21-4.23 and Tables 4.12, 4.13 and 4.15).

Overall, the two consensus trees (Figures 4.24 and 4.25) have very similar patterns of taxa grouping, with the main difference between the two being in the placement of *T. rohssii*. The consensus tree for the 386 and 387 column datasets (Figure 4.25) has a larger number of edges for which there are no conflicts (solid internal edges) than the one for the 374 and 368 column datasets (Figure 4.24). This is to be expected, considering the larger number of trees which were generated and compared for the 374 and 368 column datasets (spectral analysis as well as maximum parsimony and minimum evolution trees) and the fact that many of these trees were subtrees rather than 'complete' trees (with the accompanying possibility of edges which would not have been selected if all of the taxa could have been analysed together).

However, Figure 4.25 also contains a larger number of edges for which one or more of the set of trees had 50% or greater bootstrap support (edges with asterisks). This undoubtedly reflects the larger number of sites in the 386 and 387 column datasets, but may also indicate that the relationships in Figure 4.25 are more reliable than any alternative relationships in Figure 4.24 (similar to the findings from analyses of the different *Cyclodina* datasets). In particular, the position of *T. rohssii* in Figure 4.24 must be regarded with caution, placed as it was on the basis of analyses of the 298 column dataset (which has a still further reduced number of sites and analyses of which did not support many of the relationships consistently selected in analyses of the 374 and 368 column datasets).

Relationships of the Cyclodina Taxa.

The relationships of the *Cyclodina* taxa remain incompletely resolved after analysis of the complete skink dataset. Indeed, the results of these analyses (see Figures 4.24 and 4.25) are very similar to those from analyses of these taxa alone (see Figure 4.13 and below), but once again certain conclusions can be drawn. Again, all of the taxa form clearly defined lineages, including the two proposed new species (*C. oliveri* [Mo] and *C. aenea* [PK]). However, on the basis of these analyses, it is still unclear as to whether *C. oliveri* (Mo) warrants separate species or only subspecies status. As mentioned above, the grouping {*C. oliveri* [Mo] and *C. oliveri* [GM]} was selected in all analyses of these taxa in the 374, 368, 386 and 387 column datasets, but the level of spectral analysis (signal-clash value) and bootstrap support for this grouping was not high. The highest rank for this grouping in the spectral analyses was 23 out of 524288 possible bipartitions, while the highest bootstrap value was 47% (see ranks of bipartition 96 in Tables 4.12, 4.13 and 4.15).

On the other hand, the status of *C. aenea* (PK) as a separate species is further supported by these analyses. *C. aenea* (PK) and *C. aenea* (So) did not pair in any of the trees, instead pairing with and/or being separated by other taxa in all cases (see below and Figures 4.15-4.23 and Tables 4.12-4.15). The highest rank for the grouping of {*C. aenea* [PK] and *C. aenea* [So]} was 518231 out of 524288 possible bipartitions in the spectral analyses, while the highest bootstrap value for this grouping was <4.5% (see ranks of bipartition 3 in Tables 4.12-4.15).

As with analyses of the *Cyclodina* taxa alone (see Table 4.11 and Figure 4.13), the groupings {*C. oliveri* [Mo] and *C. oliveri* [GM]} and {*C. oliveri* [Mo], *C. oliveri* [GM] and *C. whitakeri*} are moderately well supported. As mentioned above, these groupings were always selected in analyses of these taxa in the 374, 368, 386 and 387 column datasets, but did not receive a high level of support on the basis of either bootstrap or spectral analysis (signal-clash) values in these analyses. However, the latter grouping did have 50% or greater bootstrap support in some of the trees from the 386 and 387 column datasets (see Figures 4.21-4.23 and Table 4.15).

The relationships of *C. macgregori* and *C. ornata* were constant in trees from analyses of the 386 and 387 column datasets (see Figure 4.25) and are very similar to those from analyses of the 387 and 382 column *Cyclodina* datasets (see Figure 4.13 - trees a and b). The two taxa are most closely related to each other and also to *C. oliveri* (Mo), *C. oliveri* (GM) and *C. whitakeri* in the case of *C. ornata* and to *C. alani*, *C. aenea* (So) and *C. aenea* (PK) (of the *Cyclodina* taxa) in the case of *C. macgregori*. However, the level of bootstrap support was low for these edges (see Figures 4.21-4.23 and Table 4.15).

The relationships of *C. macgregori* and *C. ornata* are more variable in Figure 4.24 (as in analyses of the 378 column *Cyclodina* dataset, see Figure 4.13 - tree c). The two taxa were closely related to each other in trees from analyses of the 374 column dataset, but *C. ornata* paired with *L. moco* (see below) in all but one of the 368 column dataset trees. In addition, the two taxa swapped places in most of the trees for these datasets - with *C. macgregori* tending to be more closely related to *C. oliveri* (Mo), *C. oliveri* (GM) and *C. whitakeri* and/or *C. alani*, while *C. ornata* was generally more closely related to *C. aenea* (So), *C. aenea* (PK) and sometimes *C. alani* (of the *Cyclodina* taxa). Bootstrap values were again low for these relationships, as were the spectral analysis values (see Figures 4.16 and 4.18 and Tables 4.12 and 4.13). As suggested above, the relationships shown in Figure 4.25 may be more reliable than those in Figure 4.24.

However, with the inclusion of the New Zealand *Leiolopisma* and the three possible outgroups in the analyses, the relationships of *C. alani*, *C. aenea* (So) and *C. aenea* (PK) become somewhat different from those in Figure 4.13 (see Figures 4.24 and 4.25). For while the *Cyclodina* taxa do group closely together (see Figures 4.24 and 4.25), they did not form a monophyletic group in any of the trees obtained in this study. The highest rank for the edge putting the *Cyclodina* taxa together was 1218 out of 524288 possible edges in the spectral analyses, while the highest bootstrap value was <4.5% (see ranks of bipartition 255 in Tables 4.12-4.15).

L. fallai fell within the group of *Cyclodina* taxa and paired with *C. alani* in all analyses of these taxa (see Figures 4.15-4.23 and Tables 4.12-4.15). However, as mentioned earlier, the {*L. fallai* and *C. alani*} grouping did not receive a high level of support from either spectral analysis or bootstrap values (the highest bootstrap value for this grouping being 37% - see Figure 4.18). Nor were any of the edges flanking this pair well supported (see Figures 4.16, 4.18, 4.20-4.23 and Tables 4.12-4.15). *L. zelandicum* and *C. aenea* (PK) formed a pair in most analyses of these taxa (again, see above), with some support from both spectral analysis and bootstrap values. The highest bootstrap value for this pair was 77% in the 386 and 387

column dataset analyses (see Figure 4.21); the highest bootstrap value in the 374 and 368 column dataset analyses was 51% (see Figure 4.16). In addition, this pair of taxa fell either on the boundary between the *Cyclodina* and *Leiolopisma* taxa or just outside the *Cyclodina* cluster in most of the trees from these analyses, but the spectral analysis and bootstrap support for these placements was low (see Figures 4.15-4.18, 4.21-4.23 and Tables 4.12, 4.13 and 4.15).

L. moco also appears to be closely related to the *Cyclodina* taxa, falling either within the group of *Cyclodina* taxa or on the boundary between *Cyclodina* and *Leiolopisma* in all trees from analyses of the 374, 368, 386 and 387 column datasets (see Figures 4.15-4.18, 4.21-23 and Tables 4.12, 4.13 and 4.15). However, the exact relationships of this taxon are not clear. In all but one of the trees from analyses of the 368 column dataset, *L. moco* and *C. ornata* paired (bipartition 16777344), but the grouping did not have a high level of either spectral analysis or bootstrap support (the highest bootstrap value was 23% - see Figure 4.18). Furthermore, this relationship was not observed in trees from the 374, 386 or 387 column datasets, and as mentioned above, relationships from analyses of the latter may be more reliable. Instead, *L. moco* separated *C. aenea* (So) and *C. aenea* (PK) in all 374, 386 and 387 column dataset trees (although again with only a low level of bootstrap and spectral analysis support).

In addition, *L. telfairi* and *C. aenea* (So) paired in many of the analyses of these taxa (bipartition 33554433), but this relationship received only a low level of spectral analysis and bootstrap support. The highest bootstrap value for this grouping was 25% in analyses of the 368 column dataset (see Figure 4.18); the highest bootstrap value in analyses of the 387 column datasets was 19% (tree not shown). These taxa also fell on or near the boundary between the *Cyclodina* and *Leiolopisma* taxa in most trees, but again all placements for these taxa had a very low level of spectral analysis and bootstrap support (see Figures 4.18, 4.22 and 4.23 and Tables 4.13 and 4.15).

These results are comparable to those of Hickson (1993), who, as mentioned previously, included one *Cyclodina* taxon in his analyses of the New Zealand *Leiolopisma* (*C. aenea* from Somes Island - see *C. aenea* [R] in Tables 4.1and 4.2). His analyses could not resolve the relationships of this taxon, but did suggest links between *C. aenea* and both *L. fallai* and *L. telfairi* (*L. moco* also grouped with *C. aenea* and *L. telfairi*). In addition, *L. telfairi* appeared to have links to *L. zelandicum* and *L. otagense*, while *L. fallai* was also observed to group with *L. microlepis*, *L. smithi* and *L. zelandicum*.

Internal and External Edgelengths.

In all of the skink phylogenies most of the internal edges are relatively short compared to the external edges (data for the external edges is not given, but see Figures 4.16, 4.18 and 4.20-4.23 in which the edgelengths of the trees are proportional to the number of changes along each edge and Tables 4.13-4.15 in which the lengths of the internal edges are given). This agrees with the findings for the *Cyclodina* dataset (in which all internal edges were short relative to the external edges) and again helps to explain why the skink phylogeny is so difficult to resolve while simultaneously indicating that many of the skink lineages are relatively old.

With regards to external edgelengths, *T. rohssii* has much the longest external edge, despite the number of sites which were absent in the *T. rohssii* sequence relative to the other sequences (see Figures 4.20 and 4.23). *L. otagense*, *L. telfairi*, *L. moco* and *L. maccanni* have the next longest edges in all trees, along with *L. infrapunctatum* when *La. guichenoti* is not present (see Figures 4.16, 4.18, and 4.20-4.23). *L. microlepis*, *C. oliveri* (GM), *C. oliveri* (Mo), *L. smithi*, *L. stenotis*, *La. guichenoti*, *C. whitakeri*, *L. n. nigriplantare*, *L. notosaurus*, *L. grande* and *C. ornata* consistently have the shortest external edges (in approximately this order, from shortest to longest; see Figures 4.16, 4.18, and 4.20-4.23).

Considering the length of the external edge leading to *T. rohssii*, any placement of this taxon must be regarded with caution, as it would not appear to be closely related to any of the other skink taxa examined here. Both of the positions selected for this taxon are near other taxa with long external edges (either *L. maccanni* or *L. infrapunctatum*; see below). Similarly, one of the other potential outgroups, *La. guichenoti*, also groups with *L. infrapunctatum* (again, see below). Another interesting feature is that two of the non-*Cyclodina* taxa which group with *Cyclodina* also appear to be two of the older lineages (*L. telfairi* and *L. moco*).

Biogeography of the New Zealand Skinks and the Origin(s) of Cyclodina.

The New Zealand skinks are divided into two broad groups in both Figure 4.24 and Figure 4.25 (separated by the edges marked with the symbol §). In addition, the taxa comprising each group are basically identical for the two trees. When the geographical distribution of these taxa is examined (see Table 1.2), it is found that one of these groups consists largely of the northern skinks (the *Cyclodina* taxa, *L. fallai* and *L. moco*; on the left side of Figures 4.24 and 4.25), while the second group contains the southern taxa plus *L. infrapunctatum*, *L. n. polychroma* and *L. lineoocellatum*/ *L. chloronoton* (which have a north/south distribution - that is, are found throughout New Zealand). Exceptions to this general pattern include *L. suteri* (a northern skink) which falls with the southern-north/south group and *L. zelandicum* (with a north/south distribution) which is placed with the northern skinks.

In addition, *L. microlepis* and *L. smithi* (both northern skinks) sometimes fall with the northern skinks and sometimes with the southern-north/south group. Similarly, *L. acrinasum* (southern), *L. infrapunctatum* and *La. guichenoti* move within the northern group in some trees from the 368 column dataset (see Figure 4.24), but this placement is not seen in trees from the 386 and 387 column datasets (see Figure 4.25) which may be more reliable. However, although the same basic groupings of taxa are observed in all trees from the 374, 368, 386 and 387 column datasets, the edges separating the two groups are not well supported in any of the trees (see Figures 4.16, 4.18 and 4.21-4.23 and Tables 4.12, 4.13 and 4.15).

With regards to the three possible outgroups, the Mauritian skink *L. telfairi* falls with the northern group - being placed on or near the *Cyclodina*/ *Leiolopisma* boundary in most trees and pairing with *C. aenea* (So) in many of the analyses (see above). The remaining two outgroups, however, fall within the southern-north/south group. The Australian skink *La. guichenoti* pairs with *L. infrapunctatum* in both Figure 4.24 and Figure 4.25, while the New Caledonian skink *T. rohssii* groups with three southern skinks (*L. maccanni*, *L. notosaurus* and *L. inconspicuum*) in Figure 4.24 and with *La. guichenoti* and *L. infrapunctatum* in Figure 4.25.

As mentioned earlier, the grouping {*L. infrapunctatum* and *La. guichenoti*} was selected in all analyses of these taxa for both the 368 and 387 column datasets (see Figures 4.17, 4.18, 4.22 and 4.23 and Tables 4.13 and 4.15), with a reasonable degree of both spectral analysis and bootstrap support. The highest bootstrap value for this grouping in the 368 column dataset analyses was 66% (see Figure 4.18), while analyses of the 387 column datasets gave a top bootstrap value of 84% (see Figure 4.22). Hickson (1993) also found that *La. guichenoti* paired most closely with *L. infrapunctatum* and with a reasonable degree of support.

The placement of *T. rohssii* with *L. maccanni* and *L. inconspicuum* (bipartition 135331840) received only a small amount of support in spectral analyses of the 298 column dataset (data not shown but see Table 4.14 for the ranks of this bipartition) while the highest bootstrap value for this grouping was only 20% (see Figure 4.20). *T. rohssii* grouped with *L. infrapunctatum* and *La. guichenoti* in both trees for the 387 column (28 taxa) dataset (bipartition 205520896), but received only 50% or less bootstrap support. As mentioned earlier, the latter of these relationships may be more likely, but both placements should be regarded with caution. Neither of these groupings agree with the findings of Hickson (1993) that *T. rohssii* was most strongly linked with *L. fallai*. However, Hickson (1993) did note that this did not appear to be a close phylogenetic relationship.

The fact that *Cyclodina* does not appear to be monophyletic with regards to the New Zealand *Leiolopisma* and also seems to have links with *L. telfairi* may suggest a Gondwanan origin for the genus, especially when considered in conjunction with the estimated divergence dates for these taxa (see earlier this chapter). However, due to the lack of resolution for the skink dataset, the possibilities that *Cyclodina* diverged from *Leiolopisma* within New Zealand or arrived independently after the breakup of Gondwanaland (see Chapter One) cannot be ruled out. The biogeographic distribution and the possible origin(s) of the *Cyclodina* taxa are discussed further in Chapter Five.

Comparison to the 12S rRNA Phylogenies of Hickson (1993).

The phylogenies presented here compare very well with those of Hickson (1993; parsimony, neighbor-joining and closest tree phylogenies), with basically the same groupings of taxa if not quite the same branching order for all groups. All but one or two of the differences involve *C. aenea*, *L. fallai*, *L. telfairi*, *L. zelandicum*, *L. moco* and *L. otagense* which, while closely related, do not form a separate branch in the trees of Hickson (1993). In the trees presented here, they do form part of a single group (shown to the left of Figures 4.24 and 4.25), with the exception of *L. otagense*. Differences in the placement of these taxa are to be expected with the addition of the remaining *Cyclodina* taxa into the analyses, the fact that Hickson (1993) found these *Leiolopisma* taxa to be the most closely related to *C. aenea* and to the taxa linked with *C. aenea* (see earlier and Hickson, 1993) and the low level of resolution for the edges involved. The remaining differences almost certainly result from the fact that Hickson (1993) used a lower number of bootstrap replicates (either 10 or 100) and a 382 column dataset for his analyses.

Hickson (1993)'s phylogenies were also only partially resolved. Those relationships which were well supported in his analyses were also consistently selected in all trees presented here, with a reasonable degree of support (see earlier). In addition, he also found that the relationships of the northern taxa were better resolved, while the relationships of the southern taxa were more complex and less well supported. The fact that most of the northern taxa form one group while the southern and most of the north/south taxa cluster together to form a second in Figures 4.24 and 4.25 (again, see earlier) would tend to support this finding (although none of the edges in the northern group have a high level of support).

Comparison to the Allozyme Analyses of Vos (1988).

Comparison of these phylogenies to those of Vos (1988) from allozyme data reveal that the relationships of *C. oliveri* (Mo), *C. oliveri* (GM), *C. ornata* and *C. whitakeri* are again similar in the two sets of trees (although *C. ornata* pairs with *C. oliveri* [GM] in the allozyme phylogenies). Similarly, both datasets suggest that *C. oliveri* (GM), *C. oliveri* (Mo), *C. whitakeri* and *C. ornata*

are the most recent *Cyclodina* lineages. However, there is one obvious difference - *Cyclodina* is monophyletic with regards to *Leiolopisma* in the allozyme trees, but does not form a monophyletic group in any of the sequence-based phylogenies. *L. telfairi* and *L. zelandicum* are placed close to *C. aenea* (So) and *C. aenea* (PK) in the one of the allozyme analyses, but {*L. telfairi* and *C. aenea* [So]} and {*L. zelandicum* and *C. aenea* [PK]} pairs were not observed. Further comparison is difficult however, as Vos (1988) only included three possible outgroups in her analyses of *Cyclodina* - *L. zelandicum*, *L. n. nigriplantare* and *L. telfairi* (see Figures 1.1 and 1.2) - and because neither sequences nor allozyme data produced a completely resolved phylogeny.

At present these analyses of Vos (1988) represent the only published skink phylogenies from allozyme data. Hickson (1993) compared the results of his analyses to the unpublished data of C. H. Daugherty and G. B. Patterson and found that while the sequence and allozyme information was not in complete disagreement, the genetic distances between the two datasets (pairwise D values for the allozyme data versus number of nucleotide differences in pairwise comparisons for the sequence data) were not well correlated overall (see Hickson, 1993, for possible explanations). Unpublished allozyme analyses also suggest that the current diversity of the New Zealand *Leiolopisma* resulted from divergence in New Zealand during the Miocene and that these taxa are not as closely related to *L. telfairi* and *La. guichenoti* as the analyses of 12S rRNA sequence data suggest (again, see Hickson, 1993). Allozyme analysis to date also provides some support for a monophyletic origin for *Cyclodina* with regards to the New Zealand *Leiolopisma* (G. B. Patterson, pers. comm.).

With this study, preliminary investigations of the New Zealand skinks using sequence data are complete. The individuals examined here and by Hickson (1993) form a subset of those investigated with allozyme data - thus once complete allozyme phylogenies are available, the results of analysis of the two datasets can be directly compared. This will facilitate the testing of conclusions drawn from each of the datasets, particularly when examined in conjunction with other information such as ecological data. Due to the low number of available allozyme loci (see Chapter One) and the lack of resolution among the sequence-based phylogenies, however, more sequence data is required before many of the long-standing questions about the origin(s) and evolution of the New Zealand skinks can be addressed with confidence.

Chapter Five: Discussion.

In this chapter the main conclusions of this study are summarised, the sequence phylogenies are examined with respect to morphological and allozyme data for *Cyclodina*, the possible origin(s) and the biogeography of the *Cyclodina* taxa are discussed and the molecular research required in future investigations of the skinks is outlined.

A unique DNA sequence was successfully obtained for each of the eight *Cyclodina* taxa under investigation using a phenol-chloroform extraction procedure, the polymerase chain reaction (PCR) and a direct double-stranded thermocycling-based sequencing method. These sequences are very similar to but nevertheless distinct from the 12S rRNA skink sequences of Hickson (1993), indicating that they do indeed represent *Cyclodina* 12S rRNA sequences rather than a contaminant. Furthermore, the exhaustive sequencing and checking procedure used makes it extremely unlikely that these sequences contain any sequencing or reading errors. The aligned *Cyclodina* sequences are presented in Tables 3.2 and 3.3, while an alignment of known skink sequences (Hickson, 1993, and this study) is presented in Table 4.2.

The *Cyclodina* sequences were fitted to the secondary structure models of Hickson (1993; domain III and part of domain II of vertebrate 12S rRNA) and Hickson *et al.* (1996; domain III of 12S rRNA in multicellular animals). Comparison of the *Cyclodina* sequences to a 'conserved sites' version of Hickson's 1993 model (see Figure 4.1) gives a further estimate of sequencing fidelity - as approximately less than one error per 1000 bases. The *Cyclodina* secondary structure is fundamentally similar to the vertebrate (Hickson, 1993) and multicellular animal (Hickson *et al.*, 1996) models and is almost identical to that of *Leiolopisma*. In addition, the patterns of variability among the *Cyclodina* and *Leiolopisma* sequences are very similar and are in good agreement with those observed by Hickson (1993) for other vertebrates. Overall, this would tend to suggest that the *Cyclodina* sequences, like the skink sequences investigated by Hickson (1993), do not have a significantly different pattern of evolution from other vertebrate 12S rRNA sequences.

Phylogenetic Analysis of *Cyclodina* and Comparison to Other Datasets.

Phylogenetic analyses (closest tree, maximum parsimony, minimum evolution and maximum likelihood) based on this sequence data could not fully resolve the *Cyclodina* taxonomy. The separate taxonomic status of the six recognised (Hardy, 1977) and the two proposed (Vos, 1988) *Cyclodina* species is well supported however, with all eight taxa forming genetically distinct lineages. Moreover, *C. aenea* (PK) is clearly distinct from *C. aenea* (So) - the two do

not group together in some analyses and the external edgelengths of these taxa represent two of the three longest in the *Cyclodina* trees. *C. oliveri* (Mo) is definitely distinct from *C. oliveri* (GM) but cannot be excluded from subspecific status. The results of these analyses are presented in several different formats (see Figures 4.8-4.13, Table 4.11 and Appendix Two, Figures 2.1-2.4).

Comparison of the trees derived from sequence data to those of Vos (1988) from allozyme data (see Figures 1.1 and 1.2) reveals some differences (the placement of *C. ornata* and the relationship of *C. alani* and *C. macgregori*) and some similarities (particularly in the respective age of the *Cyclodina* lineages). Hardy (1977) suggested close phylogenetic relationships between *C. ornata* and *C. aenea* and between *C. oliveri*, *C. alani* and *C. whitakeri* based on his analysis of lactate dehydrogenase (LDH) isozymes. On morphological grounds, *C. whitakeri*, *C. oliveri* (Mo) and the other *C. oliveri* populations are similar, as are *C. alani* and *C. macgregori* (Vos, 1988). *C. aenea*, *C. aenea* (PK) and *C. ornata* also show morphological similarities (Vos, 1988). Some of these groupings are seen in the sequence data phylogenies and some are not. With the exception of the fairly well supported {*C. oliveri* (Mo) and *C. oliveri* (GM)} and {*C. whitakeri*, *C. oliveri* (Mo) and *C. oliveri* (GM)} groupings, the general lack of resolution in the *Cyclodina* sequence data phylogenies means that most groupings proposed on the basis of other datasets cannot yet be conclusively confirmed or rejected.

The combined skink dataset (New Zealand *Leiolopisma* and *Cyclodina*) was also analysed phylogenetically (closest tree, maximum parsimony and minimum evolution) both with and without the possible outgroups *L. telfairi*, *La. guichenoti* and *T. rohssii*. As above, the results of these analyses are presented in a number of formats (see Figures 4.15-4.25 and Tables 4.12-4.15). The resulting trees are comparable to those of Hickson (1993), with many of the same taxa groupings (see Chapter Four). Although the *Cyclodina* taxonomy still could not be fully resolved, again certain conclusions can be drawn.

Firstly, the *Cyclodina* taxa, while still forming clearly distinct lineages, do form a closely related group. However, the boundaries of this group are not completely clear. *L. fallai* pairs with *C. alani* and falls within the *Cyclodina* group in all trees (however, the highest bootstrap value for this pair is only 37%, while the level of support for the placement of these taxa is also low). The proposed new species *C. aenea* (PK) pairs with *L. zelandicum* in most analyses, with a reasonable level of both spectral analysis (signal-clash value) and bootstrap support (the highest bootstrap value for this pair is 77%). These taxa fall either on the *Cyclodina/ Leiolopisma* boundary or just outside the *Cyclodina* group in most trees (although the level of spectral analysis and bootstrap support for these placements is low). In addition, *L. moco* appears to fall either within the *Cyclodina* grouping or on the *Cyclodina/Leiolopisma* boundary in all trees (but with a very low level of both spectral analysis and bootstrap support for these placements).

Thus, unlike the allozyme trees of Vos (1988), the results of these analyses do not support a monophyletic origin of *Cyclodina* with regards to *Leiolopisma* under the current taxonomic classification. However, the low level of support for most of these placements means that this possibility cannot be entirely ruled out. The allozyme trees of Vos (1988) are currently the only published allozyme phylogenies, but unpublished allozyme data for the New Zealand skinks also provides some support for a monophyletic origin for *Cyclodina* (G. B. Patterson, pers. comm.).

Apart from these differences, the relationships of the *Cyclodina* taxa are very similar to those in trees from analyses of the *Cyclodina* taxa alone. The proposal that *C. aenea* (PK) is a distinct species (Vos, 1988) gains even stronger support from the combined dataset phylogenies, *C. aenea* (PK) being separated from *C. aenea* (So) by other taxa and/or pairing with *L. zelandicum* in all trees. Interestingly, *L. zelandicum* from Maud Island (geographically very close to the Outer Chetwode islands, from which Hickson, 1993's *L. zelandicum* sample came) is very similar to *C. aenea* in both morphology and habitat (P. Anderson, pers. comm.; T. H. Whitaker, pers. comm.; T. Jewell, pers. comm.). *L. zelandicum* did not group with *C. aenea* (PK) in the allozyme trees of Vos (1988), but it is possible that this is due to the inclusion of only three non-*Cyclodina* taxa in her analyses.

C. oliveri (Mo) again forms a distinct lineage, although not to the same extent as *C. aenea* (PK), and the exact taxonomic status of this population remains unclear. The differences between *C. oliveri* populations (off the north-western shore of the North Island) have been suggested to represent a cline (D. R. Towns, pers. comm.) and further investigation of these populations is required - particularly of the *C. oliveri* populations from Little Barrier and the Hen and Chicken Islands which are morphologically very similar to *C. oliveri* (Mo) (Vos, 1988). *C. aenea* (PK) and *C. oliveri* (Mo) are both now included on the latest edition of the priority list for the conservation of threatened species (SRARNZ Notes, 1995b).

Finally, while the currently available molecular and morphological data is not sufficient to produce a fully resolved taxonomy, the closer relationships among these skinks could be resolved and general groupings of taxa could be identified based on these analyses of 12S rRNA sequence data (see Figures 4.24 and 4.25). Complete resolution of the troubled skink taxonomy requires longer sequences from additional genes.

Lack of Resolution in the Skink Dataset and 'Rapid' Divergence of the Skink Taxa.

Hickson (1993) found that the lack of phylogenetic resolution from the skink sequences he analysed was not improved by differentially weighting variable sites (through the removal of constant columns), transversion sites or paired sites. Nor did it appear to be the result of too few variable sites in the skink dataset (similar numbers of variable sites occur in the ratite 12S

rRNA dataset of Cooper *et al.*, 1992, yet a resolved tree was obtained for these taxa; see Hickson, 1993). Both this study and the study of Hickson (1993) indicate that the type of phylogenetic analysis does not make a significant difference to the level of resolution of the skink phylogeny. Examination of the expected transversion/transition (Tv/Ts) ratio at saturation (R_∞) for both *Cyclodina* and *Leiolopisma* using equation 11 from Holmquist (1983) suggests that these sequences are unsaturated for transitions and therefore phylogenies obtained from them will not be unreliable for this reason. Furthermore, as mentioned above, the skink 12S rRNA sequences do not appear to have a significantly different pattern of molecular evolution (mechanism of change) from those of other vertebrates.

Analysis of the bovid dataset of Kraus and Miyamoto (1991) and Allard *et al.* (1992) produced a similarly unresolved phylogeny (see Hickson, 1993). This appears to be the result of rapid divergence of the bovids over a relatively short period (approximately five million years; see Allard *et al.*, 1992). Hickson (1993) analysed simulated sequence data derived using a star-tree model (as opposed to the normal binary tree model) and found that these analyses gave very similar results to those for the skinks. Based on these findings, he suggested that the New Zealand *Leiolopisma* 'rapidly' diverged from each other, and that the inability to resolve the *Leiolopisma* phylogeny was due to the assumption of an inappropriate model of evolution for the skink sequences. The similarities in the characteristics and behaviour of the *Cyclodina* and *Leiolopisma* datasets suggest that this is also the reason for the lack of resolution in the *Cyclodina* phylogenies. Sequence data from more conservative genes may be required to resolve the skink taxonomy, as in the case of the Pecoran bovids (Kraus and Miyamoto, 1991).

Nevertheless, 'rapid' divergence in the sense that it is used here does not necessarily imply recent divergence - it is likely that diversification of the skinks occurred early and that the resulting lineages have been steadily and separately evolving ever since. Support for this theory is provided by the relative length of the external edges when compared to the internal edges (observed both in the *Cyclodina* trees and in those from the combined datasets; see also Hickson, 1993) and by the estimates of early divergence times among the New Zealand skinks.

Estimated Divergence Times among the Skink Taxa.

Bull and Whitaker (1975) and Hardy (1977) suggested that *Cyclodina* arrived in New Zealand between the late Pliocene (approximately five million years ago [MYA]) and the early Pleistocene (up to two and a half MYA). Vos (1988) estimated the time of divergence between *Cyclodina* and *Leiolopisma* to be in the region of 19 MYA (Miocene) based on her allozyme studies. Hickson (1993) only investigated one *Cyclodina* sequence (*C. aenea* from Somes Island), and did not give a specific estimate of divergence time for this lineage, although he did propose a possible Gondwanan origin for the New Zealand *Leiolopisma* (see Chapter One).

Estimates of divergence times were made from the number of total differences in pairwise comparisons between the skink taxa (these numbers do not include the differences resulting from indels) using a ratite rate estimate of 0.5 changes/million years (see Chapter Four). These estimates indicate that the *Cyclodina* taxa are likely to be much older than previously thought (see Table 4.10). The two *C. oliveri* taxa (the closest *Cyclodina* pair based on number of pairwise differences) alone are estimated to have diverged approximately 10 million years ago (MYA), while an estimated divergence date of 42 MYA was obtained for the two *C. aenea* taxa. Interestingly, the estimated divergence date for *C. aenea* (PK) and *L. zelandicum* is only 30 MYA, as is the estimated divergence time for *L. fallai* and *C. alani*. The estimated divergence date for *L. telfairi* and *C. aenea* (So) (see below) is also 42 MYA.

As the skink phylogeny is not well resolved, it is difficult to estimate a divergence time for *Cyclodina* and *Leiolopisma* (assuming for the moment that the hypothesis of *Cyclodina* diverging from *Leiolopisma* within New Zealand is correct). At the latest it would appear to have been 28 MYA (this is the estimated divergence date for *L. fallai-C. oliveri* [Mo], the *Cyclodina-Leiolopisma* pair with the smallest number of pairwise differences). At the earliest, based on these estimates, it could have been 60 MYA (the estimated divergence date for the most dissimilar *Leiolopisma-Cyclodina* pairs: *L. otagense-C. alani*, *L. otagense-C. macgregori* and *L. n. polychroma-C. macgregori*). However, as these divergence times were calculated using a non-reptilian rate estimate, they may well be underestimates (see Chapter Four) and *Cyclodina* may even be Gondwanan in origin. A reliable reptile rate estimate is required to more accurately determine the *Cyclodina* (and *Leiolopisma*) divergence dates.

Origin(s) of *Cyclodina*.

As mentioned above, *L. fallai* and *L. moco* (both northern skinks) and *L. zelandicum* (found in both the North and South Islands) group with the *Cyclodina* taxa in most or all trees. One of the three outgroups investigated in this study, *L. telfairi* (the *Leiolopisma* type species, from Mauritius) also falls with this group. *L. telfairi* pairs with *C. aenea* (So) in many of the analyses of these taxa, although the grouping did not receive a high level of either spectral analysis or bootstrap support (the highest bootstrap value for this grouping is 25%). The two taxa fall on or near the boundary between *Cyclodina* and *Leiolopisma* in most trees (again these placements had a very low level of spectral analysis and bootstrap support).

With respect to the three hypotheses for *Cyclodina* origins given in Chapter One, if the placings of *L. telfairi*, *L. fallai*, *L. zelandicum* and *L. moco* are accurate, *Cyclodina* is as closely related to at least one overseas skink as to the New Zealand *Leiolopisma*. This may suggest that *Cyclodina* has Gondwanan origins, as has been proposed for the New Zealand *Leiolopisma*

(Hickson, 1993). The estimated times of divergence (see above) provide some support for this theory, as do the relatively long external edges of several of the *Cyclodina* taxa when compared to the internal edges of the trees and the possible close relationship between *Cyclodina* and *L. alazon* from Fiji (Zug, 1985; see Chapter One).

Examination of the occurrence of indels among the skink sequences also provides some support for these patterns. The occurrence of only one indel in the New Zealand *Leiolopisma* dataset compared to the relatively higher number among *Cyclodina* taxa and in the *L. telfairi* sequence may reflect a closer relationship between *L. telfairi* and *Cyclodina*. In addition, *C. aenea* (So), *C. aenea* (PK), *C. ornata*, *L. maccanni* and *L. telfairi* all have deletions in the same region (the loop of helix 26; see Table 4.2 and Figures 4.2 and 4.3) and this may indicate closer relationships between these taxa. The fact that many of the indels fall in the S26 region, which appears to be one of the most variable regions among both the skinks and the vertebrates (see Table 4.5 and Hickson, 1993), and the small sample size of the *Cyclodina* and non-New Zealand skink datasets means that all such conclusions should be treated tentatively however.

All in all, the remaining hypotheses - that *Cyclodina* diverged from *Leiolopisma* within New Zealand or arrived independently of *Leiolopisma* after the break-up of Gondwanaland (see Chapter One) - cannot be conclusively eliminated. The skink taxonomy is not well resolved and the placements of both the *Leiolopisma* taxa and *L. telfairi* are not strongly supported (although the *L. zelandicum*/ *C. aenea* [PK] grouping has a reasonable degree of both spectral analysis and bootstrap support). In addition, as mentioned, allozyme analysis to date provides some support for *Cyclodina* diverging from *Leiolopisma* within New Zealand (Vos, 1988; G. B. Patterson, pers. comm.). Longer sequences from both the New Zealand taxa and more overseas skinks are required to test the three hypotheses.

Biogeography of the New Zealand *Cyclodina*.

If *Cyclodina* has been in New Zealand since the break-up of Gondwana (approximately 80 MYA), the absence of *Cyclodina* taxa from the South Island requires explanation. The most obvious explanation is one of ecology. Bull and Whitaker (1975) noted that if the '*Sphenomorphus*' species (now *Cyclodina*) were in New Zealand during the last glaciation, their ecological requirements would have restricted them to the north of the North Island, with subsequent dispersal south interrupted by Cook Strait.

Current distribution patterns of these taxa (with all but *C. aenea* and *C. ornata* largely or completely restricted to offshore islands) appear to be fairly well explained as the result of predation, particularly by rodents (for example, see Bull and Whitaker, 1975; Hardy, 1977;

Towns *et al.,* 1985; Worthy, 1987c; Towns, 1991). However, more information on the origin(s) of the *Cyclodina* taxa and more accurate estimates of divergence times would aid in establishing the effects of these and any other factors (such as glaciations and 'drownings').

Status of the Genus *Leiolopisma* and Related Skinks.

As mentioned in Chapter One, the taxonomic status of species currently and formerly assigned to the genus *Leiolopisma* is at present unclear. In addition to the reclassification of the *Leiolopisma* species from Australia (Hutchinson *et al.,* 1990) and Lord Howe/Norfolk Island (Cogger, 1986; Cogger, 1992 - not sighted), the New Zealand *Leiolopisma* have just been reclassified as *Oligosoma* (Patterson and Daugherty, 1995; see also SRARNZ Notes 1996). This nomenclature was initially suggested by Hutchinson *et al.* (1990), based on immunological data, and was subsequently followed by Hutchinson and Donellan (1993) and Hutchinson (1993). Immunological data also suggests that the New Caledonian skinks are equally distant from the Australian and New Zealand skinks and from *L. telfairi* (Hutchinson *et al.,* 1990).

Analysis of the sequence data available to date suggests that members of the New Zealand skink species are as similar to *L. telfairi* or *La. guichenoti* as to each other (Hickson, 1993, and this study). However, further molecular research is required before any meaningful generic revision can be achieved. Longer sequences from a larger and more diverse set of skinks are essential to test the proposed taxonomy.

Future Molecular Work.

This study of *Cyclodina* completes the preliminary investigation of the New Zealand skinks using 12S rRNA sequence data commenced by Hickson (1993). Once complete allozyme phylogenies are available, the two sets of phylogenies can be compared and the conclusions drawn on the basis of sequence data can be tested. However, some New Zealand taxa have not yet been examined using DNA sequences and additional sequence data is also required due to the lack of resolution among phylogenies presented in this study and in Hickson (1993) and the low number of available allozyme loci for the skinks (see Chapter One).

Sequence data is required for the New Zealand *Leiolopisma* species which have not yet been examined at the DNA level - *L. striatum*, *L. homalonotum*, *L. waimatense*, the newly identified southern *Leiolopisma* taxa, three potential new species from the West Coast of the South Island and several other possible species (see Table 1.2 and SRARNZ Notes 1995b and 1996). The relationships of *L. homalonotum* are of particular interest as this is another large northern skink species, like *L. fallai*, with a marking similar to the 'teardrop-shaped' mark under the eye characteristic of *Cyclodina* (except for *C. aenea*) (Hardy, 1977). More members of the South Island *L. lineoocellatum*/ *L. chloronoton* complex also need to be examined.

For *Cyclodina*, DNA sequence information is also required for populations which Hardy (1977) noted as being morphologically distinct from other members of the species to which they were assigned: the *C. oliveri* and *C. ornata* populations from the Poor Knights Islands, the *C. ornata* population from the Three Kings Islands and the *C. macgregori* population from Sail Rock, in addition to the *C. oliveri* populations from Little Barrier and the Hen and Chicken Islands (as mentioned earlier). Sequence data is also required for more overseas skinks. *L. alazon* from Fiji and *L. lichenigerum* from Lord Howe/Norfolk Islands are of particular interest in investigating the origins of *Cyclodina* (see Chapter One).

As yet, only sequences from the 12S rRNA gene (and some cytochrome *b* sequences; Hickson, 1993) have been obtained for the skinks. Longer sequences from additional genes are required to both improve the reliability of trees based on sequence data and to extend the range of phylogenetic relationships which can be examined (at the population level as well as the generic and species levels).

In order to obtain longer sequences from different genes more easily, it would help to have a complete mitochondrial sequence for the skinks. It is hoped to have a complete sequence for the skink mitochondrial genome (approximately 17, 000 bp) within the next one-two years. This will make it possible to design primers for comparative studies involving 1000-2000 nucleotides. It will also allow the examination of genes such as cytochrome oxidase I and cytochrome *b* - the more conserved first and second codon positions of which may be suitable for resolving older divergences, while the third codon positions may be useful for studying closer relationships - and of the D-loop region, thus facilitating population and detailed biogeographic studies.

Nuclear sequence data is also required to test the phylogenies generated using mitochondrial DNA (for example, see Palumbi and Baker, 1994). Recent advances, such as the development of universal primers spanning intron-exon boundaries in a wide range of animal groups (Palumbi and Baker, 1994), are now making the use of nuclear sequences feasible.

With further sequence data, in conjunction with other molecular information and morphological and ecological data, resolution of the difficult skink taxonomy should finally be possible. This should aid in establishing the origin(s) of the New Zealand skinks, in investigating the obviously complex biological and ecological histories of these taxa and in better understanding the processes of molecular evolution in general. The use of these combined datasets should also aid in the identification and conservation of unique populations.

References.

Allard, M. W., Miyamoto, M. M., Jarecki, L., Kraus, F. and Tennant, M. R. (1992). DNA systematics and evolution of the artiodactyl family Bovidae. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 3972-3976.

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R., and Young, I. G. (1981). Sequence and organisation of the human mitochondrial genome. *Nature* **290**: 457-465.

Anonymous (1992). fmol™ *DNA Sequencing System Technical Manual*, Promega Corporation, Madison, WI.

Árnason, Ú. and Gullberg, A. (1994). Relationship of baleen whales established by cytochrome *b* gene sequence comparison. *Nature* **367**: 726-728.

Arnheim, N. and Erlich, H. (1992). Polymerase chain reaction strategy. *Annual Review of Biochemistry* **61**: 131-156.

Arnheim, N., White, T. and Rainey, W. E. (1990). Application of PCR: Organismal and population biology. *BioScience* **40**: 174-182.

Avise, J. C., Bowen, B. W., Lamb, T., Meylan, A. B. and Bermingham, E. (1992). Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in Testudines. *Molecular Biology and Evolution* **9**: 457-473.

Baker, C. S. and Palumbi, S. R. (1994). Which whales are hunted? A molecular genetic approach to monitoring whaling. *Science* **265**: 1538-1539.

Barwick, R. E. (1959). The life history of the common New Zealand skink *Leiolopisma zelandica* (Gray, 1843). *Transactions of the Royal Society of New Zealand* **86**: 331-380.

Bauer, A. M. (1993). African-South American relationships: A perspective from the Reptilia. In: Goldblatt, P. (ed.), *Biological relationships between Africa and South America*. Yale University. Pp. 244-288.

Bauer, A. M. and Vindum, J. V. (1990). A checklist and key to the herpetofauna of New Caledonia, with remarks on biogeography. *Proceedings of the California Academy of Sciences* **47** (2): 17-45.

Bowen, B. W., Meylan, A. B. and Avise, J. C. (1991). Evolutionary distinctiveness of the endangered Kemp's ridley sea turtle. *Nature* **352**: 709-711.

Britten, R. J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**: 1393-1398.

Bull, P. C. and Whitaker, A. H. (1975). The amphibians, reptiles, birds and mammals. In: Kuschel, G. (ed.), *Biogeography and ecology in New Zealand*. Dr. W. Junk, b. v. Publishers, The Hague. Pp. 231-276.

Cann, R. L., Stoneking, M. and Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature* **325**: 31-36.

Carroll, R. L. (1969). Origin of reptiles. In: Gans, C. (ed.), *Biology of reptilia. Volume I. Morphology A.* Academic Press. Pp. 1-44.

Casanova, J.-L., Pannetier, C., Jaulin, C. and Kourilsky, P. (1990). Optimal conditions for directly sequencing double-stranded PCR products with Sequenase. *Nucleic Acids Research* **18** (13): 4028.

Charleston, M. A., Hendy, M. D. and Penny, D. (1994). The effects of sequence length, tree topology, and number of taxa on the performance of phylogenetic methods. *Journal of Computational Biology* **1** (2): 133-151.

Cogger, H. G. (1986). *Reptiles and amphibians of Australia.* Fourth edition. Reed Books Pty Ltd., New South Wales. xxi + 682 pages.

Cogger, H. G. (1992). *Reptiles and amphibians of Australia.* Fifth edition. Cornstock/Cornell, Ithaca, New York.

Cooper, A. C. and Cooper, R. A. (1995). The Oligocene bottleneck and the New Zealand biota: genetic record of a past environmental crisis. *Proceedings of the Royal Society of London* **261**: 293-302.

Cooper, A., Mourer-Chauviré, C., Chambers, G. K., von Haeseler, A., Wilson, A. C. and Pääbo, S. (1992). Independent origins of New Zealand moas and kiwis. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 8741-8744.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life.* John Murray, London.

Daugherty, C. H., Cree, A., Hay, J. M. and Thompson, M. B. (1990a). Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature* **347**: 177-179.

Daugherty, C. H., Patterson, G. B., Thorn, C. J. and French, D. C. (1990b). Differentiation of the members of the New Zealand *Leiolopisma nigriplantare* species complex (Lacertilia: Scincidae). *Herpetological monographs* **4**: 61-76.

Daugherty, C. H., Gibbs, G. W. and Hitchmough, R. A. (1993). Mega-island or micro-continent? New Zealand and its fauna. *Trends in Ecology and Evolution* **8** (12): 437-442.

Daugherty, C. H., Patterson, G. B. and Hitchmough, R. A. (1994). Taxonomic and conservation review of the New Zealand herpetofauna. *New Zealand Journal of Zoology* **21**: 317-323.

DeSalle, R., Freedman, T., Prager, E. M. and Wilson, A. C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila. Journal of Molecular Evolution* **26**: 157-164.

Dixon, M. T. and Hillis, D. M. (1993). Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution* **10**: 256-267.

Easteal, S. (1990). The pattern of mammalian evolution and the relative rate of molecular evolution. *Genetics* **124**: 165-173.

Estes, R. and Price, L. I. (1973). Iguanid lizard from the upper cretaceous of Brazil. *Science* **180**: 748-751.

Faith, D. P. (1990). Chance marsupial relationships. *Nature* **345**: 393-394.

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

Felsenstein, J. (1988). Phylogenies from molecular sequences: Inference and reliability. *Annual review of Genetics* **22**: 521-565.

Felsenstein, J. (1993). *Phylip (Phylogeny Inference Package) Version 3.5*. Computer software package and manual. Distributed by the author, Department of Genetics, University of Washington, Seattle, Washington.

Fleming, C. A. (1958). Darwinism in New Zealand: Some examples, influences and developments. *Transcripts of the Proceedings of the Royal Society of New Zealand* **86**: 65-86.

Gill, B. J. (1976). Aspects of the ecology, morphology, and taxonomy of two skinks (Reptilia: Lacertilia) in the coastal Manawatu area of New Zealand. *New Zealand Journal of Zoology* **3**: 141-157.

Gillespie, J. H. (1986). Rates of molecular evolution. *Annual Review of Ecology and Systematics* **17**: 637-665.

Gittleman, J. L. and Pimm, S. L. (1991). Crying wolf in North America. *Nature* **351**: 524-525.

Goddard, K., Caccone, A. and Powell, J. R. (1990). Evolutionary implications of DNA divergence in the *Drosophila obscura* group. *Evolution* **44**: 1656-1670.

Gray, J. E. (1843). Descriptions of the reptiles and amphibia hitherto observed in New Zealand. In: Dieffenbach, E. (ed.), *Travels in New Zealand with contributions to the geography, geology, botany, and natural history of that country (Volume Two)*. Murray, London. Pp. 202-206.

Greer, A. E. (1970). A subfamilial classification of scincid lizards. *Bulletin Museum of Comparative Zoology* **139** (3): 151-184.

Greer, A. E. (1974). The generic relationships of the scincid lizard genus *Leiolopisma* and its relatives. *Australian Journal of Zoology, Supplementary Series* **31**: 1-67.

Gutell, R. R., Weiser, B., Woese, C. R. and Noller, H. F. (1985). Comparative anatomy of 16-S-like ribosomal RNA. *Progress in Nucleic Acids Research* **32**: 155-216.

Gyllensten, U. B. and Erlich, H. A. (1988). Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 7652-7656.

Hardy, G. S. (1977). The New Zealand Scincidae (Reptilia: Lacertilia); a taxonomic and zoogeographic study. *New Zealand Journal of Zoology* **4**: 221-325.

Hardy, G. S. (1979). The karyotypes of two scincid lizards and their bearing on relationships in genus *Leiolopisma* and its relatives (Scincidae: Lygosominae). *New Zealand Journal of Zoology* **6**: 609-612.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology and Evolution* **4** (1): 6-11.

Hasegawa, M. and Hashimoto, T. (1993). Ribosomal RNA trees misleading? *Nature* **361**: 23.

Hedges, S. B., Bezy, R. L. and Maxson, L.R. (1991). Phylogenetic relationships and biogeography of xantusiid lizards, inferred from mitochondrial DNA sequences. *Molecular Biology and Evolution* **8**: 767-780.

Hendy, M. D. and Penny, D. (1982). Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences* **59**: 277-290.

Hendy, M. D. and Penny, D. (1989). A framework for the quantitative study of evolutionary trees. *Systematic Zoology* **38**: 297-309.

Hendy, M. D. and Penny, D. (1993). Spectral analysis of phylogenetic data. *Journal of Classification* **10**: 5-24.

Hickson, R. E. (1993). *Evolutionary tails from the South Pacific: Being a wondrous and exciting account of investigations into the histories of New Zealand skinks, and in which it is shown that things are not what they seem.* Ph.D. thesis, Massey University, Palmerston North. x + 121 pages.

Hickson, R. E., Penny, D. and Scott, D. B. (1992). Molecular systematics and evolution in New Zealand: applications to cryptic skink species. *New Zealand Journal of Zoology* **19**: 33-44.

Hickson, R. E., Simon, C., Cooper, A., Spicer, G. S., Sullivan, J. and Penny, D. (1996). Conserved sequence motifs, alignment and secondary structure for the third domain of animal 12S rRNA. *Molecular Biology and Evolution* **13** (1):150-169.

Higham, T. (1995). Leaping lizards. *Forest and Bird* **276**: 14-20.

Hillis, D. M. (1987). Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics* **18**: 23-42.

Hillis, D. M. and Moritz, C. (1990). An overview of applications of molecular systematics. In: Hillis, D. M. and Moritz, C. (eds.), *Molecular Systematics*. Sinauer Associates, Inc., Sunderland, Massachusetts. Pp. 502-515.

Hillis, D. M., Mable, B. K. and Moritz, C. (1996). Applications of molecular systematics: the state of the field and a look to the future. In: Hillis, D. M., Moritz, C. and Mable, B. K. (eds.), *Molecular Systematics*. Second edition. Sinauer Associates, Inc., Sunderland, Massachusetts. Pp. 515-543.

Holmes, E. C., Zhang, L. Q., Simmonds, P., Rogers, A. S. and Brown, A. J. L. (1993a). Molecular investigation of human immunodeficiency virus (HIV) infection in a patient of an HIV-infected surgeon. *The Journal of Infectious Diseases* **167**: 1411-1414.

Holmes, E. C., Brown, A. J. L. and Simmonds, P. (1993b). Sequence data as evidence. *Nature* **364**: 766.

Holmquist, R. (1983). Transitions and transversions in evolutionary descent: an approach to understanding. *Journal of Molecular Evolution* **19**: 134-144.

Hudson, B. (1994). *Reptiles and amphibians in New Zealand: Handbook for species identification.* Print Media Specialists, Auckland. 50 pages.

Hutchinson, M. N. (1993). Family Scincidae. In: Glasby, C. J., Ross, G. J. B., Beesley, P. L. (eds.), *Fauna of Australia. Volume 2A: Amphibia and reptilia.* Australian Government Publishing Service, Canberra. Pp. 261-279.

Hutchinson, M. N. and Donellan, S. C. (1993). Biogeography and phylogeny of the Squamata. In: Glasby, C. J., Ross, G. J. B., Beesley, P. L. (eds.), *Fauna of Australia. Volume 2A: Amphibia and reptilia.* Australian Government Publishing Service, Canberra. Pp. 210-220.

Hutchinson, M. N., Donellan, S. C., Baverstock, P. R., Krieg, M., Simms, S. and Burgin, S. (1990). Immunological relationships and generic revision of the Australian lizards assigned to the genus *Leiolopisma* (Scincidae: Lygosominae). *Australian Journal of Zoology* **38**: 535-554.

Innis, M. A. and Gelfand, D. H. (1990). Optimisation of PCRs. In: Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J. (eds.), *PCR protocols: A guide to methods and applications.* Academic Press, Inc., San Diego. Pp. 3-12.

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.

Jeffreys, A. J., Neuman, R. and Wilson, V. (1990). Repeat unit sequence variation in minisatellites: A novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* **60**: 473-485.

Kessler, L. G. and Avise, J. C. (1985). A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. *Molecular Biology and Evolution* **2**: 109-125.

Kimura, M. (1983). *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge.

Kocher, T. D. and White, T. J. (1989). Evolutionary analysis via PCR. In Erlich, H. A. (ed.), *PCR Technology: Principles and applications for DNA amplification.* Stockton Press, New York. Pp. 137-147.

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 Sciences of the United States of America* **86**: 6196-6200.

Kraus, F. and Miyamoto, M. M. (1991). Rapid cladogenesis among the pecoran ruminants: evidence from mitochondrial DNA sequences. *Systematic Zoology* **40**: 117-130.

Kraus, F., Jarecki, L., Miyamoto, M. M., Tanhauser, S. M. and Lapis, P. J. (1992). Mispairing and compensational changes during the evolution of mitochondrial ribosomal RNA. *Molecular Biology and Evolution* **9**: 770-774.

Lake, J. (1988). Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. *Nature* **331**: 184-186.

Lake, J. A. (1990). Origin of the Metazoa. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 763-766.

Lake, J. (1991). Tracing origins with molecular sequences: Metazoan and eukaryotic beginnings. *Trends In Biochemical Sciences* **16**: 46-50.

Larson, A. and Wilson, A. C. (1989). Patterns of ribosomal RNA evolution in salamanders. *Molecular Biology and Evolution* **6**: 131-154.

Lento, G. M., Hickson, R. E., Chambers, G. K. and Penny, D. (1995). Use of spectral analysis to test hypotheses on the origin of pinnipeds. *Molecular Biology and Evolution* **12**: 28-52.

Li, W.-H. and Gouy, M. (1991). Statistical methods for testing molecular phylogenies. In: Miyamoto, M. M. and Cracraft, J. (eds.), *Phylogenetic analysis of DNA sequences*. Oxford University Press. Pp. 249-277.

Lockhart, P. J., Penny, D., Hendy, M. D., Howe, C. J., Beanland, T. J. and Larkum, A. W. D. (1992). Controversy on chloroplast origins. *Federation of European Biochemical Societies Letters* **301**: 127-131.

Lockhart, P. J., Penny, D., Hendy, M. D. and Larkum, A. W. D. (1993). Is Prochlorothrix hollandica the best choice as a prokaryotic model for higher plant Chl-*a/b* photosynthesis? *Photosynthetic Research* **73**: 61-68.

Lockhart, P. J., Steel, M. A., Hendy, M. D. and Penny, D. (1994). Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution* **11** (4): 605-612.

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** (2): 174-190.

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

Martin, A. P., Kessing, B. D. and Palumbi, S. R. (1990). Accuracy of estimating genetic distance between species from short sequences of mitochondrial DNA. *Molecular Biology and Evolution* **7**: 485-488.

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.

Maxson, L. R. and Maxson, R. D. (1979). Comparative albumin and biochemical evolution in plethodontid salamanders. *Evolution* **33**: 1057-1062.

McCann, C. (1955). The lizards of New Zealand. Gekkonidae and Scincidae. *Dominion Museum Bulletin* **17**: 1-127.

McCann, C. (1972). Corrections to "The lizards of New Zealand" (*Dominion Museum Bulletin* **17** - 1955). *Tuatara* **19** (3): 167.

Milinkovich, M. C., Ortí, G. and Meyer, A. (1993). Revised phylogeny of whales suggested by mitochondrial ribosomal DNA sequences. *Nature* **361**: 346-348.

Mindell, D. P. and Honeycutt, R. L. (1990). Ribosomal RNA in vertebrates: evolution and phylogenetic applications. *Annual Review of Ecology and Systematics* **21**: 541-566.

Miyamoto, M. M. and Boyle, S. M. (1989). The potential importance of mitochondrial DNA sequence data to eutherian mammal phylogeny. In: Fernhom, B., Bremer, K. and Jornvall, H. (eds.), *The hierarchy of life. Molecules and morphology in phylogenetic analysis.* Elsevier Science Publishers b.v., Amsterdam. Pp. 437-450.

Miyamoto, M. M., Allard, M. W., Adkins, R. M., Janacek, L. L. and Honeycutt, R. L. (1994). A congruence test of reliability using linked mitochondrial DNA sequences. *Systematic Biology* **43**: 236-249.

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 Hillis, D. M. (1996). Molecular systematics: context and controversies. In: Hillis, D. M., Moritz, C. and Mable, B. K. (eds.), *Molecular Systematics*. Second edition. Sinauer Associates, Inc., Sunderland, Massachusetts. Pp. 1-13.

Neefs, J.-M., Van de Peer, Y., Hendriks, L. and De Wachter, R. (1990). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Research* **18**, Suppl.: 2237-2317.

Neefs, J.-M., Van de Peer, Y., De Rijk, P., Goris, A. and De Wachter, R. (1991). Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Research* **19**, Suppl.: 1987-2015.

Neefs, J.-M., Van de Peer, Y., De Rijk, P., Chapelle, S. and De Wachter, R. (1993). Compilation of small ribosomal subunit structures. *Nucleic Acids Research* **21**: 3025-3049.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press.

Nei, M. (1991). Relative efficiencies of different tree-making methods for molecular data. In: Miyamoto, M. M. and Cracraft, J. (eds.), *Phylogenetic analysis of DNA sequences*. Oxford University Press. Pp. 90-128.

Nei, M., Tajima, F. and Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* **19**: 153-170.

Noller, H. F. and Woese, C. R. (1981). Secondary structure of 16S ribosomal RNA. *Science* **212**: 403.

Noller, H. F., Moazed, D., Stern, S., Powers, T., Allen, P. N., Robertson, J. M., Weiser, B. and Triman, K. (1990). Structure of rRNA and its functional interactions in translation. In: Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Schlessinger, D. and Warner, J. R. (eds.), *The ribosome: structure, function, and evolution*. American Society for Microbiology, Washington D. C. Pp. 73-92.

Novacek, M. (1993). Genes tell a new whale tale. *Nature* **361**: 298-299.

Pääbo, S. and Wilson, A. C. (1988). Polymerase chain reaction reveals cloning artifacts. *Nature* **334**: 387-388.

Palmer. J. D. (1990). Contrasting modes and tempos of genome evolution in land plant organelles. *Trends in Genetics* **6**: 115-120.

Palumbi, S. R. and Baker, C. S. (1994). Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution* **11** (3): 426-435.

Palumbi, S., Kessing, B., Croom, H., Martin, A., McIntosh, C., and McMillan, W. O. (1989). *The simple fools guide to PCR*. Version 1.0. Department of Zoology, University of Hawaii, Honolulu.

Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L. and Grabowski, G. (1991). *The simple fools guide to PCR*. Version 2.0. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu.

Paterson, A. M., Gray, R. D. and Wallis, G. P. (1993). Parasites, petrels and penguins: does louse presence reflect seabird phylogeny? *International Journal for Parasitology* **23** (4): 515-526.

Paterson, A. M., Gray, R. D. and Wallis, G. P. (1995a). Of lice and men: the return of the 'Comparative Parasitology' debate. *Parasitology Today* **11** (4): 158.

Paterson, A. M., Wallis, G. P. and Gray, R. D. (1995b). Penguins, petrels and parsimony: does cladistic analysis of behavior reflect seabird phylogeny? *Evolution* **49** (5): 974-989.

Patterson, C. (1987). Introduction. In: Patterson, C. (ed.), *Molecules and morphology in evolution: conflict or compromise?* Cambridge University Press, Cambridge. Pp. 1-22.

Patterson, G. B. and Daugherty, C. H. (1990). Four new species and one new subspecies of skinks, genus *Leiolopisma* (Reptilia: Lacertilia: Scincidae) from New Zealand. *Journal of the Royal Society of New Zealand* **20** (1): 65-84.

Patterson, G. B. and Daugherty, C. H. (1994). *Leiolopisma stenotis*, n. sp., (Reptilia: Lacertilia: Scincidae) from Stewart Island, New Zealand. *Journal of the Royal Society of New Zealand* **24** (1): 125-132.

Patterson, G. B. and Daugherty, C. H. (1995). Reinstatement of the genus *Oligosoma* (Reptilia: Lacertilia: Scincidae). *New Zealand Journal of Zoology* **25** (3): 327-331.

Penny, D. and Hendy, M. D. (1987). TurboTree: a fast algorithm for minimal length trees. *Computer Applications in BioScience* **3**: 183-188.

Penny, D., Hendy, M. D. and Steel, M. A. (1992). Progress with methods for constructing evolutionary trees. *Trends in Ecology and Evolution* **7** (3): 73-79.

Penny, D., Watson, E. E., Hickson, R. E. and Lockhart, P. J. (1993a). Some recent progress with methods for evolutionary trees. *New Zealand Journal of Botany* **31**: 275-288.

Penny, D., Lockhart, P., Hendy, M. and Steel, M. (1993b). *Hitchhikers guide to evolutionary trees: Traveling through tree space*. Version 0.002. Farside Research Institute, Massey University, Palmerston North. i + 84 pages.

Pickard, C. R. and Towns, D. R. (1988). *Atlas of the amphibians and reptiles of New Zealand.* Conservation Sciences Publication No. 1., Science and Research Directorate, Department of Conservation, Wellington, New Zealand. 59 pages.

Pregill, G. K. (1989). Prehistoric extinction of giant iguanas in Tonga. *Copeia* **1989**: 505-508.

Rivera, M. C. and Lake, J. A. (1992). Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* **257**: 74-76.

Robb, J. (1970). A new skink of the genus *Leiolopisma* from New Zealand. *Proceedings Section C. Nederlandse Academie van Wetenschappen* **73** (2): 228-229.

Robb, J. (1973). Reptiles and amphibia. In: Williams, G. R. (ed.), *The natural history of New Zealand*. A. H. and A. W. Reed Ltd., Wellington. Pp. 285-303.

Robb, J. (1974). New Zealand lizards (2). *New Zealand's Nature Heritage* **2** (25): 681-689.

Robb, J. (1975). Two new skinks of the genus *Leiolopisma* from New Zealand. *Proceedings Section C. Nederlandse Academie van Wetenschappen* **78** (5): 477-484.

Robb, J. (1980). *New Zealand reptiles and amphibians.* William Collins Publishers Ltd., Auckland.

Sage, E. (1995). Beyond the picture postcard. *Forest and Bird* **275**: 20-30.

Sadlier, R. A. (1986). A review of the scincid lizards of New Caledonia. *Records of the Australian Museum* **39**: 1-66.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

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

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A laboratory manual.* Second edition. Cold Spring Harbour Laboratory Press, New York.

Satta, Y. and Takahata, N. (1990). Evolution of *Drosophila* mitochondrial DNA and the history of the *melanogaster* subgroup. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 9558-9562.

Sheffield, V. C., Cox, D. R. and Myers, R. M. (1990). Identifying DNA polymorphisms by denaturing gradient gel electrophoresis. In: Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J. (eds.), *PCR protocols: A guide to methods and applications.* Academic Press, Inc., San Diego. Pp. 3-12.

Shields, G. F. and Wilson, A. C. (1987). Calibration of mitochondrial DNA evolution in geese. *Journal of Molecular Evolution* **24**: 212-217.

Simon, C. (1991). Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. In: Hewitt, G. M., Johnston, A. W. B. and Young, J. P. W. (eds.), *Molecular techniques in taxonomy.* NATO ASI Series, Vol. H57. Pp. 33-71.

Simon, C., Franke, A. and Martin, A. (1991). The polymerase chain reaction; DNA extraction and amplification. In: Hewitt, G. M., Johnston, A. W. B. and Young, J. P. W. (eds.), *Molecular techniques in taxonomy.* NATO ASI Series, Vol. H57. Pp. 329-355.

Simon, C., Frati, F., Beckenbach, A. and Crespi, B. (1994). Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87** (6): 651-701.

Sneath, P. H. A. and Sokal, R. R. (1973). *Numerical taxonomy.* W. H. Freeman, San Francisco.

SRARNZ Notes (1993). Whitaker, T. (ed.), *The newsletter of the Society for Research on Amphibians and Reptiles in New Zealand.* No. 11. 7 pages.

SRARNZ Notes (1994). Whitaker, T. (ed.), *The newsletter of the Society for Research on Amphibians and Reptiles in New Zealand.* No. 12. 18 pages.

SRARNZ Notes (1995a). Whitaker, T. (ed.), *The newsletter of the Society for Research on Amphibians and Reptiles in New Zealand.* No. 13. 16 pages.

SRARNZ Notes (1995b). Whitaker, T. (ed.), *The newsletter of the Society for Research on Amphibians and Reptiles in New Zealand.* No. 14. 28 pages.

SRARNZ Notes (1996). Whitaker, T. (ed.), *The newsletter of the Society for Research on Amphibians and Reptiles in New Zealand.* No. 15. 26 pages.

Steel, M. A., Cooper, A. C. and Penny, D. (1996, in press). Confidence intervals for the divergence time of two clades. *Systematic Biology* **45** (2): 127-134.

Stevens, G. R. (1980). *New Zealand Adrift.* Reed, Wellington.

Sullivan, J., Holsinger, K. E. and Simon, C. (1995). *Molecular Biology and Evolution* **12** (6): 988-1001.

Swofford, D. L. (1985). *PAUP: Phylogenetic analysis using parsimony.* Illinois Natural History Survey, Illinois.

Swofford, D. L. (1993). *PAUP: Phylogenetic analysis using parsimony Version 3.1.1.* Illinois History Survey, Champaign.

Swofford, D. L. and Olsen, G. J. (1990). Phylogeny reconstruction. In: Hillis, D. M. and Moritz, C. (eds.), *Molecular systematics*. Sinauer Associates Inc., Sunderland, Massachusetts. Pp. 411-501.

Swofford, D. L., Olsen, G. J., Waddell, P. J. and Hillis, D. M. (1996). Phylogenetic inference. In: Hillis, D. M., Moritz, C. and Mable, B. K. (eds.), *Molecular Systematics*. Second edition. Sinauer Associates, Inc., Sunderland, Massachusetts. Pp. 407-514.

Sytsma, K. J. (1990). DNA and morphology: inference of plant phylogeny. *Trends in Ecology and Evolution* **5**: 104-110.

Taylor, R. H. and Thomas, B. W. (1993). Rats eradicated from rugged Breaksea Island (170 ha), Fiordland, New Zealand. *Biological Conservation* **65**: 191-198.

Thomas, W. K. and Beckenbach, A. T. (1989). Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *Journal of Molecular Evolution* **29**: 233-245.

Thomas, R. H., Pääbo, S. and Wilson, A. C. (1990). Reply to: Faith, D. P. (1990) "Chance marsupial relationships" (*Nature* **345**: 393-394). *Nature* **345**: 394.

Thorpe, J. P. (1982). The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. *Annual Review of Ecology and Systematics* **13**: 139-168.

Towns, D. R. (1974). Zoogeography and the New Zealand Scincidae. *Journal of the Royal Society of New Zealand* **4**: 217-226.

Towns, D. R. (1985). *A field guide to the lizards of New Zealand*. New Zealand Wildlife Service Occasional Publication No. 7., New Zealand Wildlife Service, Department of Internal Affairs, Wellington, New Zealand. 28 pages.

Towns, D. R. (1991). Response of lizard assemblages in the Mercury Islands, New Zealand, to removal of an introduced rodent: The kiore (*Rattus exulans*). *Journal of the Royal Society of New Zealand* **21** (2): 119-136.

Towns, D. R., Daugherty, C. H. and Newman, D. G. (1985). An overview of the ecological biogeography of the New Zealand lizards (Gekkonidae, Scincidae). In: Grigg, G., Shine, R. and Ehmann, H. (eds.), *Biology of Australasian frogs and reptiles*. Royal Zoological Society of New South Wales. Pp. 107-115.

Van de Peer, Y., Neefs, J.-M., De Rijk, P., and De Wachter, R. (1993). Reconstructing evolution from eukaryotic small ribosomal subunit RNA sequences: calibration of the molecular clock. *Journal of Molecular Evolution* **37**: 221-232.

Vawter, L. and Brown, W. M. (1986). Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. *Science* **234**: 194-196.

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

Vos, M. E. (1988). *A biochemical, morphological and phylogenetic review of the genus Cyclodina.* B.Sc. (Hons.) thesis, Victoria University of Wellington. v + 70 pages.

Wallis, G. P. and Arntzen, J. W. (1989). Mitochondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**: 88-104.

Ward, R. H., Frazier, B. L., Dew-Jager, K. and Pääbo, S. (1991). Extensive mitochondrial diversity within a single Amerindian tribe. *Proceedings of the National Academy of Sciences of the United States of America*, **88**: 8720-8724.

Wayne, R. K. and Jenks, S. M. (1991). Mitochondrial DNA analysis implying extensive hybridisation of the endangered red wolf *Canis rufus*. *Nature* **351**:565-568.

Wheeler, W. C. and Honeycutt, R. L. (1988). Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Molecular Biology and Evolution* **5**: 90-96.

Whitaker, T. [A. H.] and Thomas, B. (1989). *New Zealand lizards: An annotated bibliography.* Ecology Division, Department of Scientific and Industrial Research, Lower Hutt, New Zealand. 215 pages.

White, T. J., Arnheim, N. and Erlich, H. A. (1989). The polymerase chain reaction. *Trends in Genetics* **5** (6): 185-189.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* **26**: 375-400.

Wilson A. C., Ochman, H. and Prager, E. M. (1987). Molecular time scale for evolution. *Trends in Genetics* **3**: 241-247.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews* **51** (2): 221-271.

Worthy, T. H. (1987a). Osteology of *Leiopelma* (Amphibia: Leiopelmatidae) and descriptions of three new subfossil *Leiopelma* species. *Journal of the Royal Society of New Zealand* **17** (3): 201-251.

Worthy, T. H. (1987b). Palaeoecological information concerning members of the frog genus *Leiopelma*: Leiopelmatidae in New Zealand. *Journal of the Royal Society of New Zealand* **17** (4): 409-420.

Worthy, T. H. (1987c). Osteological observations on the larger species of the skink *Cyclodina* and the subfossil occurrence of these and the gecko *Hoplodactylus duvaucelii* in the North Island, New Zealand. *New Zealand Journal of Zoology* **14**: 219-229.

Worthy, T. H. (1991). Fossil skink bones from Northland, New Zealand, and description of a new species of *Cyclodina*, Scincidae. *Journal of the Royal Society of New Zealand* **21** (4): 329-348.

Zug, G. R. (1985). A new skink (Reptilia: Sauria: *Leiolopisma*) from Fiji. *Proceedings of the Biological Society of Washington* **98** (1): 221-231.

Appendix One: Constructing a Spectral Diagram (Lento-Plot) using Microsoft Word for Windows Version 6.0.

1. Setting up a file containing the information to be graphed.

This step involves removing all information from the Hadtree output file except for the data to be graphed, then formatting this data so that it can be imported into the chart datasheet. Some of the intermediate stages in this editing process, along with the end result, are shown in Appendix Figure 1.1 in order to clarify the written instructions.

- Open the Hadtree output file (hadtree.log, or whatever the file has subsequently been named). As hadtree.log is an ASCII format file, the message 'Convert File From:' will appear - select the Text Only option. Then use Ctrl+A to highlight the entire file, and choose the Courier font so that the file is once again spaced correctly.
- Delete everything up to the heading 'new ranking based on "signal minus clashes" '. Then delete everything from the words 'tree = ...' to the end of the file.
- **At this point the file should look like a. in Appendix Figure 1.1.**
- Delete the section beginning 'new ranking based on "signal minus clashes" ' and ending 'new old partition'.
- Delete the first two columns of numbers (now headed 'rank rank') and their headings.
    - To delete a column - using the mouse, place the cursor at the start of the column to be deleted, press and hold Alt, click and hold down the control on the mouse, release Alt, move the cursor across and down until the column to be deleted is highlighted, release the control on the mouse and then press del. If a mistake is made while highlighting, release the mouse control, then click it once to clear the selection. If something is accidentally deleted, use Ctrl+Z to undo the action.
- Certain numbers in the 'index' column will have stars next to them, indicating that these are the groupings (internal edges) which have been included in the tree. The number of stars equals the number of taxa minus three (for example, with eight taxa there will be five stars, while with twenty taxa there will be seventeen stars). Delete the stars and replace each with a space (to keep the columns correctly spaced) - however, keep a record of which numbers were starred, as the signals for these will be distinguished from others in the graph.
- **At this point the file should look like b. in Appendix Figure 1.1.**
- Delete the column headed 'sigl-clsh' (including the heading).
- Move the column headed 'signal' to below the 'index' column, leaving one empty line between the two columns.
    - To move a column in Word - highlight the column using Alt and the mouse, then use Ctrl+X to excise the highlighted column, place the cursor at the new position at which the column is to start, then press Ctrl+V.
- Move the column headed 'clashes' to below the 'signal' column, again leaving one empty line between the two columns.
- **At this point the file should look like c. in Appendix Figure 1.1.**

Appendix Figure 1.1 Four stages in the editing of a Hadtree output file for which a spectral diagram (Lento-plot) is to be constructed using Word for Windows Version 6.0. The hadtree.log file used in this example contains 13 signal-supported bipartitions.

a.

new ranking based on "signal minus clashes"
1st column is new ranking
2nd column is old ranking on signal alone
3rd column is partition number (index)
4th column is (signal - clashes)
5th column is signal (usually gamma) values
6th column is sum of clashes (normalised so that sum of signals and clashes equal)
(partitions marked by asterisks are in the tree)

new rank	old rank	partition index	sigl-clsh	signal	clashes
1	1	3	0.59	1.18	-0.58
2	6	96*	0.35	0.61	-0.25
3	10	100*	0.17	0.60	-0.42
4	9	126	0.14	0.60	-0.46
5	8	12	0.06	0.60	-0.54
6	3	84	-0.02	0.61	-0.63
7	11	102*	-0.03	0.60	-0.63
8	7	25*	-0.11	0.60	-0.71
9	5	106	-0.15	0.61	-0.75
10	12	10	-0.22	0.28	-0.50
11	13	17*	-0.22	0.28	-0.50
12	4	65	-0.27	0.61	-0.88
13	2	41	-0.31	0.61	-0.91

b.

index	sigl-clsh	signal	clashes
3	0.59	1.18	-0.58
96	0.35	0.61	-0.25
100	0.17	0.60	-0.42
126	0.14	0.60	-0.46
12	0.06	0.60	-0.54
84	-0.02	0.61	-0.63
102	-0.03	0.60	-0.63
25	-0.11	0.60	-0.71
106	-0.15	0.61	-0.75
10	-0.22	0.28	-0.50
17	-0.22	0.28	-0.50
65	-0.27	0.61	-0.88
41	-0.31	0.61	-0.91

..... continued next page

c.

index
3
96
100
126
12
84
102
25
106
10
17
65
41

signal
1.18
0.61
0.60
0.60
0.60
0.61
0.60
0.60
0.61
0.28
0.28
0.61
0.61

clashes
-0.58
-0.25
-0.42
-0.46
-0.54
-0.63
-0.63
-0.71
-0.75
-0.50
-0.50
-0.88
-0.91

d.

3
96
100
126
12
84
102
25
106
10
17
65
41

1.18
0.61
0.60
0.60
0.60
0.61
0.60
0.60
0.61
0.28
0.28
0.61
0.61

-0.58
-0.25
-0.42
-0.46
-0.54
-0.63
-0.63
-0.71
-0.75
-0.50
-0.50
-0.88
-0.91

- Delete the headings ('index', 'signal' and 'clashes').
- Make sure that there is one empty line before each column of numbers (one at the start of the file, one between the first and second columns and one between the second and third columns).
- **At this point the file should look like d. in Appendix Figure 1.1.**
- Save this edited file under a new name.

2. Inserting a chart into a file and setting the type of chart.

- Open the file into which the chart is to be inserted (this can be a new or an existing file).
- Use the mouse to select the Insert Chart symbol (a small picture of a graph) on the standard toolbar (if this symbol is not present on the toolbar, then select the Customize option from the Tools menu to find it).
- Set the type of chart (note that it is best to use the 'Cancel' option, rather than the 'OK' option if the settings do not need to be changed):
    - Select the DataSeries menu (using the mouse or Alt+D) and choose the Series in Columns option.
    - Select the Gallery menu (mouse or use Alt+G), choose the Column option, and then choose option 3 (this option has two sets of values in each column, one atop the other).
    - Select the Format menu (mouse or Alt+T), choose the Chart option, and make sure that the Overlap is set as 100% and that the Gap Width is set as 50%.
    - Select the Chart menu (mouse or Alt+C) and choose the Delete Legend option.

3. Importing data from the file set up in Step 1 into the datasheet.

- If the datasheet is behind the chart, bring it to the fore of the screen (either by using the mouse to click on any visible part of the datasheet or by selecting the Windows menu using the mouse or Alt+W, then selecting the Datasheet option).
- Expand the datasheet.
- Clear the existing data out of the datasheet by highlighting the existing columns of data using the mouse, pressing del, then choosing the Clear Data option.
- If there are more than three columns in use (as indicated by the number of black panels along the top of the datasheet), remove the extra columns:
    - go back to the top of the datasheet using Ctrl+Home.
    - highlight the extra columns, press Ctrl+-, then choose the Delete Columns option.
- Place the cursor in the top box of the first column, select the File menu and choose the Import Data option.
- Select the file set up in Step 1 (either type in the file name under 'File Name' or call up the file using the 'Files' option). This places all three columns contained in the file ('index', 'signal' and 'clashes') into the first column of the datasheet.

- Shift the 'signal' and 'clashes' columns from the first column of the datasheet to the second:
  - scroll down the first column until a space is reached (the empty line between the end of the 'index' column and the start of the 'signal' column).
  - highlight from (and including) this space to the bottom of the numbers in the first column of the datasheet, then press Ctrl+X (if the message 'picture too big to copy, only cell data will be carried' is displayed, press 'OK').
  - place the cursor in the top box of the second column, then press Ctrl+V.
- Shift the 'clashes' column from the second column of the datasheet to the third:
  - Repeat the previous step, this time scrolling down the second column and placing the cursor in the top box of the third column.
- Clear any extra rows from below the imported data (as indicated by the black panels along the left-hand-side of the datasheet):
  - highlight from below the last row of the imported data to the bottom of the extra rows, press Ctrl+-, then choose the Delete Rows option.

4. Formatting the chart.

- Again, if a particular setting does not need to be changed, it is best to use the 'Cancel' rather than the 'OK' option.
- Bring the chart to the fore of the screen.
- **Turn the x axis labels vertical and place them below the negative signals in the graph, and format the x axis:**
  - select the x axis by clicking on it with the mouse (the x axis has been selected when there is a small box on either end of it), select the Format menu, choose the Text option and change the Orientation to vertical (rotated 90°anticlockwise from the horizontal).
  - select the Patterns option within the Text box and set the Tick Mark Types to None and the Tick Labels to Low.
  - set Axis to Custom and set the Style, Colour and Weight options to the desired settings.
  - select the Font option within the Patterns box and set the Font, Size, Style, Colour and Background to the desired settings.
- **If every second x axis label has disappeared:**
  - expand the chart horizontally by selecting the vertical borders of the chart box (a border has been selected when the cursor turns into a double-headed arrow), clicking on and dragging the borders horizontally until the labels reappear.
  - If the borders have been dragged to the peripheries of the screen and the labels have not yet reappeared, select the Windows menu, choose a lesser % view and repeat the expansion.
  - If the chart has been expanded as far as it can go and the labels still have not reappeared, then the page margins or page orientation of the file may have to be adjusted, or some of the smallest signals may need to be removed (see Step 6. Trouble-shooting). Dropping to a lower font may solve the problem, but may also result in labels which are too small to be read clearly.

- **Ensure that the y axis labels are horizontal and next to the axis, and format the y axis:**
  - select the y axis (as with the x axis, this axis has been selected when there is a small box on either end of it), select the Format menu, choose the Text option, and check that Orientation is set to the Automatic option.
  - select the Patterns option within the Text box and check that the Major Tick Mark Type is set to Cross, that the Minor Tick Mark Type is set to None and that the Tick Labels are set to Next to Axis.
  - set Axis and the Font parameters as for the x axis.
- **Set the colour, pattern and border of the signals (bars or columns) in the graph:**
  - select the positive signals, again by using the mouse (these signals have been selected when there is a small box on each of the end columns and another on the central column), select the Format menu, choose the Patterns option, set Area to Custom and set pattern and Foreground to the desired colour and pattern for the majority of the signals (for example, plain and white). Make sure that the 'Apply to All' option is off or when the columns representing the edges in the tree are changed to a different colour in the next step, all of the columns will change colour.
  - set Border to Custom and set Style, Colour and Weight to the desired settings.
  - repeat these steps for the negative signals.
- **Set the colour of the signals for the edges which are in the tree (those with stars next to them in the original hadtree.log file):**
  - use the keyboard horizontal arrows to select each individual column whose colour is to be changed (an individual column has been selected when there are boxes on it alone; note that positive and negative columns also have to be selected separately), select the Format menu, choose the Patterns option, set Area to Custom and set Pattern and Foreground to the desired pattern and colour for these columns (for example, plain and black).
- **Place a border around the chart (inside the x and y axis labels) and format this border:**
  - select the position at which the border is to be inserted by clicking the mouse on the area inside the x and y axis labels, but outside the positions of the columns (a square of boxes marking where the border will go will appear).
  - select the Format menu, choose the Patterns option, set Border to Custom and set Style, Colour and Weight to the desired settings; set Area to None.
- **Insert and format the x and y axis titles:**
  - select the Chart menu, choose the Titles option, choose the Category (X) Axis option and type in the desired title (such as Bipartitions).
  - click the mouse on an unoccupied area of the chart - the title should now be selected (that is, surrounded by a box; if not, click on the title), select the Format menu, choose the Patterns option and set Border and Area to None.
  - choose the Font option within the Patterns box and set the Font, Size, Style, Colour and Background to the desired settings.
  - choose the Text option within the Font box and check that the Orientation is set to the horizontal option and that the Text Alignment is set to Centre.

- select the Chart menu, choose the Titles option, choose the Value (Y) Axis option and type in the desired title (such as Support for and against bipartitions).
- again, click the mouse on an unoccupied area of the chart and set the Patterns and Font options as for the x axis; set the Text Alignment to Centre and change the Orientation to vertical (rotated 90°anticlockwise from the horizontal).

- **Place a border around the graph (outside the x and y axis titles):**
  - this has not been done for the spectral diagrams presented in this study, however if a border in this position is desired, click the mouse on the area outside the x and y axis titles (a square of boxes should appear just inside the chart box frame) and format as for the border inside the x and y axis labels.
- **Select the File menu, choose the Update option, select the File menu again and then the Exit and Return to file-name option.** Note that as it is quite easy to inadvertently add extra and unwanted formatting which can be difficult to remove, it is not a good idea to save this chart as the default chart even if several other similar charts are to be set up.
- **Toggle between the file and the chart until the chart borders (both horizontal and vertical) have been adjusted to give a chart of roughly of the desired size** (fine-tuning can be done in the next stage).

5. Formatting the chart once back in the document.

- use Ctrl+Home to get to the top of the chart, Ctrl+End to get to the bottom.
- Fine-tune the size of the chart (if there are problems with every second x axis label disappearing, go to the Trouble-shooting section before performing this step). Click once on the chart (a box with small black boxes on its sides/corners should appear around the chart), place the cursor on the appropriate side/corner box (a double-headed arrow will appear), then click and drag with the cursor until the desired adjustment has been made. It is best to do the main chart size adjustment within the chart, however, as major adjustments in the file can lead to odd-sized labels, titles and so on.

6. Trouble-shooting.

- If every second x axis label has disappeared and expanding the vertical borders of the chart has not brought them back, then the margins of the pages may need to be adjusted, the page orientation may have to be changed or some of the bipartitions with the smallest signals may have to be deleted from the datasheet.
  - **adjusting the margins of the pages - the only way to do this and ensure that the chart is influenced by the new formatting appears to be to open a new file, format it and then move the chart from the first file into the new file:**
    - go back into the chart (double-click on the chart to edit it) and adjust the vertical borders of the chart until the chart can be examined using 66% view, then update the chart and exit.

- copy the chart (select the chart by clicking on it once, then press Ctrl+C).
- open a new file, press Ctrl+A, select the File menu, choose the Page Setup option, choose the Margins option and then decrease the size of the left and right margins.
- save this file (under a new name), close it, then reopen it.
- insert the chart in the new file by positioning the cursor at the appropriate position and pressing Ctrl+V.
- save the file, close it and reopen it again, then go back into the chart and expand the vertical borders as far as they will go. Update the chart and exit, then save the file again.
- if this step has not succeeded in making all of the x axis labels visible, then go on to the next step.

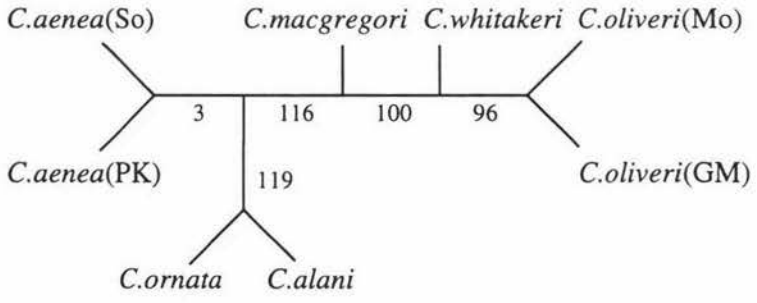
- **changing the page orientation - again, it appears that the only reliable way of doing this is to open a new file, format it and then move the chart from the first file into the new file:**
  - adjust the vertical borders of the chart and copy it as above.
  - open a new file, press Ctrl+A, select the File menu, choose the Page Setup option, choose the Paper Size option and set the Orientation to Landscape, then choose the Margins option and adjust the margins back to the appropriate values.
  - save this file (under a new name), close it, then reopen it.
  - insert the chart in the new file using Ctrl+V.
  - Save the file, close it and reopen it again, then go back into the chart and expand the vertical borders as far as they will go. Update the chart and exit, then save the file again.
- **If neither of the previous options has succeeded in making all of the x axis labels visible, then some of the bipartitions with the smallest signals (the ones on the bottom of the datasheet and on the right-hand-side of the chart) will have to be deleted from the datasheet:**
  - highlight the 'index', 'signal' and 'clash' values to be removed, press del and choose the Clear Data option.

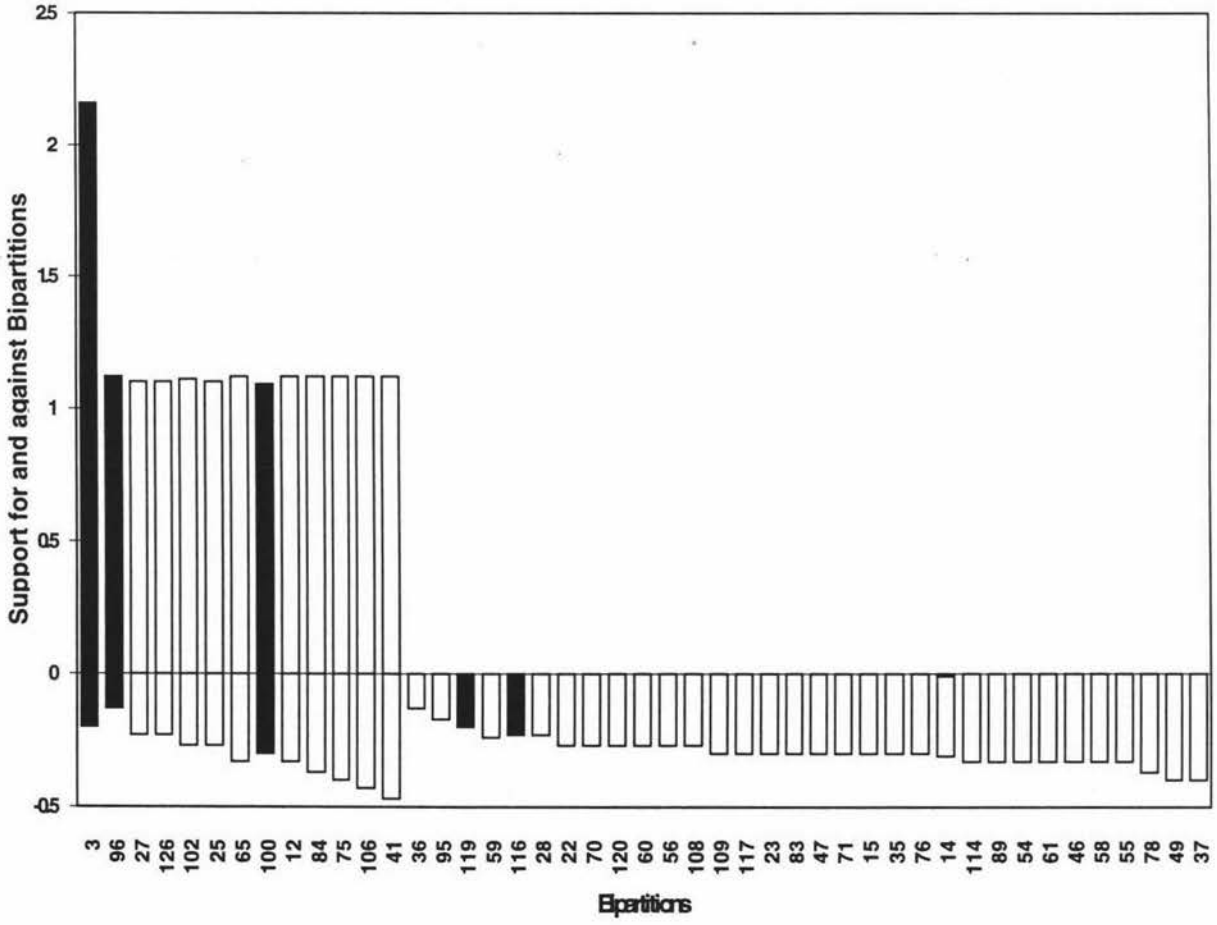
- One other problem was experienced. When the diagrams were printed from a particular computer, all of vertically oriented labels were far fainter than the horizontally oriented ones (although all of the settings for these labels were identical). Printing from a different computer was found to be the solution to this problem.

Appendix Two: Four Frequency and Two Frequency *Cyclodina* Spectral Diagrams.

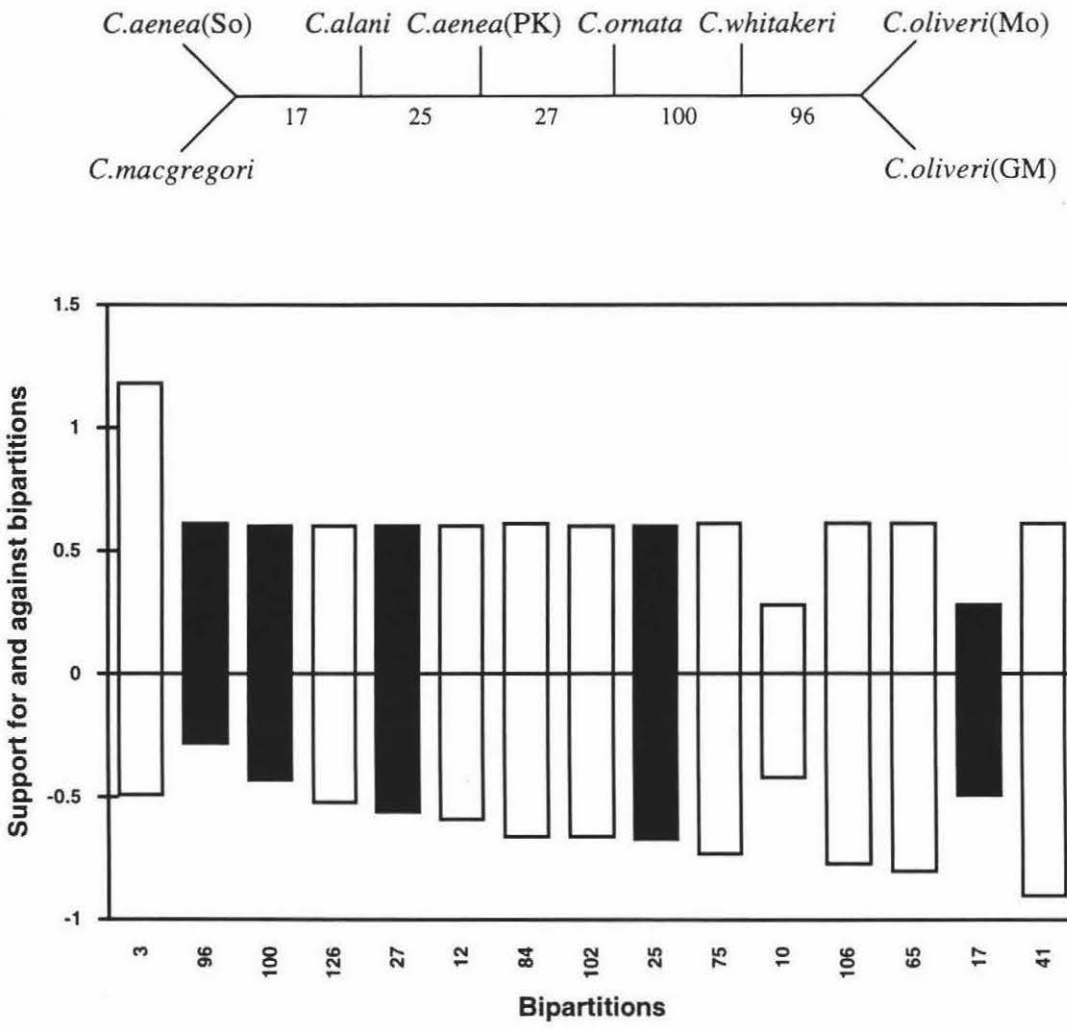
The format for each of these diagrams is the same: the unrooted tree is shown diagrammatically at the top with the appropriate bipartition number next to each of the internal edges. The main (lower) figure shows bipartitions considered for inclusion in the optimal tree (a tree being a mutually compatible set of t-3 bipartitions, where t is the number of taxa). The support for each bipartition (after any correction for multiple changes) is shown above the x axis, the sum of the contradictions for each bipartition is below the x axis. The latter values are normalised so that they sum to the same value as the support. Bipartitions in black are those selected as being in the optimal tree by the tree selection criterion (closest tree), these correspond to the numbers shown in the unrooted tree at the top of the figure.

Appendix Figure 2.1 Spectral diagram for four colour (Kimura three-parameter corrected) information from the 382 column *Cyclodina* dataset.

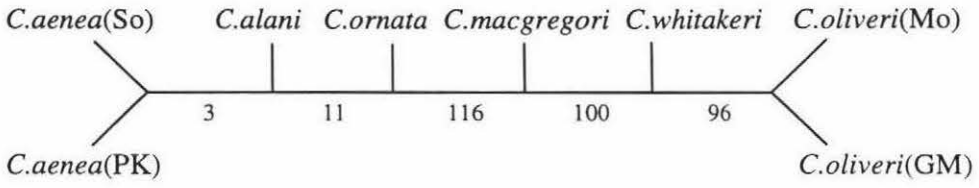


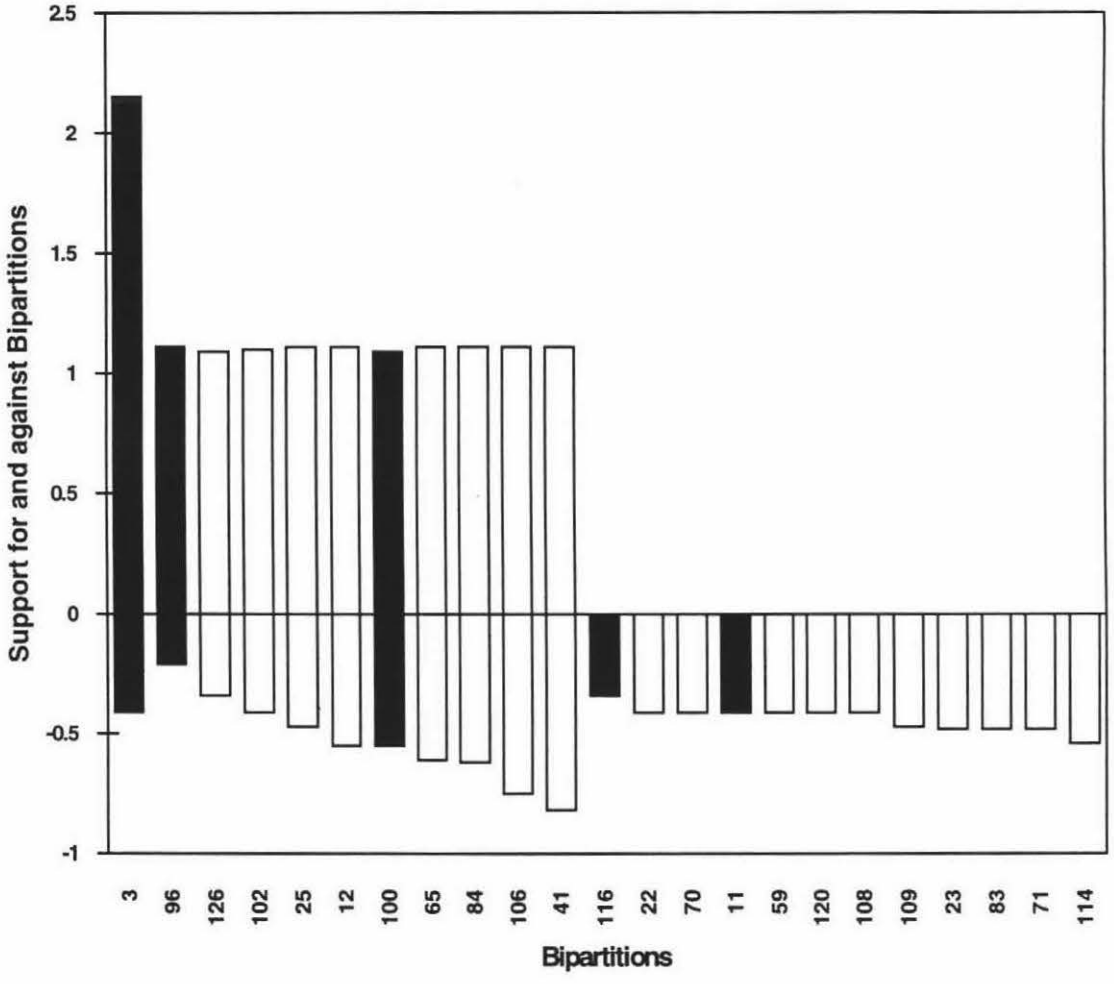


Appendix Figure 2.2 Spectral diagram for two colour (sum of seven) information from the 382 column *Cyclodina* dataset.



Appendix Figure 2.3 Spectral diagram for four colour (Kimura three-parameter corrected) information from the 378 column *Cyclodina* dataset.





Appendix Figure 2.4 Spectral diagram for two colour (sum of seven) information from the 378 column *Cyclodina* dataset.