GENERA: *EUCRYPHIA*, *GRISELINIA* AN

CORIARIA

A thesis submitted to the University of Leicester for the de
Doctor of Philosophy

By

Martha Rebecca Jane Clokie BSc. (Dundee), MSc. (Edinb

September 2001

Department of Biological Sciences
University of Leicester

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

I hereby declare that no part of this thesis has been previously submitted to this or any other university as part of the requirements for a higher degree. The work described here, unless otherwise acknowledged, was conducted by the undersigned who is fully responsible.

Signed.....

Acknowledgements

I would like to thank the many people who made this thesis possible. This starts therefore with my supervisors Colin Ferris and Richard Gornall for choosing me to carry out this project and for all the help during the work and write up.

Next, I would like to thank the dozens of people who provided me with plant material. In particular Philip Ashby and other staff at the Royal Botanic Garden of Edinburgh were extremely helpful. I would also like to thank Colin Forde from the Seaforde Garden in Co. Down Ireland for sending the samples of *Eucryphia*. Material, fossil, morphological data and advice from Richard Barnes were invaluable in producing the *Eucryphia* chapter. Ann Conolly and the botanical gardens in Canberra were also very kind in providing material. For the *Griselinia* chapter I would greatly like to thank Corporación National Forestal (CONAF) for allowing me to collect in the national parks of Chile, and the many Chilean botanists who helped me track plants down. I would especially like to acknowledge Melica-Muñoz, Patricio Peñoilillo and Marcelo Baeza for their kind assistance and hospitality. As chief enforcer of time tables, Gaynor Evans assisted greatly in the success of the Chilean field work. She was always good company and uncomplainingly carried herbarium presses up and down the length of Chile. Family helped a lot with collecting material too, cousin Bernard Honey in New Zealand and my dad and brother Ollie helped in gardens around the south of England. This was assisted in the Hillier Arboretum by Allenn Coombes. I would like to thank Pete Thompson for providing samples and for helpful discussions about the *Coriaria* chapter.

Many people helped with lab work and analysis of data during the 3 years these include. These include Michael Möller, Uta Praekelt, Jon Combes, Ezio Rosato, Dave Twell and Clive Stace. But most people in the biology department were helpful at some point!

In terms of sanity preservation I would like to thank Alma Swan, Herbert McGregor + perhaps particularly importantly Jake for his weekend visits. Also I would like to thank my friends who only said positive things when I appeared from my computer during writing up. Some friends deserve a very special mention. Cat Pashley has to be thanked for not minding too much when I borrowed her things, messed up her bench and for being totally supportive throughout the lab work and various crisis. I would like to thank Richard Reeve for initially encouraging me to apply for the PhD and for always being understanding. Mike Wilkinson deserves thanks for all his help in my third year. Paul Jenkins was a constant source of good ideas and solace. Liz and Rachael were v.i.p. in terms of pony therapy. Dan and Elliott were great house mates for the three years and provided much musical and historical input respectively. Craig, Norah and Sash, and Alison were also always positive and were there at opportune moments.

My parents and brothers Ollie, Sam and Ben, have been consistently helpful and encouraging. They all said the right things at the right times, and always seemed to have surprising confidence in me. My mother also helped with proof reading. More recently I have to thank Helena for making the last few months so painless by driving me to work everyday and bringing me back to write my thesis in the evenings. Also, all my new lab mates in 'Micro 2' at Warwick were relentlessly cheery and helpful. My boss Nick Mann was also nice and understanding about my time out for meetings back at Leicester.

Finally, I would like to give a huge thanks to Pete Wheatley who has been a source of both delightful hindrance and useful discussion and without whom my writing up would have been a lot less fun. He also wrote the program that produced the graphs to show the positions of variability in the gene regions. He has been very supportive, cheery and patient and has put up with mood swings and frantic-ness and never ceased to be lovely.

ABSTRACT

Three plant genera with South American–Australasian disjunctive distributions were studied: *Eucryphia*, *Griselinia* and *Coriaria*. The aims of the project were: a) to assess the utility of different chloroplast and nuclear DNA sequences for phylogenetic reconstruction; and b) by means of molecular clocks to establish whether or not this distribution is due to vicariance (continental drift) or more recent dispersal events. DNA was extracted, and regions of interest were amplified and sequenced. These were *rpo*A, *trn*L-F and *trn*H-K from the chloroplast genome, and rDNA 5.8S and intergenic spacers 1 and 2, *Adh* and two loci of the *G3pdh* gene from the nuclear genome. In each of the three genera, a single trans-Antarctic disjunction was indicated. Molecular phylogenies were produced using parsimony. Individual base pair variation was analysed in detail, and graphs were drawn to highlight variable regions. Sequences providing the most resolution were *trn*L-F, ITS and the two loci of the *G3pdh.*

Robust molecular phylogenies were produced which are consistent with morphological, fossil and biogeographical evidence. Constancy of mutation rate across the different branches of the trees was checked using the Tajima test. Where constancy was established, the rate of evolution for the different gene regions was used to estimate the divergence times within each genera. Calibrations were performed using fossil data for *Eucryphia* and *Coriaria*; average mutation rates derived from *Eucryphia* were applied to the *Griselinia* data set, owing to the absence of fossils.

New Zealand became separated from Antarctica (and therefore South America) 95 – 80 million years ago, whereas average divergence times for *Coriaria* and *Griselinia* were found to be around 3 and 50 million years respectively. Similiarly, with respect to *Eucryphia*, Australia split from Antarctica 40 million years ago but the divergence time is 23.23 million years. The disjunction therefore must have arisen by dispersal which may have occurred either in the water by rafting or floating, or in the air in wind currents (*Eucryphia* has winged seeds) or by means of birds (*Coriaria* and *Griselinia* have fleshy diaspores).

A review of similar disjunctions in other plants and animals revealed that vicariance explanations were more likely to apply to groups at the taxonomic ranks of genus and above, whereas dispersal explanations were more likely to apply at the level of genus and below. The generic level is consequently of great interest.

Contents

CHAPTER 1

INTRODUCTION

"Among the problems raised by the distribution of plants and animals in the southern hemisphere there is none which takes a more central position and is more stimulating to the imagination than the problem of transantarctic relationships. We have before us the broken circle of southern lands-southern South America, South Africa, Tasmania-Australia, New Zealand - separated by wide stretches of ocean, but populated by a group of biota containing numerous groups whose strongly disjunct elements are more closely related to one another than to any other group, and in the center of the scene we are faced with the dormant Antarctica continent, hiding its secrets beneath a mighty ice cap."
(Brundin, 1966; p)

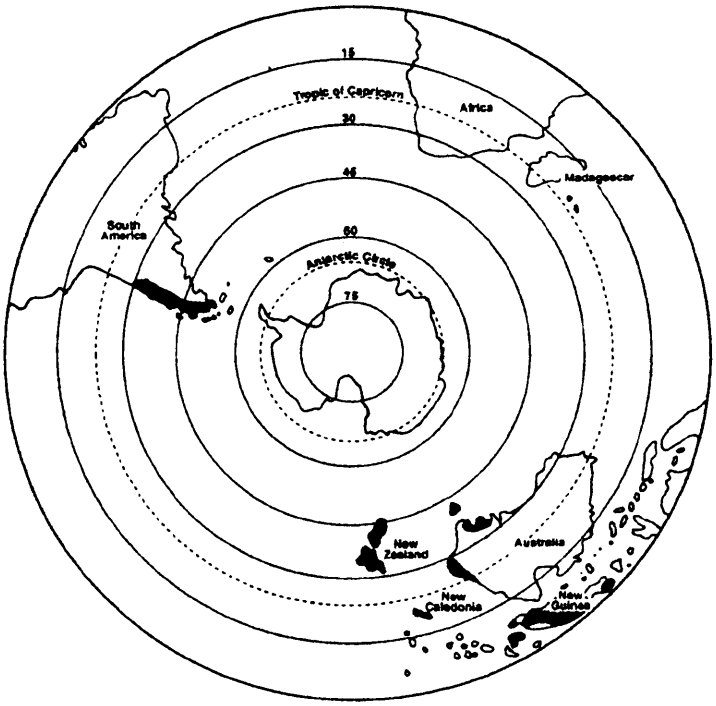

Fig. 1.1 The general distribution for *Nothofagus* to show the South American-Australasian distribution (from Manos, 1997).

1.1 The Antarctic floristic element

One of the best documented and studied biotic disjunctions in the world is that between Australasia, New Zealand and southern South America (Fig. 1.1). The pattern was first described by Hooker (1853), who examined the floras of these regions, and noticed the considerable similarity, especially at the generic level, recognising it as the Antarctic floristic element. Many biogeographers since have been intrigued by this distribution. One reason for the high level of interest is that it is shared by so many groups of organisms; not only plants (Good, 1974; Thorne, 1972), but also a diverse array of animals, including marsupials (Springer *et al.*, 1997), ratite birds (van Tuinen, 1998), freshwater fish (Helfman, 1997; Berra, 1997; Waters, 1999), insects (Brundin, 1966; Cranston and Naumann, 1991) and *Onychophora* (Monge-Najera, 1995). A summary list of many of the taxa involved was provided by Crisci *et al.* (1991) and a list of plants is given in Table 1.1.

In order to interpret the present day distribution pattern of a group of organisms, it is absolutely essential to be certain that the group is monophyletic. It must be pointed out here that few of the taxa listed in Table 1 have been formally tested for their phylogenetic status, although many are phenetically highly distinct and therefore likely to be monophyletic. Some of the groups that have been thought to show the Australasia/South America disjunction turn out to be either polyphyletic (independently evolved from different ancestors) or paraphyletic (sharing primitive characters rather than apomorphic, or derived characters). A good example of this is the plant family Epacridaceae. The Epacridaceae contains *ca.* 30 genera and 450 species. Hitherto, the family has been presumed to show a disjunction between Australasia and South America, owing to the occurrence in the latter of the monotypic genus *Lebetanthus*. A recent phylogenetic analysis of the family based on morphological evidence provided only weak support for its monophyletic status, one of the 'problem' genera being *Lebetanthus* (Powell *et al.*, 1996). Data from chloroplast DNA (*rbc*L) (Crayn *et al.*, 1996) and from 18S rDNA sequences (Kron, 1996) resolved the situation and showed that the Epacridaceae are clearly not monophyletic; in particular, the South American *Lebetanthus* is separated from the other Epacrid genera and is now placed in the Ericaceae. As a result, the disjunction in the family Epacridaceae does not exist, and attempts to explain it are irrelevant.

Where it is clear that groups are monophyletic, then a disjunct distribution may have two main explanations (which are not mutually exclusive). The distribution may be a result of either vicariance events, such as continental drift, dispersal, or a combination of the two processes. In the context of distributions extending between Australasia and South America, the debate began in 1853 when Hooker proposed vicariance as an explanation for this spectacular disjunction, whilst Darwin (1859) countered by favouring long-distance dispersal. Many botanists (excluding Hooker) assumed that the southern (and particularly the Australian) floras had been produced as a result of immigration from the north – the so-called 'invasion' hypothesis, extensively reviewed by Barlow (1981). For a long time the argument centred on the method by which the flora had arrived, either by means of land-bridges (Burbidge, 1960; Van Steenis, 1962) or by means of long-distance dispersal (Darlington, 1965). Hooker's original suggestion, reiterated and amplified later (Hooker, 1860, 1867:25-26), that the circum polar Antarctic flora might have had an origin from an ancient flora that had already existed on the various southern land-masses, was more or less ignored, except by Herbert (1953, 1967). The advent of geological proof in the 1970s and 1980s for the continental drift theory of Wegener (1915) rekindled interest in Hooker's hypothesis of an ancient flora differentiated by means of vicariance events, and focused attention more squarely on dispersal and vicariance as contrasting explanations for the disjunctions. An introduction to these two opposing views is provided in Chapter 2.

1.2 Testing the hypotheses

Timings are essential to this whole debate. As pointed out by Humphries and Parenti (1986), "Vicariance predicts that taxa in two (or more) areas and the barriers between them are the same age; whereas dispersal always predicts that the barrier predates the taxa." A critical test of the two explanations can be made by dating the divergence event. This date can be estimated by using DNA sequences and the fossil record. DNA sequences from genes and spacer regions can be likened to a "molecular clock" (Wolfe, 1989, Gaut, 1992), whereby the rate of mutation of the sequences happens at a consistent rate and thus, by calculating the number of substitutions that separate species, the time of their divergence can be estimated. As described in Chapter 3, molecular clocks must be used with care as

they do not all 'tick', or evolve, at a constant rate, but may vary in rate between the different regions of the genome, and more seriously, can vary among lineages (Li and Tanimura, 1987; Penny *et al.*, 1987). The successful calibration of their rate of 'ticking' also depends on the vagaries of the fossil record.

A second approach to testing whether vicariance or dispersal events are involved is to compare the area cladogram for the taxa involved with the area cladogram of the regions of endemism based on geological evidence. A close match indicates that vicariance could be a sufficient explanation of the distribution pattern. A mismatch indicates that dispersal is involved. This approach was pioneered by Humphries and Parenti (1986, 1999) and developed further by the work of Page (1998).

In order to test between a vicariance and a dispersal explanation for distributions, a taxon should have: a) an unambiguous and well-understood phylogeny, in particular, the group in question should be monophyletic, and b) ideally the taxon should have a good fossil record that can be used to date various nodes of the individual lineages.

1.3 Selection of genera for survey

The genera *Coriaria*, *Eucryphia* and *Griselinia* were selected for the following reasons. Most importantly they have similiar disjunct distributions. Secondly, they are all relatively small genera, with fairly well understood classical taxonomies. The relatively small size, makes it more likely for all species within the genus to be obtained. Initial investigations were performed to assess the feasibility of obtaining material. They are of interest for phylogenetic studies as little or no molecular work has thus far been carried out on them. They have slightly different life histories in that *Coriaria* and *Griselinia* seeds are dispersed via a berry and are thus more likely to undergo animal-assisted dispersal than are the dry winged seeds of *Eucryphia*. *Eucryphia* and *Griselinia* contrast nicely in their distributions in that *Eucryphia* has two species in Chile and five in Australasia and *Griselinia* has two species in Australasia and five in Chile. *Coriaria* is interesting in that although the majority of its radiation is in Austalasia it is also found in the Northern Hemisphere.

Table 1.1 Taxa of seed plants with an Australasia/South America distribution, based on lists collated by Good (1974), Thorne (1973) and Crisci *et al.* (1991). (SA = southern South America, Au = Australia, Tas = Tasmania, NZ = New Zealand)

Family	*Genus (no. spp.)*	*SA*	*Au*	*Tas*	*NZ*	*Other locations*
Araucariaceae	*Araucaria* (19)	+	+			New Caledonia, New Guinea Norfolk Island
Araliaceae	*Pseudopanax* (6)	+		+	+	
Atherospermataceae	*Laurelia* (2)	+			+	
Boraginaceae	*Plagiobothrys* (c. 50)		+			W.N.America, E. Asia
Campanulaceae	*Hypsela* (4)	+	+		+	
Caryophyllaceae	*Colobanthus* (20)	+	+	+	+	Kerguélen Is., New Amsterdam
Centrolepidaceae	*Gaimardia* (3)	+		+	+	New Guinea
Compositae	*Abrotanella* (16)	+	+	+	+	
	Celmisia (64)		+		+	
	Lagenophora (15)	+	+		+	
	Trichocline (22)	+	+			
Coriariaceae	*Coriaria* (16)	+			+	S. Europe, E. Asia, C. America
Cunoniaceae	*Weinmannia* (190)	+			+	Madagascar, Mascarenes, Malaysia, Pacific Is.
Cyperaceae	*Carpha* (11)	+	+	+		S. Africa
	Oreobolus (8)	+			+	Pacific Is.
	Uncinia (35)	+				Pacific Is., S. Indian Atlantic oceans
Donatiaceae	*Donatia* (2)	+			+	
Elaeocarpaceae	*Aristotelia*	+	+	+	+	
Eucryphiaceae	*Eucryphia* (7)	+	+	+		
Ericaceae	*Gaultheria* (150)	+	+	+	+	N. & C.America; E. & S. E.Asia
Gesneriaceae	*Negria* (1) *Drepanthus*	+	+		+	New Caledonia
Goodeniaceae	*Selliera* (1)	+	+		+	
Gramineae	*Rhytidosperma*	+	+	+	+	
Griseliniaceae	*Griselinia* (7)	+			+	
Gunneraceae	*Gunnera* (40)	+		+	+	New Guinea; C. America
Haloragidaceae	*Haloragis* (26)	(+)	+		+	(SA=Juan Fernandez Is.), New Caledonia, Rapa Is.
Iridaceae	*Libertia* (20)	+	+	+	+	
Juncaceae	*Marsippospermum* (3)	+			+	Falkland Is.
Juncaceae	*Rostkovia* (1)	+			+	
Juncaginaceae	*Maundia* (1)	+	+			

Family	*Genus (no. spp.)*	*SA*	*Au*	*Tas*	*NZ*	*Other locations*
Juncaginaceae	Triglochin (14)	+	+	+	+	N. Hemisphere
Labiatae	*Tetrachondra* (2)	+			+	
Leguminosae	*Sophora microphylla* (17)	+			+	Lord Howe Island, Chatham Islands + other Pacific Islands
Liliaceae	*Astelia* (25)	+			+	Mascarenes, New Guinea, Polynesia to Hawaii
Nothofagaceae	*Nothofagus* (21)	+	+	+	+	New Guinea, New Caledonia
Onagraceae	*Fuchsia* (100)	+			+	Andes, C. America, W. Indies
Podocarpaceae	*Dacrydium* (25)				+	S.E. Asia to New Zealand
	Eupodocarpus (10)	+			+	New Caledonia
	Stachycarpus (8)	+			+	New Caledonia
Polygonaceae	*Muehlenbeckia* (15)	+	+		+	
Portulacaceae	*Calandrina* (150)	+	+			
Proteaceae	*Lomatia* (12)	+	+			
	Oreocallis (5)	+	+			E. Malaysia
	Orites (5)	+	+			
Ranunculaceae	*Anemone* (120)	+		+	+	N. Hemisphere
Restionaceae	*Leptocarpus* (16)	+	+		+	Indomalaysia
Rhamnaceae	*Discaria* (15)	+	+		+	
Rosaceae	*Acaena* (*c.* 100)	+	+	+	+	Andes to California; New Guinea; S. Africa
	Geum (65)	+		+	+	S. Africa; N. Hemisphere
Rubiaceae	*Nertera* (6)	+	+		+	China to Java, Society Is., Hawaii, Madagascar
Scrophulariaceae	*Euphrasia*	+	+	+	+	New Guinea, New Caledonia, N. Hemisphere
	Hebe (65)	+			+	
	Jovellana (6)	+			+	
	Ourisia (25)	+		+	+	
Stylidiaceae	*Phyllachne* (4)	+		+	+	
Thymelaeaceae	*Drapetes* (4)	+		+	+	New Guinea
Umbelliferae	*Azorella* (70)	+				Antarctica, Falkland Is.
	Lilaeopsis (15)	+	+		+	
	Oreomyrrhis (25)	+	+	+	+	New Guinea
	Schizellema (13)	+	+		+	
Winteraceae	*Drimys* (14)	+	+		+	Malaysia, New Guinea, New Caledonia, C. America

The overall aim of the present work is to study molecular evolution in three plant genera with disjunct distributions in the Southern Hemisphere with a view to dating the divergence events associated with the disjunctions.

In particular, I aim to:-

- Establish well-supported molecular phylogenies for three genera of flowering plants (*Eucryphia* (Eucryphiaceae), *Griselinia* (Griseliniaceae) and *Coriaria* (Coriariaceae)), using sequences from the chloroplast and the nuclear genomes.

- Compare the rates and processes of evolution found in the various sequences of DNA that are recovered and to evaluate their relative utility at recovering phylogenetic trees.

- Compare the taxon area cladograms with those inferred from geological evidence, to determine whether vicariance events are consistent with the patterns of distribution.

- Test for evolutionary rate-constancy in the various sequences among the lineages in the three genera.

- Calibrate a molecular clock for those sequences that are rate-constant.

- Date the divergence events that separate the Australasian from the South American gene pools in the three genera and thereby test whether vicariance or dispersal is a more likely explanation for the distribution patterns observed.

- Compare the two genera (*Griselinia* and *Coriaria*) that have fleshy diaspores, and are therefore potentially dispersed long-distances by birds, with each other and with a third genus (*Eucryphia*) which has dry, capsular fruits, unlikely to be so dispersed.

CHAPTER 2

AN INTRODUCTION TO VICARIANCE AND DISPERSAL

2.1 Vicariance explanations for biogeographical distributions

Between the years of 1839 and 1843, in his role as Assistant Surgeon with the Erebus and Terror expedition under Captain James Clarke Ross, J.D. Hooker visited Australia, New Zealand, Patagonia and many sub-Antarctic islands. He was struck by the close floristic affinities, particularly at the generic level, between these southern lands, despite the fact that they were widely separated by enormous expanses of ocean (Crisp, 1999). He concluded that he had discovered a circumpolar Antarctic flora, and that the present day floras had partially diverged owing to the development of physical and climatic barriers between land which had once been continuous. In a letter to Darwin (Gardiner, 1997) he stated:

"Enough is here given to show that many of the peculiarities of each of the three great areas of land in the southern latitudes are representative ones, effecting a botanical relationship as strong as that which prevails throughout the lands within the Arctic and Northern Temperate zones, and which is not to be accounted for by any theory of transport or variation, but which is agreeable to the hypothesis of all being members of a once more extensive flora, which has been broken up by geological and climatic causes."

Despite the strength with which Hooker held his view that the distribution was caused by vicariance, it was not popular and indeed was scarcely discussed until later the next century (Croziat, 1952; Croziat, 1958; Herbert 1950, Herbert 1953; Nelson and Rosen 1981). The basis of modern support for vicariance explanations is exemplified best by the work of Humphries and Parenti (1999). These authors were dogmatic in their support for a vicariance explanation for the disjunctions observed in the southern Hemisphere. Their argument focused on the fact that so many disparate biota share the same or very similar geographical distributions. They concluded that there must be a common explanation for

all such taxa showing this distribution, and made it plain that they regarded vicariance events, as manifested by continental drift, as the answer. They asserted that "conformation to pattern is of primary importance in biogeographic analysis. The dispersal capabilities of individual taxa are secondary, and become irrelevant if all the taxa conform to a particular pattern" (Humphries and Parenti, 1986, p. 67, no.7); and "a general explanation for a pattern applies equally to all taxa supporting the pattern" (p. 67, no. 18)). They go on, "groups of plants and animals in the same area (a biota) that share a cladistic pattern share a history" (p. 67, no. 21). And all this, irrespective of the fossil record or the rank of the taxa concerned.

2.1.1 Continental drift and geological history

The resurgence of interest in vicariance explanations for southern Hemisphere disjunct distributions began after Wegener (1915) proposed his theory of continental drift, in which the continents are not fixed in space, but move over time, splitting and uniting. Although geomorphologists did not immediately accept Wegener's theory, some biologists immediately saw it as an explanation for many observed patterns of biological distribution (Crisp *et al.*, 1995). Although at this stage there was no known geophysical mechanism to explain the continental movement, several influential plant geographers including Cain, (1944) and Good, (1974) saw continental drift as the predominant mechanism for the movement of plants; long-distance dispersal was a much less important mechanism. In the1970s and,1980s, an increased knowledge of mid-ocean ridges, sea-floor spreading, palaeomagnetism and geomorphology provided confirmation of Wegener's ideas (Crisp *et al.*, 1999). Although there are uncertainties, the course of the main events are clear (Cox and Moore, 2000) and are summarised in Table 2.1 and Fig. 2.1.

About 245 Mya the Earth's land masses coalesced to form the supercontinent of Pangaea. At this time separate floras and fish faunas can be distinguished in the northern Siberian part, the equatorial Euramerican part and the eastern region of the southern part (Edwards, 1990; Young, 1990). Subsequently, Pangaea began to break up. The first split that occurred resulted in the formation of the northern supercontinent of Laurasia, and a southern one of Gondwanaland which contained what is now Africa, India, Madagascar, South America, Antarctica, Australia, New Guinea and New Zealand.

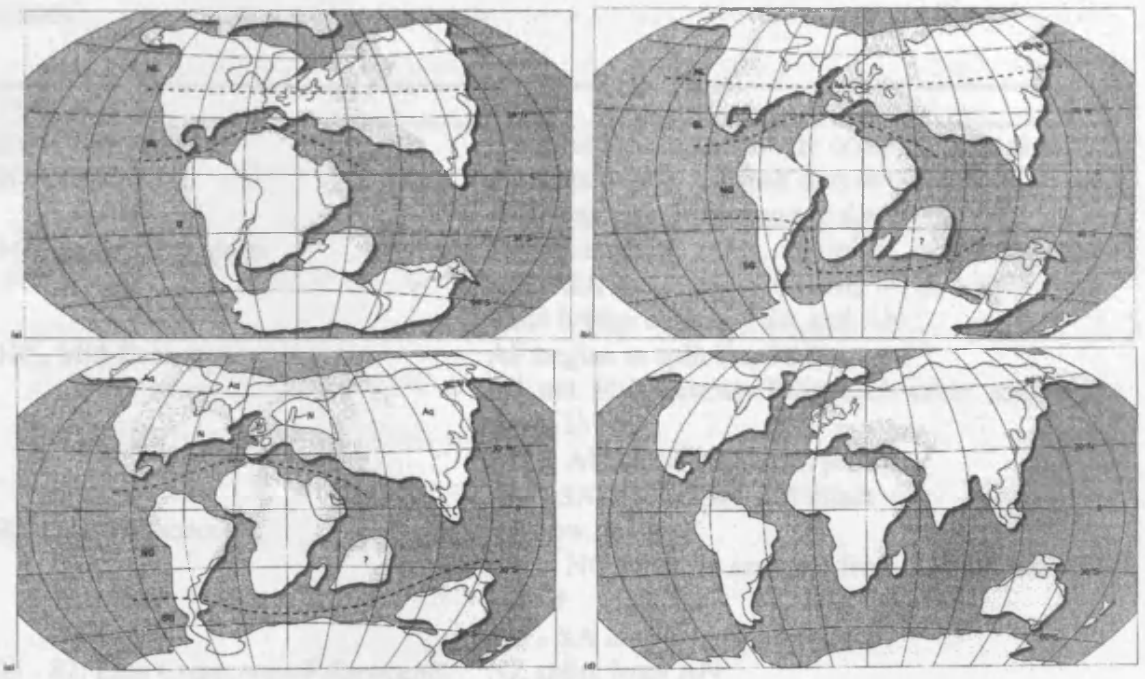

Fig. 2.1 World geography in: a) the Early Cretaceous, about 140 million years ago; b) the mid-Cretaceous, about 105 million years ago; c) the late-Cretaceous, about 90 million years ago and d) the Eocene, about 40 million years ago: Floral provinces in a – c are: Aq, Aquilapollenites; G, Gondwanan; NG, northern Gondwanan; NL, northern Laurasian; SG, southern gondwanan; SL, southern laurasian, ?, indicates that the affinities of the Cretaceous flora are unknown. Diagrams are from Cox and Moore (2000).

It is the break-up of Gondwanaland that is of most importance to the present work. This began in the early Cretaceous (*ca.* 140 Mya). The major events that occurred are illustrated in Fig. 2.1b-d. Initially, in the late Mesozoic, Gondwanaland split into East Gondwana (Antarctica, Australia and New Zealand), West Gondwana (South America and Africa), and India. Importantly, however, a land bridge (the Scotia Arc) connected Antarctica to South America up until *ca.* 35 Mya. In West Gondwana, Africa separated from South America in the middle Jurassic (LaBrecque and Barker, 1981), and India in the early Cretaceous (Johnson *et al.*, 1980; Owen, 1983). The break-up of East Gondwana began in the late-Cretaceous, more or less coinciding with a phase of angiosperm radiation and dispersal. Angiosperms first appear in the fossil record in the early-Cretaceous, 135 Mya. The diversification into families was relatively rapid; by 95 Mya many modern angiosperm families are recognisable in the Northern Hemisphere Cox and Moore (2000).

Table 2.1 Time line showing dates of important tectonic and climatic events. (AF = Africa, AN = Antarctica, AU = Australia, IN = India, MG = Madagascar, NC = New Caledonia, NG = New Guinea, NZ = New Zealand, SA = S. America.). Compiled from text within Cox and Moore (2000).

Time, Mya	*Event*
260 - 245, Jurassic	Land-masses form a single continent, Pangaea
200, Jurassic	Pangaea begins to break into two; Laurasia in the north and Gondwanaland in the south
140, Early Cretaceous	Gondwanaland splits, forming two connected parts: SA + AF and AN + AU + NZ + IN + MG; Land bridge connects SA and AN
105, Mid-Cretaceous	AF begins to split from SA IN and MG separate from each other and from AN/AU/NZ AN + AU + NG + NZ still joined AN - SA land bridge still intact
90, Late Cretaceous	AF now isolated AU + NG begin to separate from AN but are still joined AN - SA land bridge still intact
95 - 82, Late Cretaceous/Palaeocene	NZ splits from AN
40, Eocene	IN joins Asia AU + NG now isolated AN - SA land bridge begins to break
36,	Global decline in temperature; sea cools, glaciers form on AN
35, Oligocene	NZ substantially isolated from AN; AN-SA land bridge gone; circumpolar sea current and W Wind drift form
35 - 14	AN coasts and moutains covered with *Nothofagus* forests. Subtropical climate in NZ
14, Miocene	AN with maximum glaciation
10, Miocene	AU reaches present position; TA separates from AU
6.5 - 5	Global drop in sea-level
5 - 3	NZ mountains begin to uplift
1.6 - 0.01, Pleistocene	Glaciations in NZ. True alpine vegetation develops

At the time when flowering plants and mammals first evolved, radiated and dispersed in the Cretaceous, Australia was still joined to South America and Antarctica (Fig. 2.1a). The climate in this region was warm and humid (Cox and Moore, 2000) and it is known that at least the coastal regions of the land were forested (Beadle, 1981). The southern Gondwanaland flora included many podocarps, conifers, and ferns, as well as the

flowering plant genus *Nothofagus*. Although these rainforest floras once covered much greater areas of land, today they are only found in isolated, scattered patches. This is partly due to tectonic history. Australia began to split from Antarctica around 96 Mya but, for the first 50 million years, movement was very slow so Australia and Antarctica shared a single weather system Cox and Moore (2000). Between 46 Mya and 34 Mya, the northwards movement of Australia was far more rapid and indeed by the early-Oligocene the separation between the continents was sufficient for a deep-water circumpolar current of cooler water and westerly winds to become established Cox and Moore (2000). There was a global temperature decline at around the time of the Eocene/Oligocene boundary (*ca.* 36 Mya). At this time, the deep sea began to cool and glaciers formed on Antarctica (36-23 Mya) (Kennett, 1977). Maximum glaciations, with glaciers occurring at sea level, first occurred on the eastern Antarctic mainland in the Miocene at around 14 Mya (Grant-Mackie, 1979; Mercer, 1983). Although it has reduced substantially in size, this ice cap has remained a semi-permanent feature. Most of the melting of western Antarctica is thought to have occurred during the last interglacial age, *ca.*, 125,000 years ago (Mercer, 1983). Australia continued to move northwards until the late-Miocene (6.5-5 Mya) when it reached a high-pressure zone of low rainfall at a longitude of 30°S. Also, in late-Miocene, the sea ice built up suddenly around the western Antarctic archipelago (Brain, 1984), this coincided with a global drop in sea level.

At this time, large tracts of Australia became highly arid and the rainforests became fragmented, remaining only in isolated patches Cox and Moore (2000). Australia continued to move northwards until it hit a large oceanic trench where old crust material was sinking downwards. This crust remained underneath the north of the continent and caused it to rise from the mid-Miocene onwards to form the mountains of New Guinea. These mountains were high, with a cool, wet climate, and were colonised by elements of the Australian rainforest flora, including *Nothofagus*. Australia is not thought to have reached its present position, adjacent to Asia, until as recently as 10 Mya, which is when Tasmania became separated.

The islands of New Zealand split from Gondwanaland in the late-Cretaceous or early Palaeocene, 95-82 Mya (Owen, 1983; Lingen *et al.*, 1994). By 35 Mya, New Zealand was substantially separated from Antarctica. Around 26 Mya, after the Scotia Arc land bridge became an island chain, and after Tasmania and New Zealand had become separated from

eastern Antarctica, a current began to flow in a clockwise direction around the Antarctic; this is known as the Antarctic circumpolar current (Lüning, 1990). The flora of New Zealand has been influenced by three main proceses (Cox and Moore, 2000). The first is that in the early Cenozoic, New Zealand was progressively submerged by the sea until the late-Oligocene, at which point there was little sign of land Cox and Moore (2000). Secondly, there was a lot of volcanic activity with extensive lava flows during the Pliocene. The final reason is that due to the large number of mountains and the southerly location, New Zealand was heavily glaciated during the Pleistocene. Therefore, partly because of its unique combination of geographic and temporal isolation, and partly because of its climatic and geological history and high level of habitat diversity, the flora of New Zealand is relatively diverse and largely endemic (Wardle, 1991). This flora has all the hallmarks of a young one: 85% of the New Zealand angiosperm species are endemic, but only 10% of the genera are endemic and no endemic families exist. Interestingly, there are no native terrestrial mammals.

New Zealand has several genera of plants and animals that many biologists believe cannot have reached the islands by trans-oceanic dispersal Godley (1975). These include *Nothofagus*, *Podocarpus*, the reptile *Sphenodon*, and the flightless moa and kiwi. Therefore, they claim, some parts of New Zealand must have remained emergent from the sea in order to provide refugia for these biota Godley (1975). Other biologists, for example Pole (1994), believe that all biota have arrived in New Zealand by long-distance dispersal. Cox and Moore (2000) believe that the more warmth-loving plants must have arrived in New Zealand by long-distance dispersal after the end of the Pleistocene glaciations.

2.1.2 Generalised tracks

Platnick and Nelson (1978) argued that the fundamental issue facing biogeographers is not whether a particular pattern is the result of vicariance or dispersal, but rather whether the pattern corresponds to a general pattern of area interconnections (and thus reflect the history of those areas). They stated that "what is needed is a method of analysis that will pass to determine whether two given distribution patterns correspond to each other or not, so that we can test the general hypothesis that the pattern of relationships of areas indicated by one group is a general one. After a hypothesised general pattern is corroborated, we may be able to ascribe it to vicariance or dispersal by independent evidence of a history"

In vicariance biogeography the aim is to discover patterns of distribution that are congruent with each other and that therefore may share a common explanation. These shared distributions are called generalised tracks. Their explanation is likely to be vicariance if the patterns can be shown also to be congruent with the pattern of geological events. In practice, taxon cladograms are used to produce area cladograms in which the taxa have been substituted by their areas of endemism. These area cladograms are then compared with each other and also with an area cladogram produced for the geographical regions involved based on geological data. A good fit is consistent with a vicariance explanation (Humphries and Parenti, 1999).

A vicariance explanation predicts that numerous taxa will share the same basic pattern of geographical relationships and dates. Studies conducted to date in search of generalised tracks have produced mixed results. In a recent study Crisci *et al.* (1991) investigated explicitly the relationship of South American biota to counterparts in Australia and New Zealand, they examined in detail 17 taxa, including plants, animals and fungi. Analysis revealed no common pattern of interrelationships, and the intersection of the 17 cladograms did not lead to a unified hypothesis. But instead of inferring dispersal events, the authors proposed that southern South America must be a composite area, i.e., a geological mosaic. Although it is not known if they are correct, it should be pointed out that at least two of their 17 phylogenies are probably erroneous (*Embothriinae* and *Nothofagus*). The entire study may therefore have been an unfair test of the methodology and hence of the vicariance hypothesis.

2.2 Dispersalist explanations

Although Darwin was a keen follower and friend of Hooker, he did not accept Hooker's views on a vicariance explanation for the southern Hemisphere disjunctive distributions and, in his *Origin of Species,* he proposed long-distance dispersal as the mechanism for the generation of similarities in these southern Hemisphere floras:

"I am inclined to look in the southern, as in the northern Hemisphere, to a former and warmer period, before the commencement of the Glacial period, when the Antarctic lands,

now covered with ice, supported a highly peculiar and isolated flora. I suspect that before this flora was exterminated by the glacial epoch, a few forms were widely dispersed to various points of the southern Hemisphere by occasional means of transport, and by the aid, as halting places, of existing and now sunken islands, and perhaps at the commencement of the glacial period by icebergs. By these means, I believe the southern shores of America, Australia, New Zealand have become slightly tinted by the same peculiar forms of life" (Darwin, 1859).

Dispersal potential is therefore a key factor, and it is true that the geographic distribution of taxa often corresponds to their ability to disperse. Bats, for example, are the only mammals are native to New Zealand and to Hawaii. Similarly, there are animals which are incapable of surviving in salt water, such as frogs and salamanders, and these are absent from most oceanic islands.

In the case of plants, there are three main ways by which their diaspores become dispersed: by air, by water and by animal. In the latter case, seeds may either be carried internally and subsequently passed through the gut, or the seeds may adhere to the animal externally. As well as theorising about long-distance dispersal being the mechanism behind southern Hemisphere plant distributions, Darwin worked on all three of these explanations. He performed numerous experiments on the survival of seeds after immersion in salty water. He also made observations on whether seeds could be transported in the mud on the feet of vertebrates, and survive the passage through the digestive system of birds (Darwin, 1859).

Most seeds only move short distances of zero to a few tens of metres (Howe and Smallwood, 1982). Nevertheless, the unusual events which move seeds long distances are very important in plant biogeography (Cain, 2000). However, with a few notable exceptions (e.g., Carlquist, 1966, 1967; Sorensen, 1986) in recent years, very little attention been paid to the significance of this phenomenon (Howe and Smallwood, 1982). One of main problems with long-distance dispersal is that it is so difficult to document directly that few people have tried to do so. Despite practical difficulties in measuring it, however, its importance should not be ignored. (Cain, 2000). Reviews of long-distance dispersal and its relation to phytogeography are given in Beck (1976) and Raven and Axelrod (1974), and a discussion of its importance in terms of colonisation events has been presented by Cain *et al.* (2000).

2.2.1 Dispersal by air

Organisms may be carried by winds, in updrafts and storms, and/or by secondary dispersal over the substrate (Cain, 2000). There is a huge variability in the ability of organisms to disperse large distances. Many small organisms can survive the extreme conditions that exist to enable them to be blown long distances through the air. These include encysted protozoans, tardigrades, and rotifers and the spores of fungi, liverworts, mosses and ferns (Cox and Moore, 2000). Because of the variability that exists in plant seeds in size, shape, design and dormancy capacity there are many different mechanisms by which dispersal can occur. Some seeds are very small and light, e.g. those of the Orchidaceae, and can potentially be dispersed long distances by the wind (Ridley, 1930).

Seeds and fruits have a variety of adaptations to wind-mediated dispersal. These include those with a rigid or membranous wing at one end, which include seeds or one-seeded fruits (samaras), as in the Proteaceae and Aceraceae. Other kinds have a papery wing around the entire seed or at each end, for example, Ulmaceae. There are also seeds that have two lateral wings that resemble the wings of an aeroplane, as in members of the Cucurbitaceae. It is not known how efficient these devices are at increasing the distance of travel (Norberg, 1973).

Diaspores with parachute-like appendages are potentially more efficient for the purposes of long-distance dispersal . These include most members of the Compositae (van der Pijl, 1972). Each individual parachute contains an umbrella-like plumose crown of hairs or pappus above a small, one-seeded fruit or achene. These parachutes can be carried in the slightest gusts of wind and they have been shown to travel hundreds of miles (van der Pijl, 1972). Another adaptation to dispersal by wind is the presence of cotton-like hairs the fruit itself or on the surface of the seeds. These seeds are not thought to be carried quite as far as those that employ parachutes, but can travel considerable distances nonetheles (van der Pijl, 1972). Examples of families with this seed dispersal method are Salicaceae and Platanaceae (van der Pijl, 1972).

2.2.2 Dispersal by water

Many organisms are capable of travelling large distances on board rafts of floating vegetation and debris. Organisms carried in this way again include plant seeds and the eggs of animals, such as snails, but organisms more sensitive to desiccation including earthworms and lizards can also be transported (Cox and Moore, 2000).

As well as being able to travel by rafting, some angiosperm seeds have special modifications that aid their buoyancy and thus allow a greater distance when not attached to a raft, to be travelled (Murray, 1986). For example, air is retained in chambers surrounding the seed as in *Xanthium* species, *Atriplex inflata* and *A. sponiosa.* Alternatively, air spaces may develop within specialised cell-layers in the seed coat, as in *Menyanthes trifoliata* and *Calla palustris* (Murray, 1986). Another attribute is the presence of thick and impermeable seed coats.

Cain *et al.* (2000) pointed out that even where no special dispersal mechanism is evident, dispersal may nevertheless occur. Thus diaspores can be moved through rafting or some such hitch-hiking mechanism.

Long-distance dispersal in the ocean has been demonstrated by many species of plants, e.g. coconut (Ridley, 1930). The successful flotation of topical fruits and seeds in sea-water has been extensively documented by Guppy (1906, 1917) and more recently by Gunn and Dennis (1976). Many species of legumes from all three subfamilies have also been shown to be extensively distributed by ocean currents (Gunn and Dennis, 1976; Raven and Polhill, 1981). A good example of this involves *Sophora microphylla,* a legume with buoyant seeds that have been shown to survive immersion in sea-water for at least three years (Sykes and Godley, 1968). As explained later, the distribution of this species from New Zealand, across various Pacific islands to Chile is likely to be a consequence of long-distance dispersal by water (Hurr *et al.*, 1999).

In a recent study by Green (1999), the strand line of Christmas Island was surveyed for the presence of seeds over a period of four years. Seeds from 63 species of plants in 49 genera were found. Leguminous seeds were the most common but many others were also often encountered. Green concluded that only around a third of the seeds originated from the

island and most came from the Indonesian islands of Sumatra, the Moluccas and the eastern coast of Kalimantan.

Darwin was quite convinced that even seeds more sensitive to desiccation could have travelled long distances in the sea, he demonstrated this by submerging as many different seeds as he could in salt water and testing their subsequent ability to germinate (Desmond and Moore, 1992). He took great delight in teasing Hooker with the results. Hooker at this time was the director of Kew Gardens and supplied Darwin with the most sensitive seeds that he could find (Desmond and Moore, 1992).

2.2.3 Dispersal by animals

Numerous plant diaspores are dispersed by animals, either internally in the gut, or externally attached to the feet, fur or feathers. Some freshwater animals may also be dispersed by this means, for example snail eggs in mud may become attached to birds (Cox and Moore, 2000). External dispersal relies on the propagule of the organism being able to withstand highly desiccating conditions for a prolonged period of time. This usually means that the propagule is in a perennating form (Cox and Moore, 2000). As far as long distances are concerned, birds and bats are likely to be the most important of the animal vectors. A combination of wind and animal dispersal can occur in cases where the animal vectors are occasionally blown long distances in the wind.

In a fascinating review, Sorensen (1986) stated that far more attention has been given to seed dispersal by frugivores, ants and scatterhoarding birds and mammals than has been given to seed dispersal by adhesion to animal fur or feathers. She pointed out that in a sample of 10 regional floras, 5.2-12.8% of the families and a total of 0.8-5.8% of species within these floras contained adhesive diaspores. The genus *Uncinia* (Cyperaceae) was cited as a classic example of an island plant which has been observed adhering to seabird feathers. This has been postulated as a mechanism to explain the distribution of the different species on many different mountain tops on Pacific islands in New Zealand (Hamlin, 1959).

2.2.4 *Diaspore morphology in genera with trans-Antarctic disjunctions*

The list presented earlier of genera that have a trans-Antarctic disjunction (Table 1.1) is used here as the basis for a survey of the different types of diaspore that are involved. The obvious question to be addressed is whether this group of genera shows a disproportionate number of adaptations to dispersal or not. The results are presented in Table 2.2.

Of the 67 taxa in Table 2.2, 18 have fleshy diaspores and are therefore potentially dispersed by birds or bats. Two have adaptations that would allow attachment to the fur or feathers of an animal and thereby facilitate external transport. Six have adaptations of one sort or another to wind dispersal, one to water dispersal and 40 appear to have no special adaptations. Obviously this is simply a survey taken from the available literature, and therefore it may be that on further investigation more specialisations are realised. This is still a useful preliminary survey though as it shows that amongst the diaspores of these plant genera with disjunctive distributions in the southern Hemisphere, in terms of adaptations to dispersal, there is a predominance of fleshy berries. However it is also apparent that most diaspores appear to have no specialised dispersal mechanism.

Table 2.2 Types of diaspore among plant genera with disjunctive distributions in the southern Hemisphere.

Family	*Genus*	*Diaspore morphology*
Araucariaceae	*Araucaria*	Cones, wind-dispersed seeds
Araliaceae	*Pseudopanax*	Berry
Atherospermataceae	*Laurelia*	Small, hairy achenes with long plumose styles
Boraginaceae	*Plagiobothrys*	Dry nutlets
Campanulaceae	*Hypsela*	Dry capsule
Caryophyllaceae	*Colobanthus*	Dry capsule
Centrolepidaceae	*Gaimardia*	Membranous capsule
Compositae	*Abrotanella* /	Achene with a rudimentary or no pappus
	Celmisia /	Achene with pappus of about two series of rather few, unequal rigid, scabrid bristles
	Lagenophora /	Achene often imperfect, no pappus
	Trichocline /	Achene often imperfect, no pappus

Family	*Genus*	*Diaspore morphology*
	Vittadinia	Achene with pappus
Coriariaceae	*Coriaria*	Indehiscent, compressed, oblong, crustaceous achenes, enclosed by fleshy petals
Cunoniaceae	*Weinmannia*	Coriaceous, septicidal capsule
Cyperaceae	*Carpha /*	Dry capsule
	Oreobolus /	Nut enveloped above the middle in the appressed scales, crustaceous, obovoid, trigonous with a broad terminal depressed area.
	Uncinia	Adhesive dispersal to the feathers of seabirds
Donatiaceae	*Donatia*	Dry, indehiscent capsule
Elaeocarpaceae	*Aristotelia /*	Fleshy berry
	Crinodendron /	Dry capsule loculicidlly 3-5-valved, few seeds
	Dubouzetia /	Short, woody capsule
	Peripentadenia	Dry capsule, subglobose, single seed with reddish aril
Ericaceae	*Gaultheria*	Fleshy capsule
Eucryphia	*Eucryphia*	Dry capsules with winged seeds
Gesneriaceae	*Negria*	Loculicidal capsule
	Drepanthus	Loculicidal capsule
Goodeniaceae	*Selliera*	Berry, two-celled, variable in size, indehiscent
Gramineae	*Rhytidosperma*	Dry achene
Griseliniaceae	*Griselinia*	Berry
Gunneraceae	*Gunnera*	Fleshy berry
Haloragidaceae	*Haloragis*	Dry capsule
Iridaceae	*Libertia*	Capsule, coriaceous or membranous
Juncaceae	*Marsippospermum /*	Dry capsule
	Rostkovia /	Dry capsule
Juncaginaceae	*Triglochin /*	Dry follicle
	Maundia /	Dry follicle
	Tetroncium	Dry follicle
Labiatae	*Tetrachondra*	Fruit usually dry, more or less a schizocarp
Leguminosae	*Sophora*	Dry nuts with buoyancy adaptations
Liliaceae	*Astelia*	Berry
Nothofagaceae	*Nothofagus*	Dry capsule

Family	*Genus*	*Diaspore morphology*
Onagraceae	*Fuchsia*	Berry
Podocarpaceae	*Dacrydium /*	Seed ovoid, erect, outer coat short, sheathing at its base, sometimes fleshy
	Eupodocarpus /	Seed ovoid, erect, outer coat short, sheathing at its base, sometimes fleshy
	Stachycarpus	Seed ovoid, erect, outer coat short, sheathing at its base, sometimes fleshy
Polygonaceae	*Muehlenbeckia*	Fleshy berry
Portulacaceae	*Calandrina*	Dry capsule
Proteaceae	*Lomatia /*	Coriaceous seeds with a terminal truncate samara-like wing
	Oreocallis/	Coriaceous seeds
Proteaceae	*Orites*	Fruit follicular dry seeds with terminal or falcate wing
Ranunculaceae	*Anemone*	Dry achene
Restionaceae	*Leptocarpus*	Nut enclosed in the perianth
Rhamnaceae	*Discaria*	Drupe dry, coriaceous, of three cocci
Rosaceae	*Acaena /*	Achene enclosed in the indurated often barbed calyx tube; pericarp membranous, bony or coriaceous (barbed calyces of some species form sticky burrs)
	Geum	Achene, with persistent hooked style
Rubiaceae	*Nertera*	Drupe, red fleshy
Scrophulariaceae	*Euphrasia /*	Dry capsule
	Hebe /	Dry capsule
	Jovellana /	Dry capsule
	Ourisia	Capsule, loculicidal
Stylidiaceae	*Phyllachne*	Two-valved dry capsule
Thymelaeaceae	*Drapetes*	Sub-baccate seed with fleshy endosperm
Umbelliferae	*Azorella /*	Dry schizocarp with no spines or hooks
	Lilaeopsis /	Dry schizocarp
	Oreomyrrhis /	Dry schizocarp
	Schizellema	Dry schizocarp
Winteraceae	*Drimys*	Indehiscent berry

It is important to remember that dispersal is only the initial factor responsible for the establishment of new populations. Even if newly dispersed individuals survive when they arrive in a new territory, they may be unable to find mates. Colonising species of plants therefore often have life-histories that involve hermaphroditism, self compatibility, apomixis or vegetative reproduction (Stebbins, 1974). Similarly parthenogenesis is common in colonising animals, otherwise a single fertilised female is required in order to establish a new population (Begon *et al.*,1996). Another feature that enhances the chances of animal survival is the tendency to travel in flocks.

The new physical environment to which the organisms move to may be so dissimilar to the environment from which the organisms originated, that even if mates are available, they may be out-competed by established native species.

It is important to remember that plant distributions are not solely attributable to the ability of plants to negotiate small barriers. The fact that over 200 different immigrant flowering plants have reached very small isolated island groups in Hawaii shows that plants can cross even large stretches of ocean, especially when intermediate island stepping stones are available. Given the millions of years that we are dealing with, even very small chances cannot be ignored (Ashton, 1986) and rare events take on a disproportionate significance (Cain *et al.*, 2000).

2.3 Some case studies

In the following pages, a review is presented of various studies that have attempted to elucidate the nature of the trans-Antarctic disjunction. The studies have been selected on the basis that: a) a well-documented and supported phylogeny of the group exists that clearly shows it to be monophyletic; and b) a fossil record is available whereby historical events may be inferred.

2.3.1 Araucaria (Araucariaceae)

The distribution of *Araucaria* is generally considered to be a relict type that has arisen through vicariance (Pole, 1994; Macphail, 1997; Hill, 1994). To put the genus into context, the family Araucariaceae consists of three genera: *Araucaria* (19 species), *Agathis* (13 species) and the monotypic and recently discovered *Wollemia* (Setoguchi *et al.*, 1998). A recent study based on *rbc*L evidence showed that *Araucaria* and *Agathis* are sister genera and *Wollemia* is their outgroup (Setoguchi *et al.*, 1998). On the basis of its morphology, *Araucaria* has been classified into four extant sections: *Araucaria*, *Bunya*, *Eutacta* and *Intermedia* (Wilde and Eames, 1952; Stockey, 1982) and one extinct section *Yezonia* (Ohsawa *et al.*, 1995). The genus has a disjunct distribution in the southern Hemisphere being found in Chile, Argentina, southern Brazil, New Caledonia, Norfolk Island, Australia, and New Guinea.

Based on both morphological and molecular data, *Araucaria* is undoubtedly monophyletic (Setoguchi, *et al.*, 1998). Fig. 2.2 shows the consensus tree from Setoguchi's *rbc*L data which consisted of 1322 bases from 29 species representing almost the entire family. According to this, the four extant sections fall into two distinct clades. The major clade consists of section *Eutacta* (whose species occur in New Caledonia, Australia, New Guinea and Norfolk Island) and its sister contains the remaining three sections. Within this latter clade, the two South American species form a clade (sect. *Araucaria*) that is sister to a clade containing the two remaining monotypic sections: *A. bidwillii* from Australia (sect. *Intermedia*) and *A. hunsteinii* from New Guinea (sect. *Bunya*). The phylogeny thus indicates a single trans-Antarctic disjunction.

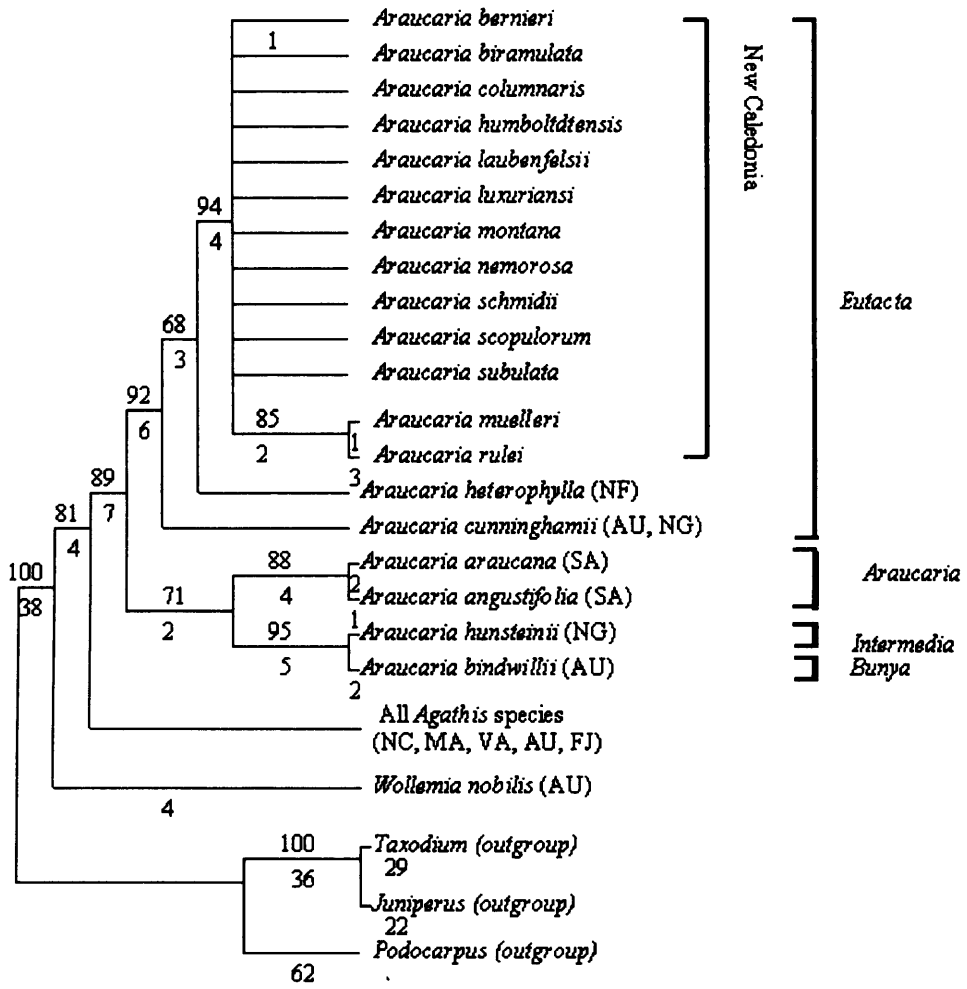

Fig. 22 Strict consensus tree of the 20 equally parsimonious trees for Araucariaceae based on cpDNA *rbc*L sequences. Numbers below internodes are mutations and bootstrap values of >50% are included above the relevant internode. Where AU = Australia, FJ = Fiji, MA = Malaysian region, NC = New Caledonia, NF = Norfolk Island, NG = New Guinea, SA = South America, VA = Vanuatu. (Redrawn from Setoguchi *et al.*, 1998).

The fossil record of *Araucaria* is extensive and has been studied by Miller (1977, 1988), Stockey (1982) and Hill (1994). The earliest fossils are of Jurassic-age and are attributable to sections *Bunya* and *Eutacta*. They show that the genus was a major component of Mesozoic forests in both the northern and southern Hemispheres. Fossils from sections *Araucaria* and *Intermedia,* however, are of more recent, late-Cretaceous origin (and possibly only Tertiary in the case of sect. *Araucaria*) and have only been found in the southern Hemisphere. It is hypothesised that the common ancestor of sections *Araucaria* and *Intermedia*, and possibly also of section *Bunya,* was distributed across Gondwanaland during the Mesozoic and early-Tertiary, and that these sections were distinct by the time of the Eocene at the latest, before South America had separated fully from Antarctica. The fossil record is clearly compatible with a vicariance explanation for the current disjunction

in the genus between South America and Australia. In terms of the molecular data, there are 10-13 differences (out of 1322 base pairs) between the South American and the Australian species. This suggests that in *Araucaria*, rate of evolution of the *rbc*L gene is relatively slow.

2.3.2 Atherospermataceae (Laurales)

The Atherospermataceae (Laurales), or southern sassafrases, are a small southern Hemisphere family consisting of two genera in Chile and 12 in Australasia. The family has caused much controversy over the number of trans-antarctic disjunctions that have occurred and about the role of transoceanic dispersal in its evolution (Renner *et al.*, 2000). The morphological evidence suggests that the two monotypic Chilean genera, *Laurelia* and *Laureliopsis* are related to different Australopacific species, thus implying that two disjunctions have occurred. The fossil history of Atherospermataceae dates back to the Upper Cretaceous (pollen, 88-86 Mya) and there are Atherosperm fossils spanning the period from then up to the Pleistocene (Renner *et al.*, 2000). Owing to the undoubted age of the family and its ecological association with putative Gondwanan relict groups (e.g. *Nothofagus*), its disjunctive distribution has been attributed to vicariance events (Schodde, 1969).

Renner *et al.* (2000) conducted a phylogenetic study on the Atherospemataceae based on an analysis of six chloroplast genes: *rbc*L (1,434 bp), *rpl*16 (918 bp), *trn*T-*trn*L (681 bp), *trn*L-*trn*F (401 bp), *atp*B-*rbc*L (724 bp) and *psb*A-*trn*H. The phylogeny that resulted from analysis of these sequences is shown in Fig. 2.3.

Firstly, it can be seen that *Laureliopsis* is nested within *Laurelia*. Secondly, the phylogram shows that there are two Australasia/South America disjunctions present in the family, viz. between *Laureliopsis philippiana* in Chile and *Laurelia novae-zelandiae* in New Zealand; and between *Atherosperma/Nemuaron* in Australia/New Caledonia and the *Laurelia /Laureliopsis* clade in Chile. From a molecular clock calibration of the *rbc*L sequence it was concluded that the origin of the genera *Laurelia/Laureliopsis* in Chile occurred at *ca.* 65 Mya, contemporary with vicariance events associated with the break-up of Gondwana. In contrast, the origin of the *Laurelia novae-zelandiae* lineage was dated at 30-50 Mya and hence is likely a product of long-distance dispersal from South America to New Zealand.

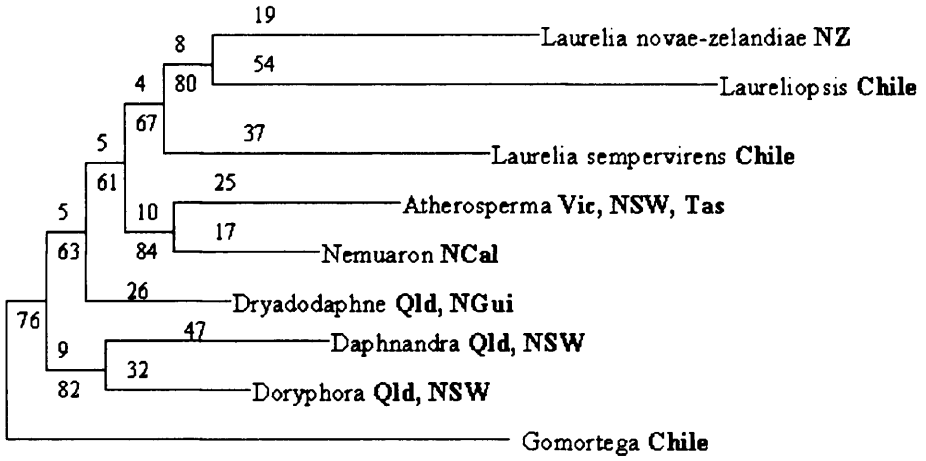

Fig. 2.3 Phylogram of the cpDNA parsimony tree for Atherospermataceae, where NCal = New Caledonia, NGui = New Guinea, NZ = New Zealand, NSW = New South Wales, Tas = Tasmania, Qld = Queensland, Vic = Victoria. Numbers above internodes are mutations and bootstrap values of >50% are included below the relevant internode (Redrawn from Renner *et al.*, 2000).

2.3.3 Nothofagus (Nothofagaceae)

One of the best-known examples of what is thought to be a relict Gondwanaland genus is *Nothofagus* (southern beech). This is a set of wind-pollinated, evergreen or deciduous trees. The genus has attracted much attention from taxonomists and biogeographers (Gandolfo and Romero, 1992; Hill and Read, 1991; Philipson and Philipson, 1988; Premoli, 1996; Rozefelds and Drinnan, 1994; Hanks and Fairbrothers, 1976; Melville, 1973; Dettman *et al.*, 1990; Hill, 1991; Humphries, 1981; Linder and Crisp, 1995; Martin and Dowd, 1993; Tanai, 1986; Van Steenis, 1971; Cranwell, 1963). All have sought to understand the pattern of variation circumscribed by its 35 species distributed between SE Asia, Australia, New Zealand and South America. Early work assumed that *Nothofagus* belonged in the Fagaceae and that there was a close relationship with *Fagus* (e.g., Kubitzki, 1993). This led several workers to erect phylogeographic hypotheses that involved an origin of *Nothofagus* from northern-Hemisphere *Fagus*-like stock. For example, Darlington (1965) thought that continental drift occurred too early to affect *Nothofagus* and so argued that many long-distance dispersal events had occurred from the northern Hemisphere to the southern Hemisphere; and Van Steenis (1962) suggested that *Nothofagus* arrived in South America via land bridges .

It has recently become clear, however, that the traditionally conceived Fagaceae are polyphyletic and that *Nothofagus* should be removed to its own family (Nixon, 1989; Manos *et al.*, 1993; Soltis *et al.*, 1995; Manos and Steele, 1997). Although there is still doubt about its outgroup. Recently, Manos (1997) presented a fully-resolved, well-supported phylogeny based on the morphological data of Hill and Read (1991), the *rbc*L cpDNA sequence data of Martin and Dowd (1993) and his own ITS nDNA sequences for the genus (Fig. 2.4). Inspection of the phylogenetic tree reveals four monophyletic groups, each recognised as a subgenus, *Fuscospora, Lophozonia, Nothofagus* and *Brassospora,* corresponding to the four extant pollen-types that had long been recognised. The area relationships of these subgenera are shown in Fig. 2.4 where it can be seen that there are three trans-Antarctic disjunctions. Two of these occur within subgenera *Fuscospora* and *Lophozonia,* and the third occurs between subgenus *Nothofagus* and subgenus *Brassospora.*

The fossil record shows that each of the four subgenera had a continuous austral distribution from Australia to South America by the Eocene (Dettman *et al.*, 1990). There are eight distinct pollen types in *Nothofagus*, four of which are found in extanct species (Dettman *et al.*, 1990). All four extant types have been found from the early Campanian of the late-Cretaceous (around 70 Mya) in both western Antarctica and South America. Slightly younger deposits in Australia and New Zealand have also been found to contain these four pollen types (Dettman *et al.*, 1990).

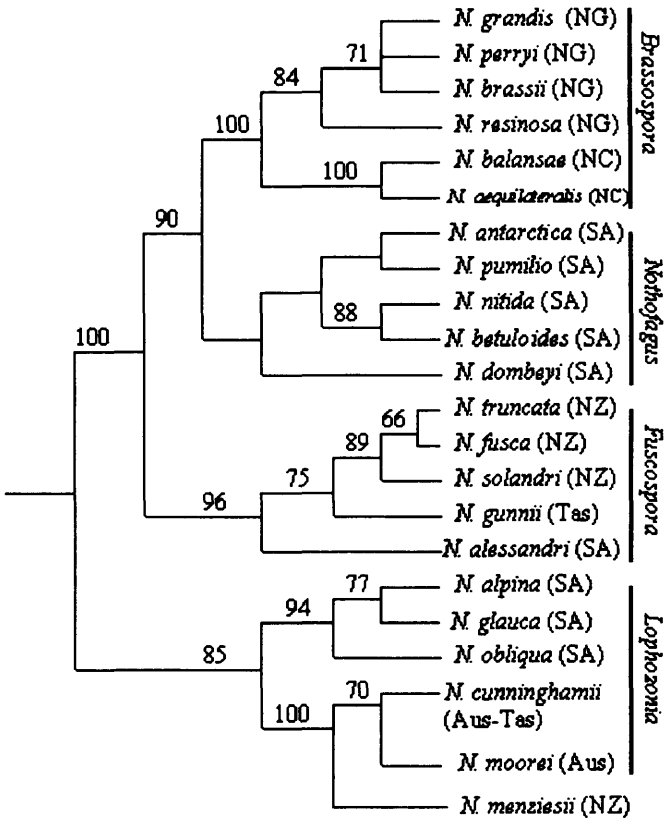

Fig. 2.4 Phylogeny of *Nothofagus* based on the morphological data of Hill and Read (1991), the *rbc*L cpDNA sequence data of Martin and Dowd (1993) and ITS nDNA data. Numbers above internodes are bootstrap values of >50%. Where SA = South America, Tas = Tasmania, Aus = Australia, NZ = New Zealand , NC = New Caledonia, NG = New Guinea. [Redrawn from Manos, 1997].

The distribution and diversity of pollen types, combined with the presence of much older pollen types (*ca.* 83 Mya), suggests that *Nothofagus* evolved and radiated in the southern Hemisphere. Thus, according to Manos (1997), the disjunctions present in subg. *Fuscospora* and *Lophozonia* are likely to be a result of vicariance events associated with the break-up of Gondwanaland. However, the area relationships within each of these two subgenera are (SA(AU,NZ)), a pattern that is incongruent with the area cladogram based on geological evidence (NZ(SA,AU)); see also Table 2.1. This discrepancy has given rise to dispersalist explanations for at least the *Nothofagus* connection between Australia and New Zealand. These views are supported by the fact that certain pollen types are absent

from early-Tertiary deposits of New Zealand, leading to the suggestion that *Nothofagus* species crossed the Tasman Sea from Australia several times in the early-mid-Tertiary (Mildenhall, 1980; Truswell *et al.*, 1987; Hill, 1994; Mcphail *et al.*, 1997). Proponents of a vicariance explanation argue that the absence of a fossil record is scarcely a strong basis for a dispersalist hypothesis, and prefer an ancestral presence followed by extinctions as a more parsimonious explanation. However, calibration of an *rbc*L cpDNA clock by Martin and Dowd (1993) using the first fossil *Nothofagus* pollen as a reference point (83 Mya) showed that the disjunction between South America, Australia and New Zealand within subg. *Lophozonia* and within subg. *Fuscuspora* occurred at a maximum of 66 Mya and 48 Mya, respectively. Both of these times are much less than the estimated dates of separation of Australia and New Zealand and are strongly suggestive of long-distance dispersal (Hill, 1991) (or of systematic errors in the estimates). In his recent paper, Manos (1997) did not apply a molecular clock and stated that "problems concerning the use of ITS sequences for biogeographic hypothesis testing include relatively low numbers of substitutions between trans-Antarctic species, a feature shared with *rbc*L sequences in comparisons of the same taxa (Martin and Dowd, 1993)." However, the fact that between the trans-Antarctic subgenera there are very low numbers of substitutions could alternatively suggest that the species have not been separated for the length of time needed to satisfy a vicariance explanation. Manos (1997) went on to suggest that the answer is to find regions of the chloroplast genome that have faster levels of evolution. It is important however to try to interpret the data as they stand instead of seeking further data to satisfy fixed views.

To explain the disjunction between the sister subgenera *Nothofagus* and *Brassospora,* Manos (1997) argued that the overall fossil evidence is best interpreted as supporting the extinction of subg. *Brassospora* from South America, Antarctica, New Zealand and Australia, and the extinction of subg. *Nothofagus* from Antarctica, New Zealand and Tasmania. The disjunction is thus one caused by massive extinctions, rather than by vicariance or dispersal events.

2.3.4 Proteaceae

The Proteaceae contain 79 genera and some 1700 species (Johnson, 1998). They are predominantly distributed in the southern Hemisphere where they are found in south-east Asia, Australia, New Caledonia, New Zealand, South Africa, Madagascar and South

America. Therefore they encompass an extended version of the Australasia/South America disjunctive distribution. The taxonomy of the family is complex and contentious (Hoot and Douglas, 1998), although a recent review, covering morphological and DNA evidence, showed that the family is clearly monophyletic (Johnson, 1998). Although the fossil history for the family is fragmentary, it can be deduced that the taxon dates back to the mid-Cretaceous at least.

Molecular phylogenetic work was published by Hoot and Douglas (1998) on the *atp*B gene and *atp*B-*rbc*L intergenic spacer sequences. The authors sequenced 46 species, including at least one species from each major subgroup of the Proteaceae. The study demonstrated that the family is monophyletic and that so also are the subfamilies within the Proteaceae; most tribes and subtribes, however, are not. The molecular data generally pointed to some unusual alliances within Proteaceae. Indeed, the strict consensus tree resulting from the molecular data does not agree with morphology or chromosome number. It does fit, however, with previous hypotheses of multiple aneuploidy and chromosome doubling events. In terms of biogeography, information from these data suggests that the major groups (subfamilies) within the Proteaceae diverged prior to or during the break-up of Gondwana in the early-mid-Cretaceous. This evidence is consistent with the fossil data. Seven cases of the South America/Australasia disjunction are shown in the strict consensus tree of the molecular data. To what extent these are explicable on the basis of vicariance is unknown, although one case requires special mention; *atp*B and spacer sequences from the Australian *Floydia* are *identical* to those in the South American *Roupala*. If this lack of variation is correct, this is surely suggestive of a recent common ancestry, one that post-dates the break-up of Gondwanaland and implies a dispersal event.

2.3.5 Abrotanella (Compositae)

Abrotanella contains, 19 species distributed between southern South America, New Guinea, Australia, Tasmania, New Zealand, Stewart Island (south of New Zealand), and the sub-Antarctic Campbell and Auckland Island (also south of New Zealand). Swenson and Bremer (1997) produced a well-supported phylogeny based on a range of morphological features that demonstrated the monophyly of the genus and revealed the relationships of its constituent lineages. Comparison of fully resolved taxon cladograms with area cladograms based on the break-up of Gondwanaland revealed various degrees of

mismatch, which led the authors to conclude that there must have been several instances of long-distance dispersal as well as vicariance events. In particular, the arrival of the genus in New Zealand was suggested to be a result of long-distance dispersal from South America, based on evidence from the cladograms and from the fossil record. The authors declined to put a date on this dispersal, except to say that it must have been "in comparatively recent geological times", "long after" the separation of New Zealand from other continents. In this regard the Asteraceae themselves are believed to be only *ca.* 40 million years old, and *Abrotanella* is a derived genus within the family, and so is younger still.

Several other examples of the South America/Australasia disjunction in the Compositae, viz. between sister species in *Cotula, Microseris, Centipeda, Lagenophora,* and other Senecioneae genera, may also have arisen from long-distance dispersal events by ancestors in South America (Bremer, 1992).

2.3.6 *Sophora microphylla* agg. (Leguminosae)

There are around 17 closely related taxa in the *Sophora microphylla* complex, distributed in Lord Howe Island, New Zealand, the Chatham Islands, Raivavae, Rapa, Marquesas, Masafuera, Masatierra, Chile, Easter Island, Gough Island and Reunion. The taxa are known to produce buoyant seeds which are viable after three years of immersion in sea-water (Sykes and Godley, 1968).

Hurr *et al.* (1999) used data from the intergenic spacer region, *atp*B-*rbc*L, to investigate relationships within the complex. The molecular clock for the gene was calibrated against the sister species *S. tomentosa*, whose first known fossils occur 30 Ma (Herendeen, 1992). One problem that does occur is that there is no rate consistency over the whole tree. However, the relationship between the substitution rate and time was demonstrated by the greater number of changes between older lineages within the group than within younger lineages. They found that *S. microphylla* differs from its sister species, *S. tomentosa,* by 14-15 base pairs which, assuming a rate of synonymous substitution based on the rate of evolution for the gene, indicates an origin for the complex at between 9.6 Mya and 8.9 Mya.

Although there were not enough mutations within the complex to estimate precisely the relationships of the constituent taxa, the 1-3 mutations that there are, are consistent with radiation over the past 0.5-2 million years. Hence, the trans-Antarctic disjunction must have occurred by means of long-distance dispersal, presumably with the aid of ocean currents, rather than by means of vicariance events.

2.3.7 Marsupials

The South American-Australasian distribution in marsupials is at the taxonomic level of order. Three extant orders are found only in South America (with a few recent immigrants in North America), and four occur in Australasia (Fig. 2.5). This geographic delimitation has strongly influenced the interpretation of morphological and anatomical characters of marsupials which in turn has had a large effect on their classification.

Three of the key features of marsupial taxonomy are illustrated. Diprodonty (a dental state with procumbent lower incisors) and syndactyly (where the second and third digits of the foot are joined), and the continuous lower ankle joint pattern (CLAJP) are all derived character states found in marsupials which are used for classification purposes. A further derived marsupial feature used in their classification is that of epididymal sperm pairing.

Springer *et al.* (1998) investigated the phylogenetic relationships within marsupial orders based on sequence data from multigene data sets including both mitochondrial and nuclear genes. The genes that they sequenced were 12S rRNA, tRNA valine, 16S rRNA, cytochrome b and IRBP (interphotoreceptor retinoid-binding protein) sequences. Relationships revealed by sequence data are also supported by single copy DNA hybridisation results (Kirsch *et al.*, 1991). The phylogenetic tree (Fig. 2.6) shows that there are two trans-Antarctic disjunctions at the order level within the marsupials.

Although marsupials are hailed as a classic example a southern Gondwanaland distribution, they are thought to have originated in Laurasia (Springer *et al.*, 1997). The oldest skeleton thought to be from a marsupial is around 101 million years old and was found in the Albian of Utah (Cifelli, 1993). Many definite fossils dating from the late-Cretaceous have also been found in North America (Cifelli, 1990 a,b, 1993; Eaton, 1993). Marsupial fossils from this period have also been found in Asia (Szalay, 1994; Trofimov

and Szalay, 1994). The first fossils found in South America date from the early-Palaeocene (Bonaparte, 1990; Gayet *et al.*, 1991). These fossils were also originally considered to date from the late-Cretaceous (Springer *et al.*, 1997). Similarly, although once thought to be considerably older, the oldest marsupial fossils found in Australia are now dated as late-Palaeocene or early Eocene (Godthelp *et al.*, 1992; Archer *et al.*, 1993). There is no geological or paleontological evidence for the presence of marsupials in Gondwanaland before the beginning of the Tertiary.

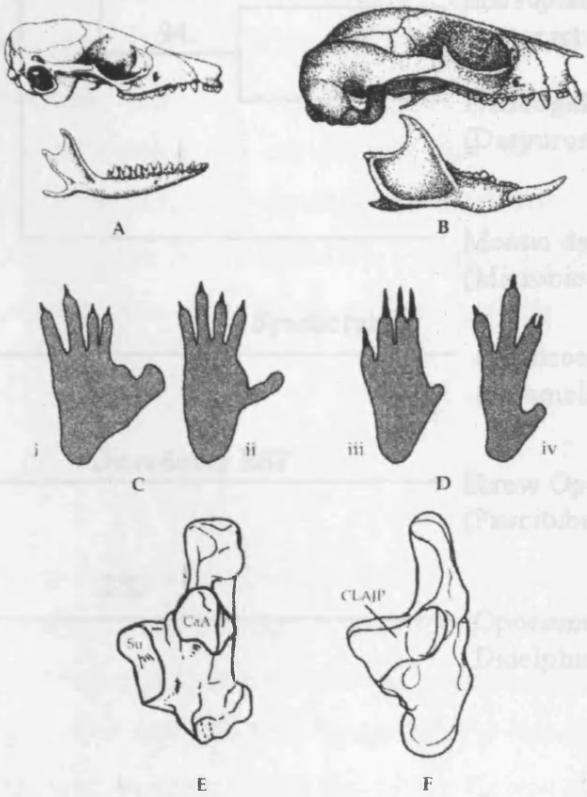

Fig. 2.5. Alternative states of three key marsupial characters; the presumed primitive state is shown on the left in each case. A,B, lateral view of the skulls and lower jaws of a polyprotodont and the diprodont condition respectively. C,D, hindfeet of the didactyl and the syndacty state respectively. E,F, the separate lower ankle joint pattern and the continuous lower ankle joint pattern. (Picture from Springer *et al.*, 1997).

It is thought that marsupials migrated from South America into Australasia. However, in order to explain the fact that the montio del monte is nested inside the Australasian marsupials, there must have been a back-migration event of an ancestor from this lineage from Australia to South America. In order for this to be the case, there must have been

microbiotheres, or their ancestors in both Australia and Antarctica with subsequent extinctions on both continents.

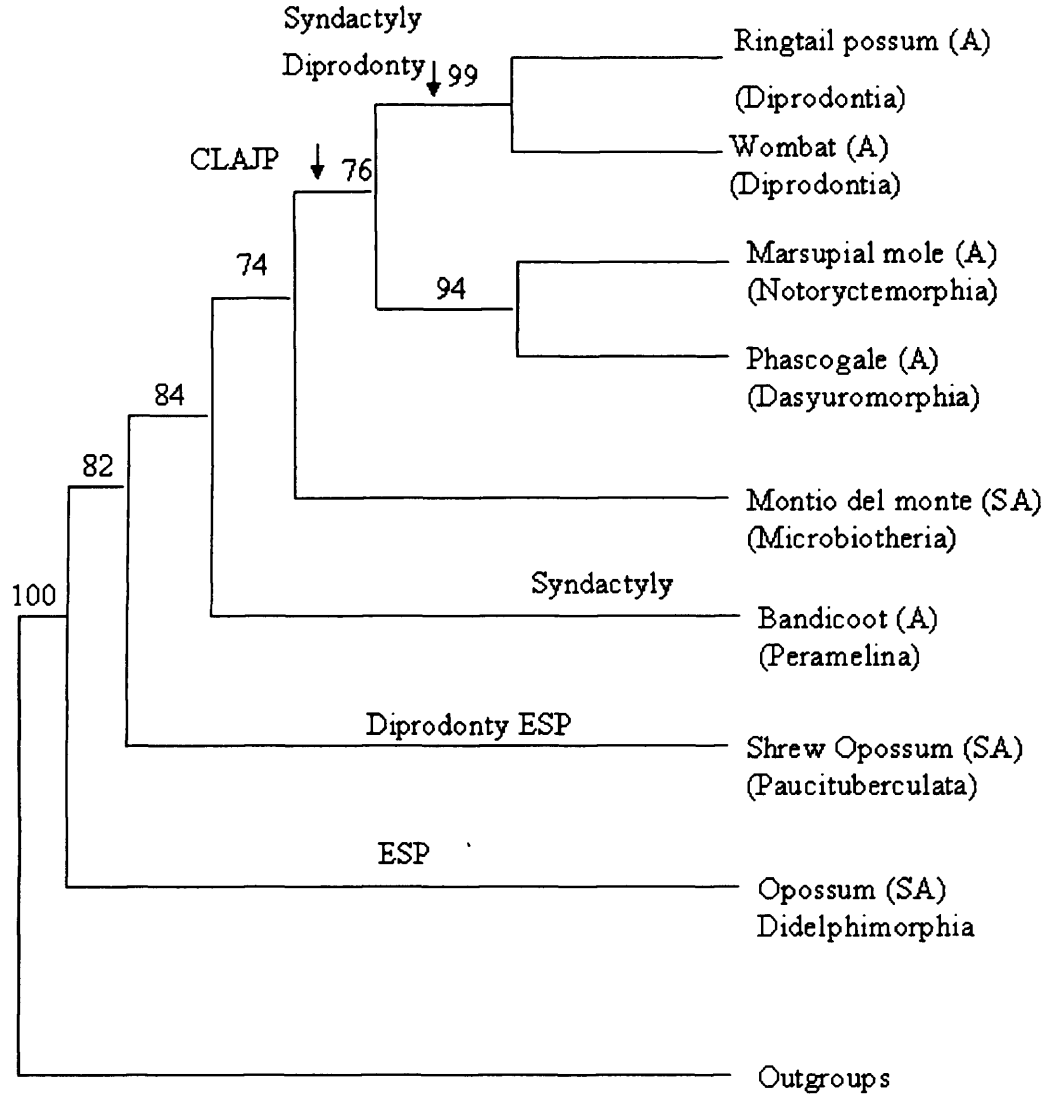

Fig. 2.6 Maximum likelihood bootstrap tree showing clades supported above the 50% level. The tree is based on a combination of 12S rRNA, tRNA valine, 16S rRNA, cytochrome b and IRBP sequences. N.B., the outgroups were human, cow, horse and mouse. Abbreviations: A, Australasia; SA, South America; CLAJP, continuous lower ankle joint pattern; ESP, epididymal sperm pairing. In all cases the derived state has been marked on the individual clades, whereas for simplicity, the alternative primitive state (Fig 2.5) has been left off the diagram (Redrawn from Springer *et al.*, 1998, with added information included from Springer, 1997, 1998).

Molecular work has provided a useful framework in which the fossil and anatomical marsupial data can be interpreted. Work by Springer *et al.* (1998) and Kirsch *et al.* (1997) suggests that the distribution of marsupial orders within Australia and South America is

indeed due to vicariance. The story most consistent with the data is that before the fragmentation of Gondwanaland the marsupials migrated from their centre of origin in Laurasia to South America. One lineage of what was to become the australidelphians (Australasian taxa + the South American Microbiotheria) then migrated to Australia and Antarctica. This lineage diverged in Australia and Antarctica to form at least two of the Australasian lineages (the proto-bandicoot and the proto-eometatherian lineage) before Australia became separated from Antarctica at around 64 million years ago. Antarctica seems to have been the theatre in which marsupials diverged and, before it became separated on either side from South America and Australia, there were multiple dispersals of australidelphian marsupials back into Australia, as well as at least one migration back into South America. This explains the paraphyletic distribution of characters in the australidelphian marsupials. In summary, therefore the distribution of marsupial orders in Australia and South America can be explained by a number of pre-Gondwanaland split dispersal events followed by the splitting up of the continents and by a number of extinctions.

2.3.8 Ratite birds

Ratite birds have traditionally been considered to be an example of a distribution that is disjunct at the family level. Living members of this group are shown below in Fig. 2.7. Morphological data had been misleading from an evolutionary standpoint primarily because anatomical analysis of ratite birds is fraught with problems due to difficulties in determining the polarity of characters (Kurochkin, 1995). Nevertheless, several molecular studies have shown that the living ratites are indeed monophyletic (as had been supposed) and that their closest living relatives are the weakly flying tinamous from South America (van Tuinen *et al.*, 1998).

Although it is agreed that the ratite birds constitute a monophyletic group, the relationships within this group are contentious (van Tuinen *et al.*, 1998). An understanding of these relationships is necessary to interpret the biogeography of these birds. Results from DNA hybridisation studies and from sequence analysis of some protein-coding genes were conflicting (van Tuinen *et al.*, 1998). Therefore, in order to try and resolve the relationships of the three major lineages of ratites: rheas (from South America), ostriches (from Africa) and Australian ratites, van Tuinen *et al.* (1998) sequenced an entire 2.8 kb

region of mitochondrial DNA. Their data suggested that the African ostrich is basal to the South American rhea which in turn is sister to the Australian emu and cassawary (Fig. 2.8). The mtDNA sequences agree with DNA hybridization data and with immunological data, but conflict with some earlier data based on protein-coding sequences. It has been suggested however that such sequences, e.g., cytochrome b, are not useful for dating divergences that are earlier than the Miocene (Moore and DeFilippis, 1997). This is thought to be due to transversion saturation, base composition bias and rate variation among lineages.

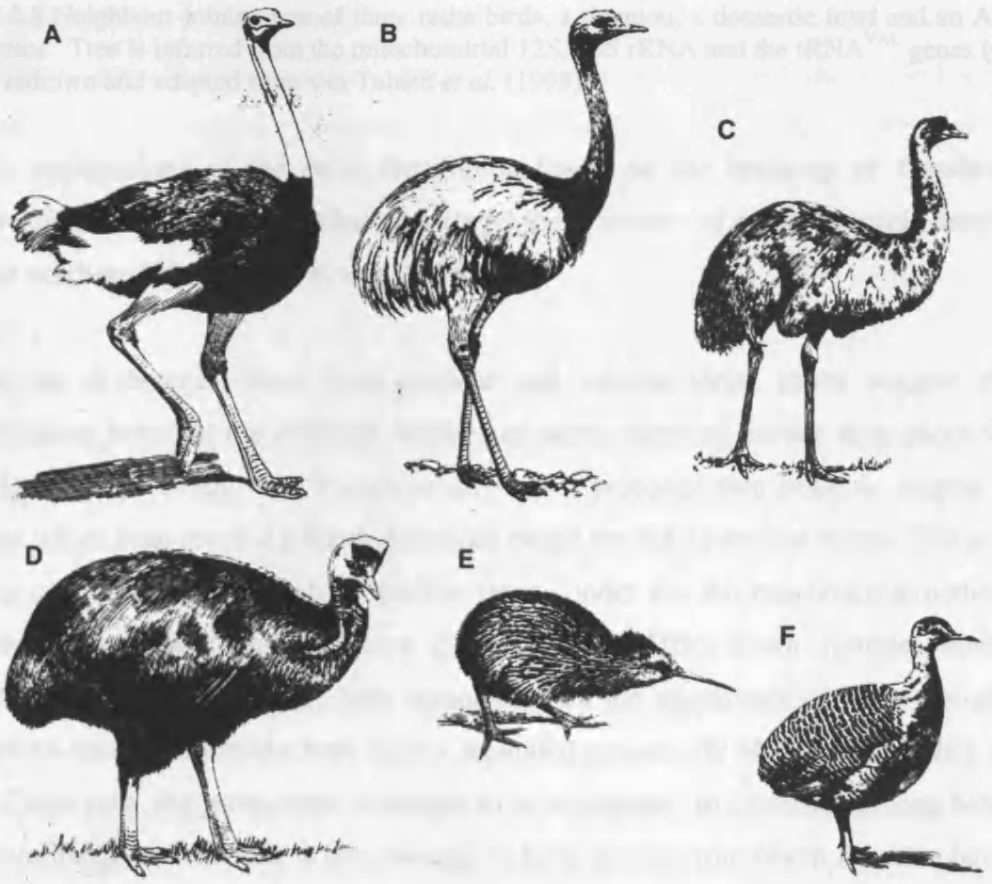

Fig. 2.7 Families of livingratite birds. (A) Struthionidae, ostrich from Africa; (B) Rheidae, rhea from South America; (C) Dromiceiidae, emu from Australia; (D) Casuariidae cassowary from Australia and New Guinea; (E) Apterygidae, kiwi from New Zealand; (F) Timamidae (tinamou) from tropical America. Not drawn to scale, taken from Futuyma, (1986).

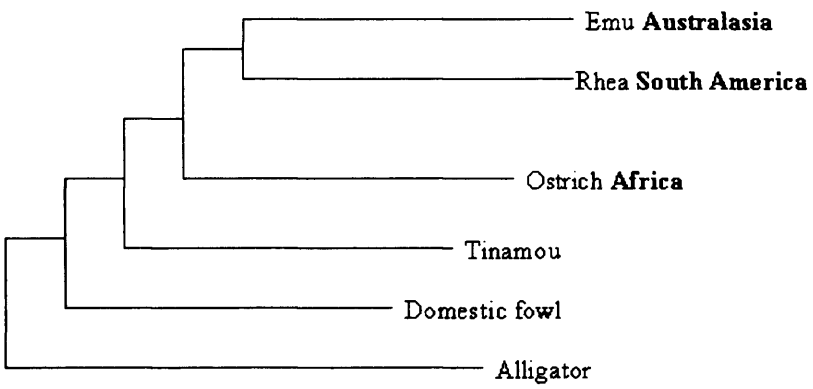

Fig. 2.8 Neighbour-joining tree of three ratite birds, a tinamou, a domestic fowl and an American alligator. Tree is inferred from the mitochondrial 12S, 16S rRNA and the $tRNA^{VAL}$ genes (total 2.8 kb), redrawn and adapted from van Tuinen *et al.* (1998).

Early explanations of the ratite distribution based on the break-up of Gondwanaland (Cracraft, 1974) were confounded recently by the discovery of early-Cenozoic ratite fossils in the northern Hemisphere (Houde, 1986).

Tests on divergence times from nuclear and mitochondrial genes suggest that the divergences between the different families of ratites were no earlier than about 90 Mya (Hedges *et al.*, 1996). Van Tuinen *et al.* (1998) proposed two possible origins for the ratites which both involve a South American origin for the Australian ratites. The scenarios differ only in the location of the earliest ratite. Under the African-origin hypothesis, the ancestral birds are thought to have existed on the Africa-South America land mass. Vicariance is thought to have been responsible for the appearance of the proto-ratite and the proto-tinamou when the land masses separated around 100 Mya. Between then and the late-Cretaceous, the proto-ratite is thought to have migrated to Laurasia, leaving behind the ostrich lineage in Africa. It is then thought to have moved from North America into South America in the late-Cretaceous to establish the rhea lineage which, shortly after, migrated to Australia across Antarctica via the Scotia land bridge.

Under the alternative hypothesis of a South American origin, the ratite lineage is thought to have originated in South America after it separated from Africa. After the diversification of the tinamous, an ostrich lineage may have arisen by migration northwards, across the proto-Antilles to North America in the late-Cretaceous and then subsequently through

Laurasia and Africa. The South American stock is also thought to have migrated across Antarctica via the Scotia land bridge to Australia to produce the emus and cassowaries.

The New Zealand kiwi is sister to the Australian ratites (van Tuinen *et al.*, 1998) and may have had an origin by means of vicariance from an Antarctic stock or by means of an early dispersal from Australia. The New Zealand moas may have a similar history but good data are lacking. The Madagascan elephant birds, in contrast, most likely dispersed there from ostrich-like stock in Africa after the separation of Madagascar in the early-Cretaceous. The oldest fossil ostrich has been found in Africa and dated to the early-Miocene (Mourer-Chauvire *et al.*, 1996).

In conclusion therefore, the ratite birds form a monophyletic group which, if either of the two above hypotheses are correct, spread widely before the break-up of Gondwanaland. Therefore their subsequent diversification and distribution is consistant with vicariance. Hopefully as more work is carried out, the discrepancy between data from the mitochondrial and nuclear genomes will be understood.

2.3.9 Fish

The best studied example of the trans-Antarctic disjunction in fish concerns *Galaxias maculatus* (Galaxiidae). This species is a small diadromous fish found in the streams of eastern and western Australia, New Zealand, South America and some Pacific oceanic islands (Berra *et al.*, 1997).

Rosen (1978) believed that galaxiid fishes owed their distribution to vicariance events and that they were originally part of a pan-austral Gondwana biota. The alternative explanation of dispersal was suggested by McDowall (1978) in the same issue of *Systematic Zoology*. *G. maculatus* is the only galaxiid that breeds in brackish water; its larvae grow in the ocean before returning to freshwater. McDowell therefore suggested that a relatively recent dispersal event at the juvenile stage was the most likely explanation for the distribution. Recently collected allozyme data showed very little differentiation between populations of *G. maculatus* from South America and Australia, and the consequent small genetic distance thus supports dispersal as an explanation (Berra *et al.*, 1997). In further work, a molecular phylogeny was recently constructed from mitochondrial cytochrome b and 16S

RNA gene sequences (Waters and Burridge, 1999). Sequences were obtained from six populations of *Galaxias maculatus* from Tasmania to New Zealand and from Chile to the Falkland Islands. Divergence for this disjunction was a maximum of 14.6% for cytochrome b and a maximum of 6% for 16S RNA. However, divergence between Tasmanian and New Zealand haplotypes was also high (maximum 5.1%) and within the Chilean and Falkland Islands populations divergence was up to 3.8%. Molecular clock calibrations suggest that this pattern of haplotypes post-dates the fragmentation of Gondwanaland (Waters and Burridge, 1999). From these findings, therefore, the authors favoured marine dispersal rather than vicariance as an explanation for the distribution.

2.4 Conclusions

In drawing conclusions from this Chapter, it is clear that there is no single explanation for trans-Antarctic disjunctions. Superficially similar patterns of distribution, on close examination, turn out to have different histories and time scales. When we are dealing with marsupial disjunction, at the taxonomic level of order, vicariance is indeed a convincing explanation for the disjunction. Similiary, at the level of family as in the case of the ratite birds, an interpretation from molecular and fossil evidence has strongly suggested that vicariance can explain the distribution. Vicariance may also be a the most convincing explanation for the distribution of the species within *Araucaria* and for the major groups in Proteaceae. To what extent vicariance can explain disjunctions in the Proteaceae below the level of subfamily is not yet known.

In the Atherospermataceae, where more extensive molecular work has been carried out, vicariance can explain one of the disjunctions but not the other which must be attributable to dispersal. The story regarding *Nothofagus* is also complicated and, some would argue, still unresolved. The available evidence is consistent with some vicariance events although at least the New Zealand species appear to have originated following long-distance dispersal. *Abrotanella* (Compositae) also probably owes its distribution to a combination of dispersal and vicariance events. The *Sophora microphylla* complex (Leguminosae) provides us with a well documented example of a distribution which is attributable to dispersal by oceanic currents. Similarly, the disjunct pattern apparent in the fish *Galaxias maculatus* has also to be explained in terms of recent oceanic dispersal events.

This review has shown that the disjunct distribution patterns exhibited at different levels of the taxonomic hierarchy clearly relate to different periods in time. In particular, disjunctions involving higher level ranks are likely to have vicariance as a major component of their explanation. In contrast, disjunctions involving genera and, especially, species may post-date the break-up of Gondwanaland, and dispersal must be invoked. The rank of genus appears to sit at the cusp of this dichotomy and is therefore of special interest. The review has also highlighted the importance not only of fossil evidence, but also of establishing monophyly, and of the new insights that may be gained from gene sequences.

CHAPTER 3

DNA SEQUENCE EVOLUTION

The point has already been made that in order to understand the phytogeographical distribution of a plant group, it is necessary to understand its phylogeny. This project aimed to assess the utility of different gene sequences for use in providing phylogenetic information for species within a genus. Too often, this type of work is based on information from a single gene. In many cases, however, it has been shown that different genes produce different phylogenetic trees (Renner and Chanderbali, 2000, Qui *et al.*, 1999). One of the aims of this project was to compare the phylogenetic behaviour of different gene sequences (chloroplast, mitochondrial and nuclear) and their relative rates of evolution.

3.1 Chloroplast genomes

The chloroplast genome is a circular molecule which typically has two inverted repeat regions that separate a large and a small single-copy section (Fig. 3.1). The analysis of chloroplast DNA has been particularly useful for resolving phylogenies above the genus level as, on average, it evolves relatively slowly (Downie and Palmer, 1992; Clegg and Zurawski, 1992). Many authors (e.g. Soltis *et al.*, 1992; Gillham, 1994) however have pointed out, that cpDNA variation can also exist at the intraspecific level, similar to the way that mitochondrial DNA variation occurs among animal populations and likewise can be used in order to understand evolution.

3.1.1 Advantages of cpDNA

The chloroplast genome is relatively small, typically between 120 and 200 kb which makes it relatively easy to study the whole genome using restriction site analysis (Soltis *et al.*, 1992). Most of the genes, including those discussed below, are single copy in contrast to most nuclear genes which belong in multigene families.

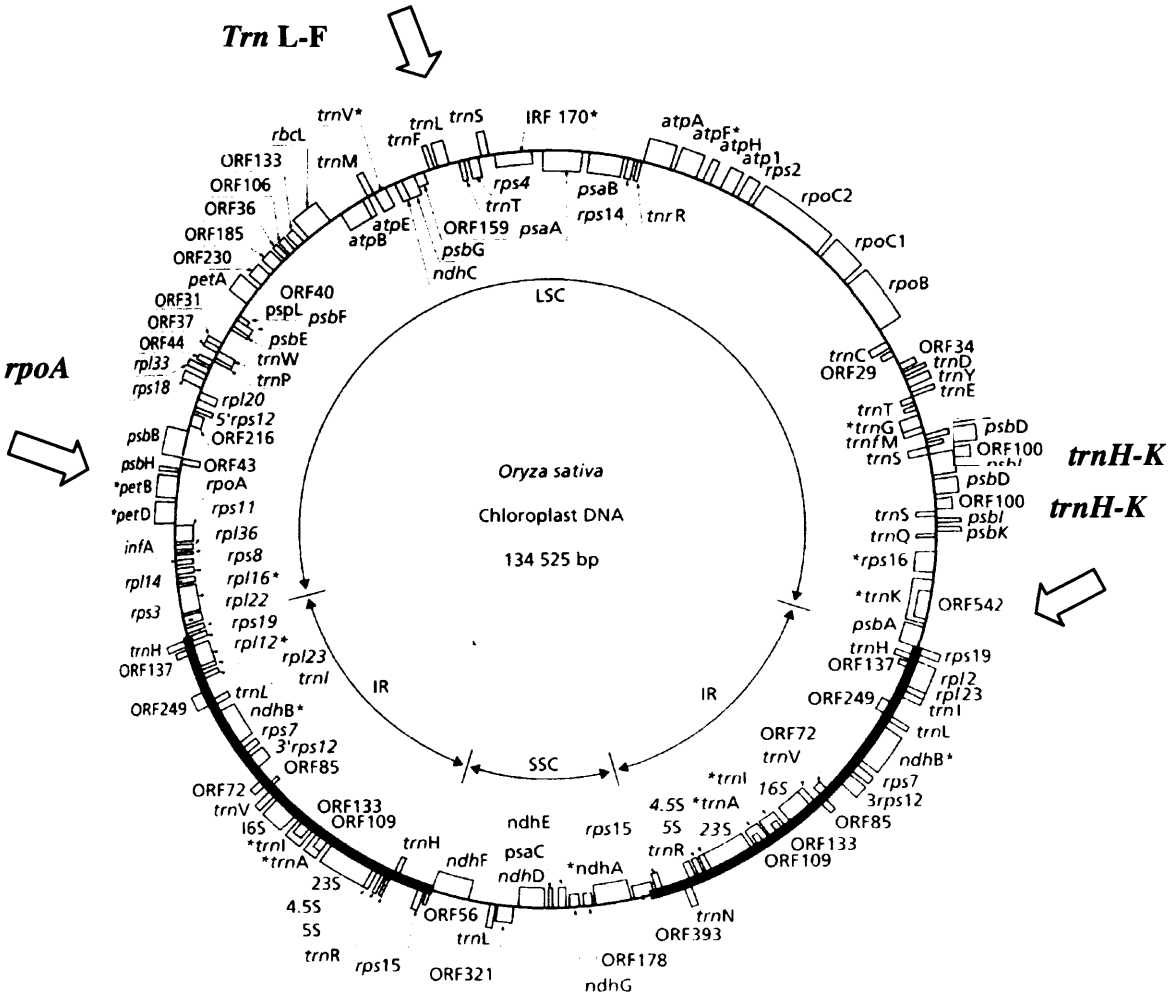

Fig. 3.1 The chloroplast genome (cpDNA) of rice (*Oryza sativa*). Genes are located on two strands with those on the outside transcribed anticlockwise and those on the inside transcribed in a clockwise direction. The bold regions indicate the inverted repeats (IR) and the large and small single-copy (LSC and SSC) DNA are above and below respectively. From Page and Holmes (1998). Regions of DNA used in this thesis are indicated with arrows.

The genome is both structurally conservative and evolves relatively slowly at the nucleotide sequence level (Downie and Palmer, 1992). The genome is generally uniparentally inherited, predominantly through the maternal line in angiosperms (Page and Holmes, 1998).

Analysis of mutational changes and rearrangement patterns are largely uncomplicated by recombination. This contrasts with nuclear chromosomes which normally recombine at meiosis (Page and Holmes, 1998).

All these factors combine to increase the ease of determining orthology and if phylogenetic analysis is to reconstruct accurately the history of an organism, orthologous sequences need to be compared. That is, the sequences must share the same history. Two homologous genes are orthologous if their most recent common ancestor did not undergo gene duplication, otherwise they are termed paralogous.

3.1.2 Disadvantages of cpDNA

The main disadvantage of cpDNA is that the conserved nature has meant that sometimes it has not been able to resolve relationships between closely related species or at the level of the population (Soltis and Soltis, 1998). However, in some cases the assumption of low levels of chloroplast variation is not at all true (Harris and Ingram, 1991). Indeed, cpDNA variation between populations of a species has been used to interpret post-glacial migrations (Ferris *et al.*, 1993).

Another disadvantage of the chloroplast genome is the occurrence of introgression, whereby the chloroplast genome may be moved from one species to another. This problem has been extensively reviewed (e.g., Rieseberg and Brunsfeld, 1992). Although this phenomenon is problematic, when it is identified then the information can be useful. If undiagnosed, such chloroplast transfer can inadvertently lead to an inadequate assessment of species relationships. A case in point is shown by Ferris *et al.* (1997), who studied *Spartina anglica*, a natural amphidiploid between *S. alterniflora* and *S. maritima*. As chloroplasts are inherited maternally in 70% of angiosperms, a chloroplast phylogeny does not reflect the species but the history of the gene in the maternal *S. alterniflora*. It is therefore very useful to have other lines of evidence from different genes.

Other complications that can arise from using cpDNA for phylogenetic analysis stem from a strong codon bias (particularly for A- and T-ending codons). 70% of all the nucleotides in chloroplasts are A or T, and this therefore does not reflect random mutation. In terms of practical difficulties, this makes the design of primers difficult and the actual sequencing

itself can be problematic due to slippage caused by runs of As and Ts (Kruglyak *et al.*, 2000). In the noncoding regions, patterns of nucleotide substitution are strongly influenced by the base composition at adjacent sites; in other words there are internal processes that affect the molecules, so again mutation is not random.

3.1.3 Examples of commonly used regions of cpDNA

rbcL is the most commonly used chloroplast gene; it has been very useful in studies of relationships among plant families (Chase *et al.*, 1993) and can sometimes provide substantial resolution of genera or even species (for reviews see Olmstead and Palmer., 1994, Chase and Albert, 1998). The *rbc*L gene is positioned in the large single copy region of the chloroplast genome (Fig. 3.1). It codes for the large unit of ribulose 1,5 bisphosphate carboxylase/oxygenase (Rubisco) and is typically between 1,428-1,434 bp in length with insertions or deletions being rare (Soltis and Soltis, 1998). However at lower taxonomic levels *rbcL* often does not have enough variation for robust resolution of relationships.

More rapidly evolving chloroplast genes that have been used at lower taxonomic levels include *ndhF* (Olmstead and Reeves, 1995), *matK* (Johnson and Soltis, 1995; Xiang *et al.* 1998), and the non-coding spacer region from the transfer RNA gene *trnL-trnF* (Gielly and Taberlet, 1994, 1996; McDade and Moody, 1999). Non-coding regions are presumed to be more useful at lower taxonomic ranks because they are less functionally constrained and are therefore freer to vary and so provide more phylogenetically useful sites.

3.2 Mitochondrial genomes

Like the chloroplast genome, the plant mitochondrial genome is inherited uniparentally, mainly through the female line. Although extensive structural reorganisation is observed in mitochondrial genomes, particularly with respect to gene order, little variation in gene and intron sequences at the species level has been observed, thus making mtDNA unsuitable for phylogenetic studies at this level (Laroche *et al.*, 1997). Instead, the mitochondrial genome has been useful for inferring ancient phylogenetic relationships and in estimating

the time of early diversification events in seed plants. Laroche *et al.* (1997) pointed out that as well as looking at structural arrangements and changes in the genetic constitution of mitochondrial genomes, a third potential level of heterogeneity is between mitochondrial genes and their paralogs which are transferred to the nuclear genome. Variation can be looked for both in pseudogenes in the mitochondrial genome and in the deleted paralogs.

3.2.1 Advantages of mitochondrial genomes

The mitochondrial genome is potentially useful as it represents a thus far under-exploited source of molecular variation in understanding plant relationships at the level of genus and above. The abundance of mitochondria in cells enables the genome to be examined with relative ease. In a few cases they also have the potential advantage of being inherited differently from the chloroplast genome, as in the case of some conifers where the chloroplast genome is inherited paternally and the mitochondrial genome maternally (Isoda *et al.*, 2000).

Importantly, the rate of nucleotide substitution is approximately one and two orders of magnitude respectively slower than the chloroplast and nuclear genomes of plants (Wolfe *et al.*, 1987, 1989). Hence mitochondrial genome can be used to investigate higher taxonomic levels. The structural rearrangements and loss of genes and introns can easily be observed through restriction analysis (Palmer, 1992). One advantage of mtDNA over similarly slowly evolving regions of the chloroplast genome is that homoplasy is thought to be less prevalent (Soltis and Soltis, 1998).

Despite the traditional view of the plant mitochondrial genome evolving extremely slowly, recent studies have suggested that parts of the genome may be evolving between 50 – 100 times faster than the rest of the genome (Parkinson, 1999).

3.2.2 Disadvantages of the mitochondrial genome

The reasons why the plant mitochondrial genome has not been used much in reconstructing plant phylogenies are listed in Soltis and Soltis (1998) and summarised here. The first is that in contrast to the small, stable, animal mitochondrial DNA, plant mtDNAs are much larger and more variable in size. The genomes include many foreign DNA sequences,

particularly those derived from the chloroplast genome. Large numbers of duplications occur and then are subsequently lost. Extensive recombination is thought to happen between these repeats thus creating a very complex genome structure. There are also many shorter (50-1000 bp) dispersed repeats which are scattered throughout the mitochondrial genome. A further complication is the presence of small (1-11kb) unstable extra chromosomal plasmids of unknow-n origin and function are common in plant mitochondrial DNA. Rearrangement in plant mitochondrial genomes is so rapid that even closely related species do not have the same order of genes

The disadvantages of studying the mitochondrial genome are exemplified by Laroche *et al.* (1997). These workers compared the modes and tempo of the evolution of mitochondrial genes and introns, by analysing 15-protein coding genes and six introns of 29 flowering plant species. They found that the number of synonymous changes (Ks) was six times higher than the number of nonsynonymous (Ka) changes. The proportion of synonymous changes however varied less from gene to gene than the nonsynonymous substitutions. In general, Ks dropped by an order of magnitude for comparisons within a family, confirming that the number of potentially informative sites in exon sequences is too low to be useful at a low taxonomic level, (Laroche *et al.* 1997).

3.2.3 *Examples of commonly used regions of mitochondrial DNA*

The mitochondrial genome has been extensively used in animal phylogenetic studies particularly at the infraspecific level (see reviews by Avise, 1986, 1994). The mitochondrial genes which have received the most attention in plants are *coxI* and *atpA*. These have been amplified successfully in most of the angiosperms that have been studied. They are similar both in their lengths (1,574 and 1,592 bp respectively) and in their rates of evolution, which is approximately 0.43 and 0.45, respectively of the rate of evolution of *rbc*L.

3.3 Nuclear genome: ITS sequences

There are three types of rDNA sequence which are used commonly for phylogenetic purposes in plants. There is a small subunit 18S which is 1800 bp, the large subunit 26S

which is around 3400 bp, and the small 5.8S subunit of 160 bp. As well as these coding sequences, there are two main "spacer" sequences (ITS-1 and ITS-2) which contain the signals that are needed to process the rRNA transcript (Fig. 3.2). A single RNA transcript is produced from this unit. The number of copies of this unit varies greatly between taxa, for example the protist *Tetrahymena* has only one copy whereas the lizard *Amphiuma means*, has up to 19300 copies (Page and Holmes, 1998).

Ribosomal DNA is useful at many different taxonomic levels, as it includes both highly conserved (18S, 5.8S, 26S) and highly variable (ITS) sequences, with the latter being particularly valuable at resolving relationships at the species level.

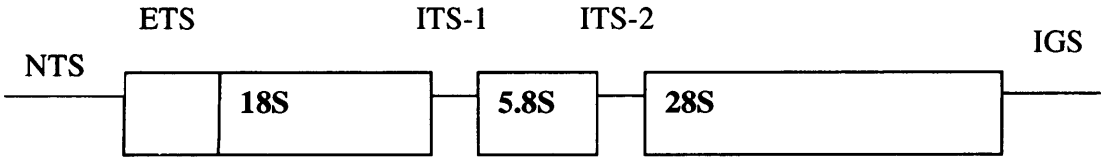

Fig. 3.2 A schematic diagram of the 18S – 28S rDNA in plants. The coding regions 18S, 5.8S and 28 S are shown by blocks and they are separated by the three regions: ETS (external transcribed spacer), ITS-1 and ITS-2 (internal transcribed spacer). At the extreme right is the intergenic spacer (IGS). The tandemly repeated arrays are separated from each other by a non-transcribed spacer (NTS). Redrawn from Soltis and Soltis (1998).

3.3.1 Advantages of ITS sequences

The most widely used nuclear regions in phylogenetic analysis for comparisons among and within genera are the ITS regions of rDNA. Reasons for the extensive concentration on the ITS regions have been much reviewed (Hamby and Zimmer, 1992; Sanderson and Doyle, 1992; Nickrent and Soltis, 1995; Baldwin *et al.* 1995).

The ribosomal RNA genes became widely used as a source of sequence data after it was realised that concerted evolution homogenises sequences so that an entire array of tandemly repeated rDNA cistrons evolves as a single locus (Hillis and Dixon, 1991). This process is known as gene conversion and occurs when the DNA sequence of one gene is replaced by that from another. This is thought to occur over DNA sequences of a few base pairs to many kilobases, and between different genes located in the same or on different chromosomes (or between alleles of the same gene) (Page and Holmes, 1998). Although

the mechanisms of concerted evolution are not known for certain, they are thought to involve gene conversion and unequal crossing over (Liao, 2000). Whatever the process involved, homogenisation depends on the extent of similarity among sequences; the more similar the sequences are, the more likely conversion will occur (Soltis and Soltis, 1998). Therefore conversion is more common in coding genes, because of the functional constraints on these genes which preserve similarity, compared to intron sequences. Similarily, newly duplicated genes, which will be very similar in sequence, are more likely to undergo conversion than genes that diverged longer ago. Anything that disturbs sequence similarity, such as the insertion of a transposable element, will lower the chance of conversion happening.

Some authors, however, have recently suggested that concerted evolution does not occur (de Miera and de la Vega, 1998). The ITS region can therefore be considered to be a population of genes rather than a single gene. When an entire PCR product is sequenced, the commonest sequence will dominate. On the electropherogram, it is common to see small peaks beneath the largest peak. This may be because of other loci of ITS that exist (de Miera and de la Vega, 1998). The only way to get around this problem is to clone the different ITS loci from the initial PCR product. Variation can exist both within the different copies at each locus, and between the different copies at the alternative locations of the ITS region (Hershkovitz and Zimmer 2000).

3.3.2 Disadvantages of ITS sequences

Firstly, in polyploids, inter-locus concerted evolution may homogenise homologous rDNA loci so that only a single one is retained. This locus may derive from any of the genomes present. Different lineages within a species therefore may have different rDNA sequences depending on the genome that was proliferated (Page and Holmes, 1998). Secondly, pseudogenes may persist within the genome and be preferentially amplified by PCR (Page and Holmes, 1998). Thirdly, there are problems both due to the presence of more than one locus, and due to variation within the different copies at each locus (Page and Holmes, 1998). Finally, it is becoming commoner for intraspecific and even intra-individual variation to be found (Hershkovitz and Zimmer, 2000), which makes it more difficult to produce phylogenies.

In summary, although concerted evolution allows genes to evolve together, it also greatly complicates the phylogenetic analysis of multigene families because it becomes difficult to discern which genes are truly homologous. It also means that the point of sequence divergence corresponds to the time of the last gene conversion event, or unequal crossing over, rather than to the time of the last gene duplication or speciation event. The complex processes that shape the evolution of multigene families such as the ribosomal DNA unit, mean that they are often composed of "mosaics" of sequences, each with a different phylogenetic history, rather than strictly homologous genes gradually diverging through time.

3.3.3 Examples of the utility of ITS

The literature is replete with phylogenies based on ITS spacer regions. A search on ITS-1 and phylogeny (in March 2001) found 2449 hits. ITS phylogenies can be particulary useful for understanding relationships between genera, one example of this is the increased understanding of the Carmichaelinae (Leguminosae) in New Zealand after an ITS study was conducted (Wagstaff *et al.*, 1999). A good further example relevant to this study is the ITS phylogeny of the Nothofagaceae (Manos, 1997).

3.4 Nuclear genome: low copy-number nuclear sequences

Although low copy nuclear genes potentially provide a source of useful information, they have thus far been under utilised (Soltis and Soltis, 1998).

3.4.1 Advantages of low copy-number nuclear sequences

The noncoding regions of low copy number nuclear genes can potentially provide multiple, unlinked allele genealogies at the species level. They are free from the constraints described for both chloroplast and the ITS genes. In other words, they are a potentially uncomplicated source of low copy gene sequence which is inherited bi-parentally and evolves at a fast enough rate to help resolve relationships at the level of the species. One part of this project was to try and exploit this and try to understand the evolution of novel, single or low copy nuclear genes.

3.4.2 Disadvantages of low copy sequences

There are difficulties because low copy nuclear genes have so much less starting material for PCR. Furthermore, because low copy nuclear genes are generally in families of genes that have similar coding sequence, it is often difficult to amplify single products. There is a further set of problems relating to studying genes on which there is thus far very little information. These facts render the design of primers troublesome, as often there are very few plant species available in sequence data banks from which primers may be designed. This information is necessary as primers, particularly those that will work for more than one genus need to be in conserved coding regions but preferably span noncoding introns or spacer regions. As more work is carried out on the exploitation of low copy nuclear genes it seems likely that these problems will diminish. The practical problems I encountered with working with low copy nuclear genes are discussed in Chapter 5.

3.4.3 Examples of the utility of low copy-number nuclear genes

Nuclear genes which have been investigated thus far include the phytochrome gene family (PHY) which has been shown to be useful at many taxonomic levels (Mathews *et al.*, 1995; Mathews and Sharrock, 1997), the ADH gene family which has shown to be useful at the level of the genus (Kosuge *et al.*, 1997) and the G3PDH gene family which has been shown to be useful at the species level (Olsen and Schaal, 1999). All three of these gene families were investigated in this project and are discussed in detail below.

The project started with an investigation of the phylogenetic potential of phytochrome genes. Work by Sharrock and Mathews (1997) has shown that these genes are useful in understanding evolution within the grasses. The five phytochrome genes in the family are termed PHYA - E. Although it is possible that other PHY genes are present, low stringency hybridization and degenerate primer PCR analyses have not found any evidence of this (Mathews and Sharrock, 1997). They have different physiological functions and are under different regulatory control stimuli. Four different protein types are produced from these genes, one from each of the genes PHYA, PHYC and PHYE, and the final type coming from the more closely related PHYB and PHYD genes.

The PHY genes are highly conserved, all containing short introns at homologous positions which divide the genes into three coding exons. The PHY genes A, B/D and E have a third much longer intron which is found towards the end of the final coding exon. All genes have a large untranslated region at the 5' end of approximately 2000 base pairs. This region in PHYA differs from the others in the family because it has an intron in it and it contains many transcription start sites. These genes have been used in phylogenetic studies by Mathews and Sharrock (1997) and Lavin *et al.* (1998). In both of these studies conserved regions of the first coding region were sequenced, and were found to give good phylogenetic resolution at the genus level. It is possible that the large intron will be a lot more interesting and give resolution at the lower taxonomic level (Mathews, pers. comm.).

The alcohol dehydrogenase (*adh*) gene (Fig. 3.3) produces an essential enzyme in anaerobic metabolism. The gene family is relatively well studied and therefore a logical choice for further work. The family has 2-3 loci in most plants and one in *Arabidopsis*.

One example of the utility of this gene is provided by Charlesworth (1998), who studied its phylogenetic behaviour in *Leavenworthia* (Brassicaceae) and showed that it provided much useful data on evolution in the genus. In *Leavenworthia* the three gene loci evolve as is shown in Fig. 3.3, with a progressive loss of introns. All three genes are expressed (despite no introns in the third allele type). Furthermore, all three loci are not thought to be closely linked to each other (Charlesworth, 1998).

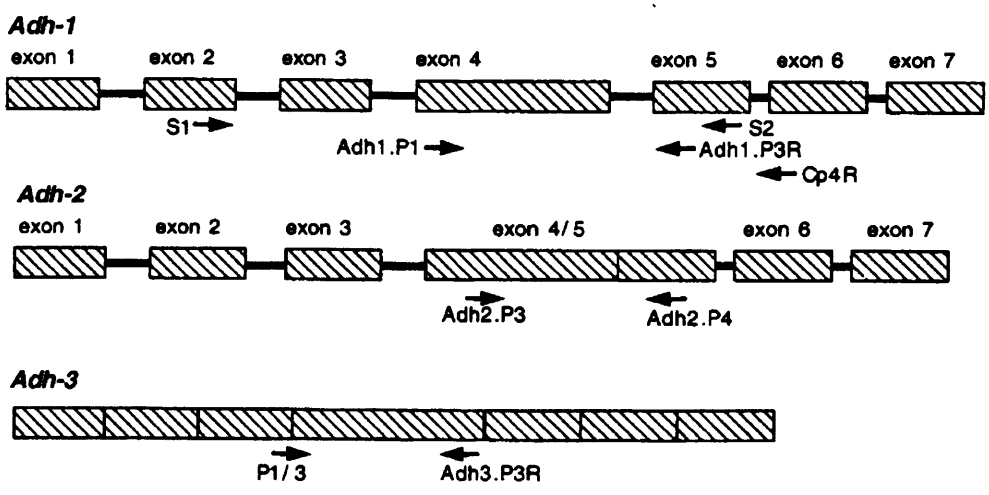

Fig. 3.3 Interpretation of the structure of the *Leavenworthia* alcohol dehydrogenase genes, showing approximate positions of primers. Hatched boxes represent exons, and solid lines represent introns (from Charlesworth *et al.*, 1998).

Primers for *G3pdh* (Fig. 3.4) were effectively used in a survey by Olsen (1999) to show variation within *Manihot esculenta*. Data were subsequently shown to be useful in establishing the biogeographical origin of cassava. In all plants examined thus far (*Arabidopsis thaliana*, *Ranunculus acris* and *Manihot esculenta*), this region spans nine introns introns and nine exons (Shih et al., 1991). The primers used correspond to exons 5 and exons 9 in *Arabidopsis* (approximately position 1620 – 2456bp in *Arabidopsis*). As is typical with nuclear introns, all four of them begin with the bases GT and end with AG.

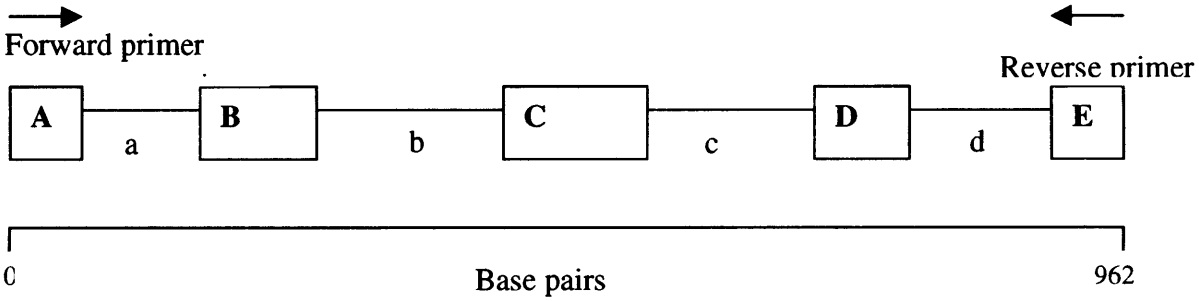

Fig. 3.5 Structure of the *G3pdh* region in *Manihot*. Uppercase and lower case designate exons and introns, respectively as follows: partial exon A (52 bp), intron a (84 bp), exon B (98 bp), intron b (265 bp), exon C (143 bp), intron c (101bp), exon D (84 bp), intron d (101 bp) partial exon E (34bp), redrawn from Olsen and Schaal (1999).

The *G3pdh* region in cassava contained 64 polymorphic sites of which only one encodes an amino acid substitution. The amount of variation in the noncoding regions therefore was high and allowed the identification of 28 haplotypes among 212 individuals within the species *M. esculenta*.

Other low copy-number nuclear genes that have been investigated sucessfully include the small, heat-shock genes which have been shown to be useful at a generic level (Waters, 1995). The gene for chloroplastic glyceraldehyde 3-phosphate dehydrogenase (*gap*A) has been shown to evolve at a comparable rate to *rbc*L and ITS and has been shown to be useful in resolving genera within the Berberidaceae (Adachi *et al.*, 1995). Genes which encode isozymes such as phosphoglucose isomerase genes (*Pgi*) may also be useful in constructing phylogenies (Gottlieb and Ford, 1996). The histone locus is part of a large multigene family whose introns are thought to be potentially useful at the species level (Doyle *et al.*, 1994). The multigene family for glutamine synthase (GS) has been shown to be phylogenetically useful at the species level (Emshwiller and Doyle, 1999). The MADS-

box genes encode a family of transcription factors which control a wide range of developmental processes in flowering plants, and have been shown to be useful both in understanding early angiosperm evolution (Becker *et al.*, 2000) and at the population level in Brassicaceae (Bailey and Doyle, 1999).

3.4.4 Desirable attributes for low copy-number genes

Although there are potentially thousands of genes which can reveal phylogenetic information, there are many complications and factors to be considered. After a review of the literature (Mathews *et al.* 1995; Strand, 1997; Charlesworth, 1998; Soltis and Soltis, 1998), it was established that in order for a gene to be investigated further it should ideally have the following properties.

1) The gene should be either single-copy, or form a small family.

2) The gene should be well-studied, with sequences available in databases; this adds to the ease with which primers can be designed.

3) The gene should have both exons and introns. Exons provide conserved areas in which to anchor the primers and allow ready alignment of sequences between species. Introns potentially provide a region of increased variability. The two together, therefore, allow comparisons to be made across a wide taxonomic scope.

4) There must be enough variability in an area of less than 1 kb, otherwise the costs of sequencing are prohibitive.

5) The gene should code for an enzyme which is fundamental to plant physiology or development, rather than a more specialised enzyme product.

6) It should be possible to infer homology of the gene between taxa that are included in the study, in other words it must be possible to ensure that the genes are descended from the same common ancestor and that different copies of a gene are not compared.

7) The gene should be able to be sequenced directly from a PCR product.

To summarise therefore, DNA sequences from the low copy fraction of the nuclear genome potentially provide a good source of phylogenetic resolution, specifically at the species level (Strand *et al.*, 1997). The reason that they may be useful is that they are free from the constraints of many processes such as uniparental inheritance and nonhomologous recombination which lead to incongruence of species and gene trees. Furthermore, single mutations become established more quickly than they would in multicopy DNA. However, single copy DNA sequences of the nuclear genome have so far been relatively under-explored in the phylogenetic study of plants (Mathews *et al.*, 1995).

3.5 Phylogenetic incongruence

Not all genes or sequences will behave in the same way and may possibly imply different phylogenies. This phenomenon is called phylogenetic incongruence. Incongruence of topologies derived from two different data sets is often treated as an unfortunate but undesirable side-effect of a phylogenetic analysis and impedes phylogenetic understanding (Wendel and Doyle, 1998). However, Wendel and Doyle propose that incongruence is an important observation that often reflects something interesting in the biology of the group under study and brings to our attention one or more evolutionary processes that may not have been suspected were incongruence not present (Table 3.1).

Table 3.1 Phenomena that may lead to conflicting phylogenetic hypotheses: From Wendel and Doyle (1998).

Technical causes	Organism-level processes	Gene- and genome-level processes
Insufficient data	Convergent or rapid	Intragenic recombination
Gene choice	morphological evolution	Orthology/paralogy conflation
Sequencing error	Rapid diversification	Interlocus interactions and
Taxon sampling	Hybridisation/introgression	concerted evolution
	Lineage sorting	Rate heterogeneity among taxa
	Horizontal transfer	Rate heterogeneity among sites
		Base compositional bias
		RNA editing
		Non-independence of sites

The first question to ask, is whether or not the conflict is significant. If particular clades are not well-supported, it may well be that slightly less parsimonious solutions do not cause conflict with other data sets (Wendel and Doyle, 1998).

One of the fundamental principles of molecular systematics is that the rate of molecular evolution for particular sequence should be optimised as much as possible to the scale of divergence in the group whose phylogenetic relationships are being investigated. If the gene evolves too slowly, there will be insufficient resolution and alternatively in a gene which evolves too quickly phylogenetic signal may be obscured by homoplasy.

To interpret results in the most satisfactory manner, it is important to be aware of all causes of phylogenetic incongruence. Where problems have occurred, they are discussed further in chapters 6-9, in the context of where they are thought to have contributed to phylogenetic incongruence.

3.6 Molecular clocks

The idea of an evolutionary clock based on a constant rate of morphological character evolution was first proposed by Simpson (1944). This idea was not favoured however as

there was no neutral theory that predicted rate constancy at the morphological level. Zuckerhandl and Pauling (1962, 1965) first proposed a molecular clock based on protein data. Although many genes and proteins do not evolve at a constant rate through time (Britten, 1986; Avise, 1994), the concept of a molecular clock, or at least one based on averages, is generally accepted as a model for most phyla (Wray *et al.*, 1996).

The fundamental assumption of a molecular clock is that the rates of mutation are equal across the different evolutionary lineages (Muse, 2000). If a molecular clock exists, then molecular sequence data can be used to estimate the dates of historical events, such as speciations. It is also useful for testing hypotheses in biogeography (Avise, 1994; Hillis *et al.*, 1996).

In order to demonstrate rate homogeneity it is necessary to perform a relative rate test on the lineages being compared. To take a simple example (Fig. 3.5), species **A** may be separated from species **B** by 10 substitutions. It is not possible to determine if all of these changes occurred on the lineage leading to species A, or to species B, or indeed if there were some changes on each branch (Sanderson, 1998). This problem is resolved by adding a third, closely related species, **O** (Fig. 3.5). Assuming that the amount of change from ancestor to A is a, and the amount of change from ancestor to B is b, then the relative rate test evaluates the null hypothesis, H_0: D(O,A) – D(O,B) = 0. This expression is true only when there is rate homogeneity.

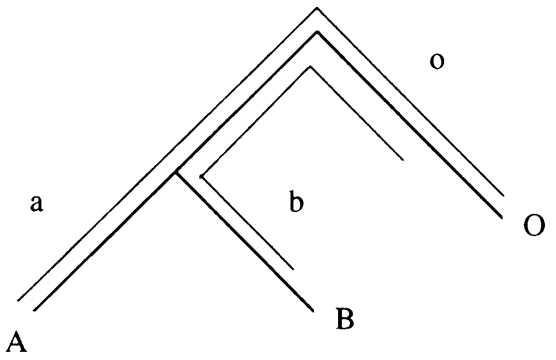

Fig. 3.5. Explanation of the relative rate test. A and B have diverged from their most recent ancestor, O, to distances *a* and *b*, respectively. The test is of whether D(O,A) – D(O,B) = 0.

The methodology involved in relative rate tests is discussed further in the data analysis section of Chapter 4.

Heterogeneity of substitution rates has been reported between lineages for several genes from each type of plant genome. One example is the *rbc*L gene of chloroplasts which has been found to evolve significantly faster in annual angiosperms than in perennial angiosperms, gymnosperms, ferns and liverworts (Savard *et al.*, 1994). Also *rbc*L evolves faster in grasses than in palms (Wilson *et al.*, 1990; Gaut *et al.*, 1992). Difference in generation time is clearly an important confounding factor when attempting to apply a molecular clock. Similarly, five different mitochondrial genes have been found to evolve significantly faster at nonsynonymous sites in angiosperms than in liverworts (Bousquet *et al.*, 1992).

If, however, rate homogeneity can be established, a molecular clock can be calibrated either using the fossil record, or using average rates of evolution for the particular DNA sequence involved. For example, based on work by Wolfe (1989), and supported by the work of Gaut *et al.* (1992), the average rate of mutation in the chloroplast genome as a whole is one mutation per thousand bases per million years. It is important to recognise, however, that each gene or sequence will have its own particular 'clock'. Many studies have shown that the rates of evolution vary among regions of genes and proteins subject to the constraints imposed by the genetic code or the protein structure and expression (Sanderson, 1998).

CHAPTER 4

METHODS

4.1 Collection of plant material

Leaf samples were collected from all plant accessions for the purposes of DNA extraction for PCR amplification and sequencing. In all cases leaves were removed from the plant using sharp secateurs. Material was either instantly frozen in liquid nitrogen or dried in silica gel. Where the latter approach was used, silica gel was consistently changed until no further absorption occurred. Samples were stored in the dark in dry conditions until they were needed. Once DNA was extracted it was stored at -20° C.

4.2 DNA extraction

Total genomic DNA was extracted from 100 mg of fresh or frozen material, or 50 mg of dried material using Qiagen DNeasy kits (Qiagen Ltd., Boundary Court, Gatwick Road, Crawley, West Sussex, RH10 2BR). Initially the C TAB method (Doyle and Doyle, 1990) were attempted but low yields of DNA were obtained.

In some cases (particularly *Eucryphia* species and *G. jodinifolia*), the plant material was too glutinous to be successfully extracted with the kits. Due to the presence of polysaccharides, the leaf extract from these species was so glutinous that it adhered to the membrane in the tube and thus blocked it and prevented any further passage of DNA into the column. The DNA from these samples was therefore extracted using a modification of the Doyle and Doyle (1990) protocol with additional ammonium acetate washes.

Leaf material was ground to a fine powder in liquid nitrogen using a clean mortar and pestle. Fine sand was added to assist in the grinding process and PVPP (polyvinyl polypyrrolidone; Sigma P-6755) was added to bind and precipitate polyphenolics. This

mixture was then added to a 5 ml 2 X C TAB (Sigma H 5882) solution which contained 10 µl of mercaptoethanol. The tubes were incubated in a water bath at 60° C for 30-60 minutes and shaken every 5-10 minutes.

The tubes were then removed from the water bath and 5 ml chloroform: iso-amyl alcohol (24:1) was added. The tubes were then shaken and centrifuged at 10,000 rpm for 10 minutes. The upper aqueous layer was removed using a 1 ml pipette with the tip cut off (this prevents the shearing of DNA) and added to a clean tube. The chloroform: iso-amyl alcohol extraction stage was then repeated. After the second centrifugation, the aqueous layer was then added to 3.5 ml of cold isopropanol and the tube was gently shaken to assist the precipitation of DNA. The DNA was then transferred using a sealed glass pipette to a clean Sterilin tube containing 5 ml wash buffer solution (76% ethanol, 10 mM ammonium acetate). Tubes were left at room temperature for at least 20 minutes.

The DNA was removed from wash buffer and air dried for a few minutes before being dissolved in 1X TE (Tris-EDTA buffer, pH 8) in a clean Steralin tube. Where the pellets of DNA were small they were re-suspended in 1 ml TE, where pellets were larger, 2 or 3 ml was added. Pellets were left to dissolve for one hour. RNase solution (1 mg/ml) was added (10 µl of RNase per ml of DNA solution), and the tubes inverted gently and incubated at 37 °C for 60 minutes.

Once the RNase digestion was complete, 100 µl 3 M ammonium acetate per ml DNA solution was added and the tubes were gently mixed. The time taken for the DNA to dissolve depends upon its concentration. The ammonium acetate stage was repeated three times in order to remove the excess polysaccharides. 2.5 volumes cold absolute ethanol were then added to the tubes, which were gently rocked to precipitate the DNA. The DNA was then removed as before and air-dried . On rare occasions the DNA would not adhere to the glass rod, therefore it was spun in a centrifuge for 2 - 3 min. at 2000 rpm and the alcohol was carefully decanted. Tubes were then inverted over some tissue to dry the DNA pellet. Once dry, the DNA was again re-suspended in an appropriate volume of TE buffer.

To determine the quantity present, DNA was run against standards and diluted to approximately 1 ng/µl for subsequent use in PCR.

4.3 DNA electrophoresis

DNA electrophoresis was carried out on agarose gels. For a standard 1% gel, 1 g agarose was added to 100 ml 1 X TBE (Tris-Boric Acid-EDTA) in a conical flask. Cling film was placed over the top of the flask and punctured before being microwaved for two minutes at maximum power. After one minute the flask was removed and swirled to dissolve the agarose before it was given another minute. The agarose solution was then allowed to cool down for ten minutes or until the flask could comfortably be hand-held and 5 µl of ethidium bromide was added. The solution was then poured into a pre-prepared former and left to set in a fume cupboard.

DNA fragments were visualised under UV illumination. 1kb ladders were used to obtain an approximation of the product size. To visualise the product, in most cases, 2 µl PCR product was added to 2 µl loading buffer and loaded into the gel. The remainder of the PCR product was then cleaned as described below.

4.4 Sequence amplification

4.4.1 Primer design

Table 4.1 gives the details of the primers used. For the sake of simplicity, details are only given for primers that yielded reliable products. Having decided upon which genes were potentially interesting, the next problem was establishing if the primers would work for the three genera of interest. Below is a discussion of the design of primers (where appropriate), the way in which the primers were screened for the three genera, and subsequently the way in which they were optimised.

Three chloroplast, one mitochondrial and thirteen nuclear sequences were investigated. The universal primers of Demesure *et al.* (1995) were used to amplify the chloroplast sequences *trn*L-F and *trn*H-K. Primers for *rpo*A (Wilkinson, pers comm.) were also used. Similarly, published primer sequences were used in the attempted amplifications of *cox*I (Lu *et al.*, 1998) *adh* (Charlesworth, 1998), *G3pdh* and other loci (Strand *et al.*, 1997) (Table 4.1) and ITS (Gardes and Bruns, 1993).

The chloroplast and mitochondrial primers had been shown to be effective and amplify well in other studies. The two internal transcribed spacers (ITS-1 and ITS-2) and the 5.8S rDNA gene were amplified for all plant accessions (Fig. 3.2). Plant-specific primers were used in order to avoid any possibility of non-target amplification attributable to fungal contamination. This occurred in my work when the standard fungal primers designed by White *et al.* (1990) were used for amplification and sequencing. Plant specific forward and reverse primers were designed from the sequence data published by Gardes and Bruns, (1993) whereby each primer had five plant-specific bases. In addition, the reverse primer was truncated as a base position in the basidiomycetes was found to be absent in higher plants. Amplification was most effective at Mg^{2+} concentration of 0.5 - 2 mM (see Table 4.1).

For the phytochrome and *adh* genes, primers were designed to match the criteria stipulated by the programme *Primer v. 0.5* (Lincoln, 1991). Where possible, primers were between 18 and 20 bp long. The programme *Primer* ensures that primer-dimer formation and runs of single or simple repeat sequences do not occur within the primer. *Primer* also calculates the stability of the primers, unstable primers are those which are less strong at the 3' end than at the 5' end and therefore lead to the occurrence of non-specific binding.

Phytochrome. Primers were designed to span the phytochrome region that included the large intron of *PHYB*. Since *Araucaria* was going to be included in this project, primers were designed which would hopefully work with both angiosperms and gymnosperms (The genus was abandoned due to time constraints). Two complete gymnosperm phytochrome sequences gymnosperms were available on Genbank: *Pinus sylvestris* and *Picea abies*. These were compared, using *Megalign,* to *Arabidopsis* sequences from the five members of the gene family. These five *PHY* genes have been found in a range of angiosperm families. Although the program managed to find part of each phytochrome gene to which it could align with the gymnosperm sequence, the section of the *PHY B* that it aligned to was the most interesting as it split the sequence at either side of the large intron. As the gymnosperm sequence was derived from messenger RNA, it seems likely that it also has an intron at this position. The primers were 17 nucleotides long and were chosen as they had the closest melting points and were not complementary to each other.

The primers were optimised and found to produce a strong single band on an agarose gel, when there was an annealing temperature of 59°C and an extention time of 1 minute with a Mg^{2+} concentration of 2.5 mM. There were problems with the phytochrome primers as despite initially looking like a single PCR product, it was never free from the contamination of other gene sequences. Under normal conditions, when 5 µl of PCR product was loaded with loading buffer on to a 1% agarose gel, one band appeared to be present. When this was sequenced however, it was apparent that multiple products were being amplified.

When the entire 50 µl product was loaded, several bands were present. The physical separation of these different fragments was attempted by loading 50 µl product onto an agarose gel and cutting out the individual bands. When this was sequenced the 4 products were again present which implies that the primers amplify more than one phytochrome gene in *Araucaria*. The same problem occurred when this was attempted for the other genera.

4.4.2 PCR reaction mixture

All regions were amplified by the polymerase chain reaction (PCR) using aproximately 2 – 3 ng of total DNA as template. The reactions were carried out under mineral oil in a total volume of 25 or 50 µl using a DNA thermal cycler (Perkin Elmer Cetus). A typical temperature profile is given below. Initially a 25 µl reaction was carried out and if successful, a 50 µl reaction would be performed, the two products being combined and purified for sequencing. The recipe for 50 µl is shown below:

Water	33.1 µl (31.1 µl, if BSA is used)
Reaction Buffer (10X)	5 – 7.5 µl
2 µM dNTP's (2 mM)	2.5 µl
Mg^{2+} (50 mM)	2 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
(BSA	2 µl optimal use as necessary)
Taq (1U μl^{-1})	0.4 µl
DNA	0.5 – 2.5 µl

Arguably one of the factors which has the biggest influence on the amplification of DNA is the concentration of magnesium. The magnesium concentration affects the annealing of the primer to the template DNA by stabilising the primer-template interaction, it also stabilises the replication complex of the *taq* polymerase with the primer-template. Therefore it is possible for the magnesium to increase non-specific annealing and produce undesirable PCR products (which appear as multiple bands in the gel). The stronger the concentration of magnesium, the less stringent the conditions become for the adhesion of the primers to DNA. Often however the reaction will work better at more stringent conditions (low concentrations of magnesium) as it prevents the amplification of these misprimed products. To determine the best magnesium concentration, a magnesium concentration gradient was performed. Generally this was the first step in optimising primers for each genus. The concentration of magnesium was often specific to a genus and therefore worked best for the different genera at different concentrations. Optimal concentrations ranged from 0.5mM – 2 mM (Table 4.1).

The concentration of the DNA template sometimes had an effect on the outcome of PCR. If the PCR did not work when DNA was at standard concentration, the reaction was repeated with a stronger or weaker concentration. In most cases the working strength aliquots of DNA were used as template DNA. In some cases where only a weak product was obtained, this was used as a template for a second PCR. In the case of *Adh* for *Coriaria*, the outer external primers were initially used and this product was used as a template for a nested PCR.

Sometimes proteins inhibited the PCR reaction and in these cases, the addition of 2 µl BSA, which binds to proteins, to a 50 µl reaction caused it to be successful.

4.2.3 PCR Temperature profiles and extension times

A typical temperature profile is as follows;
A hot start of (1 X 94° C for 7 minutes) linked to

30 X
(94 °C for 30 seconds - to ensure denaturation of DNA
55 °C for 30 seconds - to allow the primers to anneal

72 °C for 45 seconds - for the taq polymerase to extend the sequence)
this was linked to
1 X
(72 °C for 10 minutes to ensure all polymerisation is completed)

finally, this was linked to a chilling cycle of 10 °C for half an hour to keep the DNA PCR product cool.

As different primers optimally anneal at different temperatures, this parameter greatly affected the outcome of the PCR. The theoretical optimum often differed from the actual one, and this was determined by repeated trial on the Perkin Elmer PCR machines. Although it would potentially have been quicker to use a gradient cycler PCR machine this machine was not available for the majority of the work and the optimum temperature varied between machines due to differences in ramping time. Therefore the optimum temperature had to be determined on the machine in which all the reactions would be carried out. Generally the annealing temperature which gave the best results was 55°C. A higher annealing temperature was used for primers which had a high GC content or which were longer than average. A list of temperatures is given in Table 4.1.

The enzyme *Taq* polymerase is thought to add bases to the gene sequence at a rate of approximately 500 bases per 30 seconds. Therefore, for the majority of gene sequences which were studied in this project an amplification time of 45 seconds was sufficient. Specific modifications are also given in Table 4.1.

4.2.4 Number of cycles

Generally 30 cycles were sufficient for further sequencing, although occasionally to increase the strength of a PCR product this was raised to 35 cycles. It is not advisable to have many more than this number as it allows the amplification of mis-primed products which makes the actual sequences more difficult to interpret due to a high level of background noise.

Table 4.1 PCR details for the sequences amplified, NA = not applicable, as the reactions did not work at all under any conditions

Sequence	Primer sequence 5' – 3' Forward primer top Reverse primer bottom	Anneal. Temp.	Extension time	$MgCl_2$ (m mol)
Chloroplast				
*trn*H-K Demesure *et al.* (1995)	ACGGGAATTGAACCCGCGCA CAACGGTAGAGTACTTCGGCTTT TA	62°C	2 min	1.5
*Trn*L-F	CGAAATCGGTAGACGCTACG ATTTGAATGGTGACACGAG	55°C	1 min	2
*Rpo*A	CTCAAACACTACAGTGGAAGTG CAGCATCTATAGGATAACTTCC	55°C	45 sec	2
Mitochondrial				
*cox*I (Lu *et al.* 1998)	TTYTTGTTYTTYGGICARCCIGA RGTTA GGIACIGGTTGGACTYTTTAYCC ICC	55°C	1 min	2
Nuclear				
ITS 1-2 (modification from Gardes, & Bruns, (1993)	CCTTATCATTTAGAGGAAGTAA TAGGGGACTTGGGCCGGTCC	55°C	1 min	0.5 - 2
aat (Strand *et al.* 1997)	CGTATTCAAGAGAACAG TCAACACCAGTAGGGTTA	NA	NA	NA
adh external (Charlesworth *et al.* 1998)	GATGT(T,C,A)TACTTCTGGGAAA GC ATC(G,A)TGGACACATTCAAAT GC	touch-down	1 min	2

adh internal designed by me	GAGAATCCTTTAAACAATT GATTTACAAACTCAGTCACC	57°C	1 min	1
adh (Strand *et al.* 1997)	TACTTITGGGAACGIAAGGTA TCICCIACACTCCIACAAT	55°C	1 min	1
cam (Strand *et al.* 1997)	AGCCTNTTCGACAAGGATGG AGTGANCGCCATCACAGTT	NA	NA	NA
chi (Strand *et al.* 1997)	TNNTTCCTCGGCGGGCGC TCCCCATNATGGNCTCCA	60°C	1 min	2
chs (Strand *et al.* 1997)	AGGAAAAATTCAAGCGCATG TTCAGTCAAGTGCATGTAACG	NA	NA	NA
g3pdh (Strand *et al.* 1997)	GATAGATTTGGAATTGTTGAGG AAGCAATTCCAGCCTTGG	55°C	2 min	0.5
leafy	AGCATCCTTTCATTGTCACG CCAACATTTACTCCCCTCTCC	NA	NA	NA
pgi (Strand *et al.* 1997)	TCTCTICAGTAIGGCTT AATGATACATTCCATCACCT	NA	NA	NA
Phytochrome PIN designed by me	GCATTATCCTGCTACTGATATT GGCTGAACCACACATCC	57°C	45 sec	2
tpi (Strand *et al.* 1997)	AAGGTCATTGCATGTGTTGG CTTTACCAGTTCCAATAGCCC	NA	NA	NA

4.2.5 "Touchdown" PCR

The theory behind "touchdown" PCR is that it starts at a high annealing temperature which is then decreased in steps to reduce the amount of non-specific PCR product. The annealing temperature starts at 60 °C and is relaxed by two degrees every two cycles until the eighth cycle when it has reached 46 °C, at this point eight cycles are performed. Conditions are then made slightly more stringent by the temperature being raised, and 25 further cycles are carried out at 48 °C.

4.3 Separation of PCR products

4.3.1 Physical G3pdh modifications

Physical separation was possible for the 2 *G3pdh* products as they differed by a suitably large margin of 200 bp which allowed them to be easily viewed and cut from a gel. 50µl of PCR product was separated on a 1.6 % agarose gel. These products were separated on the gel then cut out using a sharp scalpel, and cleaned using Qiagen Agarose extraction kits before being directly sequenced.

4.3.2 Cloning for Adh

Cloning was carried out using Promega cloning kits. 1-3 ul of fresh PCR product was ligated into a plasmid (puc13). Competent cells were transformed overnight with the plasmid on plates containing LB buffer, ampicillin X-gal and IPTG (isopropyl – b-D-thiogalactopyranoside. The white colonies were picked from plates and added to 2 ml SOC solution again containing ampicillin. Colonies were grown overnight and were purified using Qiagen plasmid prep kits (ref). To check that an insert of the correct size had occurred, a PCR reaction was carried out using universal m13 primers. Sequencing was attempted both with the m13 primers and with the original PCR primers.

With the *Adh* work using the Charlesworth *et al.* primers of *Coriaria*, some species, only produced one band while in other species two products of similar size were amplified. These were separated using cloning. After cloning the plasmids were sequenced using universal M13 primers.

When the resulting sequences were aligned with the original data set, one locus was obviously the same as that in the original data matrix which was obtained from the direct sequencing of these primers in species where only one product was obtained (*Adh*–A). The other sequence was substantially different and was termed *Adh*–B. The two sequences could still be aligned, and from this alignment specific primers were then designed to amplify *Adh*-A, using Primer V. 0.5 (Lincoln , S., 1991).

4.4 Sequencing of PCR products

4.4.1 Purification of PCR products for sequencing

In order to remove excess primer and unincorporated dNTPs, the PCR products were purified before sequencing using QIAquick purification kits (Qiagen, catalogue number 28104) following the manufacturers instructions.

4.4.2 Sequencing reaction

PCR products were sequenced directly using the di-deoxy chain termination method. The reactions were initially carried out in a total volume of 20 µl consisting of 8 µl of terminator ready reaction mix (Perkin Elmer-ABI), 1 µl of sequencing primer (3.2 mM) and 11 µl total volume of DNA (200 - 500 ng) and dH_2O. This produced such strong sequence that the total volumes were able to be halved. The ready reaction mix contains AmpliTaq® DNA polymerase, magnesium chloride, and Big Dye terminators that have been labelled with a fluorescent dye. At all times, mineral oil was applied to overlay the reactions.

The PCR programme consisted of 25 cycles of 96 °C for 30 seconds, 50 °C for 15 seconds and 60 °C for four minutes. The reactions were kept at 4 °C for half an hour to ensure the reaction had terminated.

4.4.3 Purification of sequencing reactions

The sequencing reaction was removed from underneath the mineral oil and added to 50 µl of 100% ethanol and 2 µl of sodium acetate (2M, pH = 5.2). The tube was mixed and placed on ice for 30 minutes. The tubes were then centrifuged at 13, 000 rpm for ten minutes. Providing a pellet was observed, the supernatant was carefully removed using vacuum suction, caution was taken to ensure that the DNA pellet was not disturbed. Where no pellet appeared to be present, the tubes were replaced in the ice for a further 30 minutes and re-centrifuged at 13, 000 for a further 10 minutes. The pellet was then washed in 70% ethanol. Tubes were then spun down at 13, 000 rpm for two minutes. The ethanol was removed using vacuum suction and dried in a vacuum desiccator for at least half an

hour. The tubes were sent to the Protein and Nucleic Acid Chemistry Laboratory, University of Leicester for sequencing on an ABI 377 sequencer (Perkin Elmer).

4.5 Data Analysis

The electropherograms were analysed manually using Sequence Navigator v. 1.0.1 (Perkin Elmer). The checked sequences were exported to EditSeq v 3.75 (DNAStar) and aligned using MEGALIGN v 1.0.5 (DNAStar).

4.5.1 Gene sequence data

Gene sequence data information given in the tables in each results chapter were determined as follows:

The length range of the data and the alignable length were observed. The mean lengths are given both for the ingroup and the outgroups, individually and together. The number of indels was observed and the number of each particular size given.

The G + C content was calculated in MacClade under the Chart menu, followed by the Average no. of steps per character option.

The sequence divergence was determined in PAUP* 4 (Swofford, 1998), in the table (STATE WHICH), it is initially given in terms of number of bases, and secondly as a proportion of the alignable length. The number of constant, variable, autapomorphic and informative sites were also determined using PAUP* 4 (Swofford, 1998).

The minimum and maximum number of transitions and transversions were calculated using MacClade. This is under the Chart menu followed by state changes and stasis, (the all included characters and current tree only options were chosen). The average number of steps per character was calculated in MacClade under the Trees menu, followed by the number of steps option.

The number of trees found, the number of steps for the shortest tree and the Rescaled Consistency Index and the Consistency Index were all calculated using PAUP 4 (Swofford, 1998).

Graphs have been drawn to show the number of times each individual base pair number has to change in order to be consistent with the most parsimonious tree for that gene. These graphs give an idea therefore of where the variation occurs in the gene. They represent an extension of parsimony analysis in that they show where the variation is in order for the most parsimonious tree to be produced. A parsimony analysis was therefore performed as before on the ingroup plus the outgroup. It was also done on the ingroup alone, therefore it was unrooted. This does however show the location of the variation that exists within the ingroup.

In all cases, the bottom section (drawn in green) relates only to the ingroup data. The next section up includes both ingroup and outgroup data. This information is useful as it shows where the variation within the gene in closely related genera is. The top half of each graph shows the cumulative number of steps that are taken in order to produce the tree. Thus it is easier to see which areas are more variable. The green and red lines correspond to the ingroup data only and the ingroup plus outgroup data respectively. The blue lines represent the ingroup data which has been normalised to correspond to the ingroup plus outgroup data.

4.5.2 Phylogenetic analysis

Phylogenetic trees were calculated with PAUP* 4.0 (Swofford, 1998). All of the analyses were performed using exhaustive searches. No weighting or ordering was performed on the characters and any gaps were treated as missing data. A matrix was written to code the presence or absence of insertions or deletions. This was generally a binary matrix; where a deletion was present it was coded 0 and the corresponding presence of base pairs was given a code of 1. Where more than one size of deletion or insertion were present, the alternative sizes were coded with 2, 3 and so on.

A nonparametric bootstrap (Felsenstein, 1985) analysis of 1000 replicates was performed to test the support of the clades within the ingroup. Each bootstrap had 100 random addition applications, and the options TBR swapping and MULPARS were chosen. The

COLLAPSE but not the STEEPEST DESCENT options of PAUP were used during these searches and the optimisation selected to interpret character changes was ACCTRAN. The most parsimonious trees were calculated independently for the different data sets and the chloroplast data and the nuclear data sets were then combined to obtain a more resolved chloroplast and nuclear tree. Finally, in the absence of conflict, all of the data sets were combined; in the case of *Eucryphia*, morphological data were also included.

Maximum likelihood (ML) analysis was performed using the general-time-reversible model (GTR; Yang, 1994). This model has been shown to work well in other studies (e.g., Flook *et al.*, 1999). The model assumes that all the substitution probabilities are independent. The model does however account for a possible transition-transversions bias, and for variable substitution rates among sites which are described by a gamma distribution with four rate categories. This is one of the models incorporated in PAUP* 4.0, the model estimates the proportions of invariable sites and the shape parameters of the gamma distribution. The starting trees for ML searches were calculated via neighbour joining and the swapping strategy used was TBR.

For all the cases tried, the overall topology for the trees using ML analysis was identical to that obtained using parsimony.

4.5.3 Testing for rate heterogeneity

The necessary hurdle that has to be jumped over before a molecular clock can be applied, is the demonstration of rate consistency of evolution over the different lineages. The question that follows therefore is how to choose and estimate an appropriate measure of distance and secondly how to establish the statistical significance of differences in relative distance (Sanderson, 1998). There are numerous methods to establish this, but for their divergence estimations, most tests use distances that are derived from pair wise sequence dissimilarities. Relative rate tests also exist which are based on maximum likelihood (Felsenstein, 1988). They are described in detail by Muse and Weir (1992). Recent evidence suggests that neither method gives significantly better results (Tajima, 1993, Sanderson, 1998) in this study the test described by Tajima (1993) was used.

This test is based on a chi-squared distribution (χ^2). It states that m_A is the number of sites in which the ingroup A has one nucleotide, and where the ingroup B and the outgroup share the same nucleotide but one which differs from that of ingroup A. m_B is the opposite of this whereby it is number of sites that ingroup A and the outgroup share with each other but differ from ingroup B. Under rate consistency the following expression is distributed as a χ^2 distribution with one degree of freedom:

$$\frac{(m_A - m_B)^2}{(m_A + m_B)^2}$$

Relative rate tests were calculated as above for each genus. Where the rate was consistent a molecular clock could then be calibrated.

4.5.4 Calibrating a molecular clock and dating lineages

It is important to ensure that the genes that are of potential use as molecular clocks, do indeed function as such, and can be calibrated accurately. Where possible, this was calibrated using fossil data. The first decision therefore is the correct placement of the fossil. A simple case to illustrate this is shown in Fig. 4.1. In this example, the oldest fossil known in a genus is 50 million years old. To calculate the rate of the clock, the number of molecular differences between the genus and the closest known relative has to be established. This is done by adding the number of molecular characters that all the species in the genus share plus the average of the number of characters unique to each species. In the simple case described (Fig. 4.1) the average number of molecular characters therefore over which 50 million years of evolution has occurred in 112.5. The amount of fossil data available as for each genus in this project was different. This slightly altered the approach taken to each genus and these approaches are discussed in the relevant results chapters.

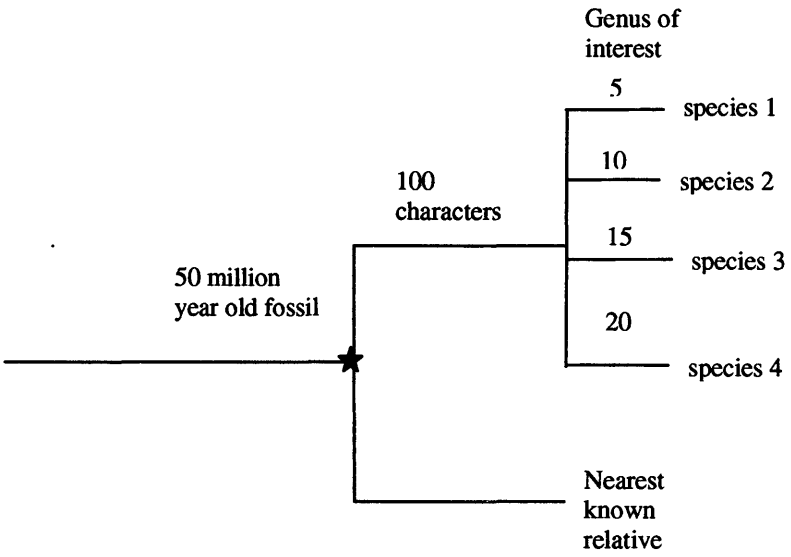

Fig. 4.1. An illustration of the dating of phylogenetic nodes, a fossil at the base of the four species can be placed on the node indicated by the red star. See text for details.

The following equation can be used to obtain rate estimates as long as at least one date is known, this may be either the date of a fossil, or the geological history. D is the number of substitutions per site, K is the overall rate of substitution per site per unit of time and t is the age of the common ancestor.

$$D = 2Kt$$

Therefore the rate (K) = the number of molecular differences over the tree (112.5 in the case in Fig. 4.1) divided by twice 50 million years.

4.5.5 Dating a biogeographical split

After the rate for a molecular clock has been determined, the final question of the divergence time of two species, or two groups of species can now be answered. To do this one final parameter is required; the number of mutations that have occurred between species either side of the biogeographical split. This value is simply multiplied by the average rate of the clock. Minimum and maximum divergence times were also calculated. Again this is discussed in further detail in the relevant chapters.

CHAPTER 5

RESULTS: SURVEY OF SEQUENCES

This chapter describes the results of the preliminary work involved in choosing appropriate DNA sequences for the phylogenetic analysis of *Eucryphia*, *Griselinia* and *Coriaria*. Initially three members of each genus were chosen to represent the range of morphological and geographical variation present and, by inference, provide a good chance of recovering at least some molecular variation. For each gene, if sufficient sequence variation was not present among the species no further work was conducted. In the following account, the results of the survey of plastid sequences are presented first and then the nuclear ones are considered.

5.1 Chloroplast DNA

Of the eight chloroplast regions examined, in all cases *trn*L-F, *rpo*A1 and *trn*H-K were amplified and these were chosen for further work as they consistently produced strong, single-banded PCR products (Table 5.1).

Table 5.1. Chloroplast sequences tested in this project, indicating the number of fragments amplified by PCR

Species	*ndh*1	*ndh*2	*psb*E-F	*rpo*A1	*rpo*A2	*rsp*4	*trn*H-K	*trn*L-F
Coriaria angustifolia	1	0	0	1	0	1	1	1
C. intermedia	1	0	0	1	0	0	1	1
C. myrtifolia	1	0	0	1	0	1	1	1
Eucryphia cordifolia	0	0	0	1	0	0	1	1
E. glutinosa	0	0	0	1	0	1	1	1
E. wilkiei	0	0	0	1	0	0	1	1
Griselinia lucida	1	0	0	1	0	1	1	1
G. racemosa	1	0	0	1	0	1	1	1
G. scandens	0	0	0	1	0	1	1	1

*trn*L-F was consistently variable enough to provide resolution at the species level. Although

the *rpo*A1 gene was not as variable as the *trn*L-F region it was useful as in no case did information from it conflict with data from the *trn*L-F region. *Trn*H-K was particularly informative in the case of *Griselinia* but less so in the case of *Eucryphia*. Although this region was successfully amplified in *Coriaria*, due to the A/T-rich nature of the sequence, only an invariable region of approximately 100 bp could be sequenced. After about 100 bp of sequence electropherograms showed not one but two sets of peaks. The second set may have been caused by the primer slipping and re-attatching at a different place. Work on *trn*H-K in *Coriaria* was consequently discontinued.

5.2 Mitochondrial DNA

A 700 bp region of this rapidly evolving region of *cox*I (Lu *et al.*, 1998) was sequenced in *Eucryphia cordifolia*, *E. glutinosa*, *E. wilkiei*, *Griselinia lucida*, *G. racemosa*, *G. scandens*, *Coriaria angustissima*, *C. intermedia* and *C. myrtifolia*. Although there were no technical problems, each of the three genera proved to be invariant, i.e. there was no infra-generic variation and so mitochondrial work was not pursued further.

5.3 Nuclear DNA

5.3.1 rDNA ITS

Preliminary trials with the standard fungus-derived primers of White *et al.* (1990) showed that they amplified more than one product in the three genera studied here. Amplification using the new primers was most effective at a magnesium ion concentration of 0.5 mM. Apart from amplifying the target product cleanly, the plant-specific primers also prevented unwanted amplification of material derived from fungal contamination. This was particularly useful when working with the leaves of *Eucryphia* which were often infected with a fungus that was difficult to remove.

Using plant-specific primers, the two internal transcribed spacers (ITS-1 and ITS-2) and the 5.8 S rDNA gene were amplified for all accessions in each of *Eucryphia*, *Griselinia* and *Coriaria*. A complete ITS phylogeny was produced for each.

5.3.2 Low copy-number nuclear genes

The lack of nuclear markers for use in plant phylogeny reconstruction was discussed by Strand *et al.* (1997). To remedy this they designed eight universal primers spanning either single or low copy-number genes. These primers were obtained and screened for their ability to amplify the genes in three species of each of *Coriaria*, *Eucryphia* and *Griselinia*. Additional genes were selected for study if five or more different plant sequences were available from which primer sequences could be designed. On this basis, a total of 13 genes were screened for their technical suitability and for their variability between species. The results obtained after optimising the magnesium concentration gradients and adopting an annealing temperature of 55°C are shown in Table 5.2. As can be seen, experience at amplifying the genes was mixed. For six of the genes (*aat, cam, chs, leafy, pgi* and *tpi*) either no products were amplified or there were multiple products. *Cam* primers, for example, frequently generated at least six products (Fig. 5.1). Although the *chi* gene apparently amplified as a single product, attempts at sequencing it revealed that it was in reality a mixture of several. Attempts to reduce the number of products by increasing the annealing temperature proved to be only partially effective. Therefore work on all of these genes was discontinued and attention directed at the remainder.

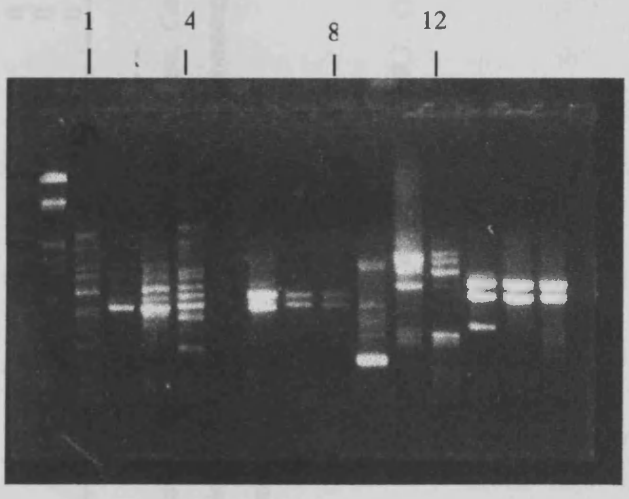

Fig 5.1. Example of gel to illustrate the multiple products from PCR of low copy-number nuclear genes. Lane 1: size ladder. Lanes 2-5: *cam* from *G. scandens*, *E. cordifolia*, *E. glutinosa* and *E. wilkiei*. Lanes 7-9: *g3pdh* products from *C. angustifolia*, *C. intermedia* and *C. myrtifolia*. Lanes 10-12: *g3pdh* products from *G. lucida, G. racemosa and G. scandens*. Lanes 13-15: *g3pdh* products from *E. cordifolia*, *E. glutinosa* and *E. wilkiei*.

Table 5.2. Low copy-number nuclear genes surveyed, indicating the number of fragments amplified by PCR. (m = multiple fragments, sm = smear.)

Species/gene*	*Aat*[1]	*Adh*[2]	*Adh*[3]	*Adh*[1]	*Cam*[1]	*Chi*[1]	*Chs*[1]	*G3pdh*[1]	*LAT 52*[4]	*Leafy*[5]	*Pgi*[1]	*PHYA*[6]	*PHYB*[6]	*PHYB*[3]	*Tpi*[1]
Coriaria angustifolia	0	1	1	0	m	1	0	2	1	2	0	1	1	m	0
C. intermedia	0	1	1	0	1	1	0	2	1	m	0	1	1	m	0
C. myrtifolia	0	1	1	0	1	1	0	2	1	m	0	1	1	m	0
Eucryphia cordifolia	0	0	0	1	m	1	2	2	1	m	0	1	1	m	0
E. glutinosa	1+sm	0	0	1	m	1	1	2	1	m	0	1	1	m	0
E. wilkiei	0	0	0	1	m	1	0	2	1	m	0	1	1	m	0
Griselinia lucida	m	1	0	2	m	1	0	m	1	m	1	1	1	m	0
G. racemosa	0	1	0	0	m	1	0	m	1	m	0	1	1	m	m
G. scandens	0	0	0	0	m	1	0	m	1	2	0	1	1	m	0

**Aat* = aspartate aminotransaminase; *Adh*= alcohol dehydrogenase; *Cam* = calmodulin; *Chi* = chalcone isomerase; *Chs* = chalcone synthase; *G3pdh* = glyceraldehyde 3-phosphate dehydrogenase; *LAT52* = late pollen promoter; *Leafy* = floral developement; *Pgi* = phosphoglucose isomerase; *PHY* = phyochrome; *Tpi* = triose phosphate isomerase.

Primers from:-

[1] Strand et al. (1997); [2] Charlesworth *et al.* (1998); [3] M.R.J. Clokie, designed here; [4] D.Twell (pers. comm.); [6] G.Whitelam (pers. comm.)

Phytochrome (PHY)

Primers had been designed within the department by Prof. Harry Smith's research group for use in understanding the mode of action of the phytochrome gene family. These spanned many regions of *phy*A and *phy*B. The initial remit of the project was to see if these regions would be of phylogenetic interest. These primers amplified a single product. Although good clean sequences were obtained, unfortunately the gene regions that they spanned were not sufficiently variable to resolve species relationships. No variation was detected among the test species in any of the three genera.

Primers for PHYB were therefore designed to span a potentially more variable region at the beginning of PHYB (see Chapter 4 for details). Initially it appeared that a single PCR product was recovered but, on sequencing, it seemed that at least four products were being amplified. In the literature there are reports that 4 and 5 copies occur in gymnosperms and angiosperms respectively (Matthews *et al.* 1997). Phytochrome work was therefore discontinued as it was felt that it would be more profitable to look for genes that belong to a smaller family.

Alcohol dehydrogenase (adh)

Initially, two sets of primers were used in attempts to amplify the *adh* gene. Those designed by Strand *et al.* (1997) only amplified a product in the genus *Eucryphia* (Table 5.2). Unfortunately I had no success in sequencing it. In contrast, primers designed by Charlesworth *et al.* (1998), which were tested on all three genera, produced products consistently in *Coriaria*. The problem of small amounts of target DNA in low copy-number genes was overcome in the case of *adh* in *Coriaria* by means of a touchdown PCR procedure as described in Chapter 4. In some species [(*C. japonica* (lane 5), *C. myrtifolia* (lane 7), *C. napalensis* (lane 9) and *C. pottsiansa* (lane 10); Fig. 5.2] a single product was obtained, and the product was sequenced directly. This product was designated as locus *Adh*-a. In all other species [e.g., *C. angustissima* (lane 4) and *C. sarmentosa* (lane 12); Fig. 5.2], two distinct products or a single smear, were obtained.

Fig. 5.2. Example of gel to show the products from different species of *Coriaria* amplified with the *adh* primers of Charlesworth *et al.* (1998). 1 *C. arborea*; 2 *C. angustissima*; 3 *C. arborea*; 4 *C. angustissima*; 5 *C. japonica*; 6, *C. kingiana*; 7 *C. myrtifolia*; 8 *C. myrtifolia*; 9 *C. napalensis*; 10 *C. pottsiana*; 11 *C. pteridoides*; 12 *C. sarmentosa*; 13 *C. terminalis*.

In the latter cases, because more than one product was being sequenced, direct sequencing resulted in a mass of indistinguishable bases. This problem was overcome by means of cloning, following the method described in Chapter 4. The clones were sequenced using universal M13 primers. The resulting sequences were aligned with the species that were sequenced directly and the locus that aligned best with *Adh*-a was used in the phylogenetic analysis. The second locus, *Adh*-b, was not studied further. Having acquired *Coriaria* sequences for *Adh*-a, primers specific to these were then designed for use on the remaining species (*Adh* superscript 3 in Table 5.2). The cloning worked well and the specific primers were effective for species *C. arborea*, *C. intermedia*, *C. pottsiana* and *C. terminalis* , but unfortunately and despite extensive effort, it was not possible to obtain amplified products for the other species.

Glyceraldehyde-3-phosphate-dehydrogenase (G3pdh)

Although in *Manihot esculenta* and the other species from which the *g3pdh* primers were designed only one copy of the gene was present, in *Coriaria* and *Eucryphia* two products were consistently amplified (Table 5.2, Fig. 5.3). In both genera, the two products were approximately 1000 bp and 800 bp in length, a size difference that allowed them to be separated manually by cutting them directly from a gel. They were respectively designated *G3pdh*-1 and *G3pdh*-2. In the few cases where the product could only barely be detected on the gel, the weak product was used as a starting template material for a second PCR.

Fig. 5.1 shows a typical gel, depicting the double products for *Coriaria* and *Eucryphia*, and the multiple products for *Griselinia* for the *G3pdh* primers. Figure 5.3 illustrates the separated loci which after extraction from the gel were run on a further gel to check that only one product was present.

Despite the promising results, regrettably, the products from *Coriaria* proved impossible to sequence. In contrast, however, in the case of *Eucryphia* the approach worked well and both loci were sequenced for all accessions of *Eucryphia* species plus their outgroups (see Chapter 6 for details).

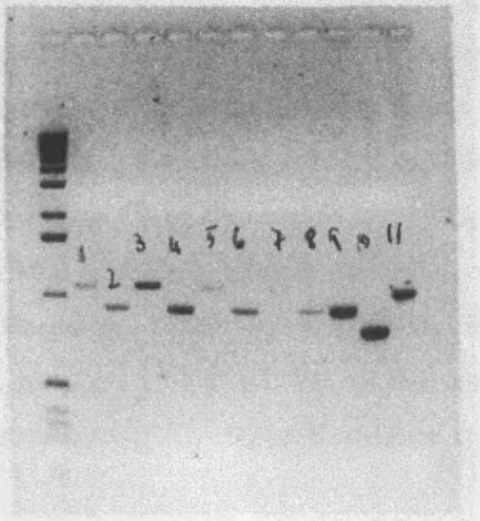

Fig. 5.3. Example of a gel to show the purified products for g*3pdh*. Left-hand lane = 1 kb size ladder. Lanes 1 - 2: *Eucryphia cordifolia*; lanes 3 - 4: *E. glutinosa*; lane 5 - 6: *E. milliganii* ssp. *Milliganii*; lanes 7 - 8: *E. wilkiei*; lanes 9 - 10: *Coriaria angustifolia*; lane 11: *C. intermedia*. In all cases the g*3pdh*-1 product is shown first followed by that of *G3pdh*-2.

CHAPTER 6

THE GENUS EUCRYPHIA

6.1 Introduction

There are seven species in the genus *Eucryphia*, consisting of bushy shrubs or trees that may grow to 40 m high. The leaves are opposite, simple or pinnate, with small, basally connate, interpetiolar stipules. The blades are thick and leathery and the margins entire to serrate or dentate. In pinnate leaves, the leaflets are opposite and usually range from 3-13 per leaf. Floral morphology is characterised by showy, white, tetramerous flowers that are usually solitary and axillary. Each flower has 4-14 multi-ovulate carpels and numerous stamens. The fruits are dry capsules that release dry, winged seeds. Two species are found in Chile (Archer *et al.*, 1993; Hoffmann, 1997; Holdgate, 1961; Rodriguez *et al.*, 1983) and five inhabit Australia, with three on the mainland and two in Tasmania (Barnsley, 1983; Bausch, 1983; Dress, 1956; Forster and Hyland, 1997; Hill and Carpenter, 1991). All the species are evergreen, except for *E. glutinosa*, which is winter-deciduous (Hill, 1991a; Rodriguez et al., 1983).

Aspects of the morphology of the genus are illustrated in Figs. 6.1a and 6.1b, and a general morphological description can be found in Hutchinson (1973). Systematic anatomical studies were initiated by Dickison (1978) and have provided a recent and extensive review of the anatomical work on the genus. The biochemistry of tissue flavonoids was studied by Bate-Smith *et al.* (1967) and recently a fuller analysis of flavonoids found in the bud and leaf exudates for all seven species has been performed by Wollenweber *et al.* (2000).

Hamel (1956) reported sporophytic counts of 2n = 32 chromosomes in *E. lucida*, *E. glutinosa*, *E. cordifolia* and *E. moorei.* He suggested that the presence of "double pairs" probably indicates that *Eucryphia* is tetraploid with x = 8. More recently, however, Goldblatt (1976) reported a sporophytic number of 2n = 30 chromosomes in *E. lucida.*

a **b**

Fig. 6.1. The stature, relative size, bark texture, flower and fruit morphology of *Eucryphia cordifolia* (**a**) and *Eucryphia glutinosa* (**b**). From Hoffmann, (1997).

A list of the currently accepted names and a synonymy is provided in Table 6.1, data for which were drawn from the revisions by Gilg (1925), Bausch (1938), Dress (1956) and from more recent work by Barnes *et al*., (2000).

Table 6.1. List of currently accepted names and synonyms in *Eucryphia*.

Eucryphia Cav., Ic. 4: 48, t. 372 (1797).

Chilean species

Eucryphia cordifolia Cav., Ic. 4: 49, t. 372 (1797).

Eucryphia cordata Loud., Encycl. Pl. 2: 1394 (1855), sphalm.

Pellinia chilensis Molina, Sagg. Chil. ed. 2, 160 (1810).

Pellinia cordifolia Molina, Sagg. Chil. ed. 2, 290 (1810).

Eucryphia patagonica Speg., in Anal. Mus. Buenos Aires 7: 281, nr. 9 (1902)

Vernacular names: roble de Chile, roble de Chilo', ulmo, toz, voyencum (Chile), muermo (Chilotes), ngulngo, gnulgnu (Mapuche Indians). (N.B. ulmo and muermo are Spanish derivatives from ngulngo).

Eucryphia glutinosa (Poepp. & Endl.) Baill., Hist. Pl. 1: 401 (1867-9), in adnot.

Fagus glutinosa Poepp. & Endl., Nov. Sp. et Gen. 2: 68, t., 194 (1838 – 45).

Fagus glandulosa Poepp. & Endl., Reiche, Flor. Chil. 1:28 (1895). sphalm.

Eucryphia glandulosa Reiche, Fl. Chile 1: 268 (1896), sphalm.

Eucryphia pinnatifolia C. Gay, Fl. Chil. 1: 852. t. 8.

Eucryphia pinnatifida Schnizl., Icon. Fam. Nat. Veg. 3: t. 4 (1854), sphalm.

Vernacular names: guindo santo, guindo, –irre, –ire, nirrhe (Chile).

Australian species

Eucryphia jinksii P.I. Forst., in Austrobaileya 4: 592 (1997).

Eucryphia moorei F. Muell., Fragm. 4: 2. (1863/4).

Vernacular names: dog-rose tree, wingecarribbee, white sallow, acacia, plumwood, plum tree (New South Wales), stinkwood (Victoria), pinnate boronia (Australia).

Eucryphia wilkiei B. Hyland, in Austrobaileya 4: 590 (1997).

Tasmanian species

Eucryphia lucida (Labill.) Baill., Hist. Pl. 1: 402 (1867-9), in adnot.
Carpodontos lucida Labill., Relat. Voy. 2:16 (1798)
Eucryphia billardieri Spach, Hist. Veg. Phan. 5: 345 (1836).
Vernacular names: pinkwood, leatherwood (Tasmania).

Eucryphia milliganii J.D. Hook., Fl. Tasm. 1: 54 (1855).
Eucryphia billardieri var. *milliganii* (J.D. Hook.) Bentham, Fl. Austral. 2:446 (1864)
Eucryphia lucida Druce var. *milliganii* Summerhayes, Handlist Tr. Shr. Roy. Bot. Gard. Kew (1934).
Vernacular name: Milligan's leatherwood (Australia).

subsp. ***milliganii***
subsp. ***pubescens*** R.W. Barnes, G.J. Jord., R.S. Hill & C.J. McCoull in Aust. Journ. Bot. 48(4): 488 (2000).

Hybrids

Eucryphia x ***hillieri*** Ivens, in J. Roy. Hort. Soc., Lond. 88: 342 (1953). (*E. lucida x E. moorei*, e.g. *E.* 'Winton').
Eucryphia x ***hybrida*** Bausch, in Kew Bull., 1938, 331. (*E. lucida x E. milliganii*).
Eucryphia x ***intermedia*** Bausch, in Kew Bull., 1937, 193. (*E. glutinosa x E. lucida*, e.g. *E.* 'Rostrevor')
Eucryphia x ***nymansensis*** Bausch, in Kew Bull., 1938, 324. (*E. cordifolia x E. glutinosa*, e.g. *E.* 'Mount Usher', *E.* 'Nymansay')
E. cordifolia x E. lucida, e.g. *E.* 'Penwith' (incorrectly distributed as a variant of *E. hillieri*).

The genus was first described in 1797 by Antonio José Cavanilles who had studied Chilean herbarium specimens which he named *Eucryphia cordifolia* (Cavanilles, 1797). The name is derived from the Greek 'eu' and 'kryphia', which means well covered, or well-hidden. This refers to the protective sepals that surround the corolla.

Following its discovery in 1793, the Tasmanian *E. lucida* was described as *Carpodontos lucida* by La Billardière in 1799. This species was transferred to *Eucryphia* by Spach (1836), based on a study of *Carpodontos* and of *Eucryphia* which led him to unite the two genera under the earlier name. As La Billardière's generic name fell into synonymy, Spach honoured him with the substitute specific epithet *Billardieri,* and this species was often known as *E. billardieri* until Baillon (1867-69) resurrected the earlier name.

Because it was first found when not flowering, *E. glutinosa* started its nomenclatural life as a beech tree: *Fagus glutinosa* (Poeppig and Endlicher, 1839). A few years later Claude Gay saw it flowering by the Biobio river in the province of Concepción in Chile, and recognised its relationship with the other Chilean *Eucryphia* species. He described it as *E. pinnatifolia,* but Baillon (1867-69) reduced this name to synonymy under *E. glutinosa.* The specific epithet refers to the sticky exudate that occurs on the young leaves and buds.

In 1855, Hooker described *E. milliganii* in his 'Flora Tasmaniae'. It was named after Dr. Joseph Milligan of Hobarton, who was a keen collector and was thought to have been the original discoverer of the species. *Eucryphia milliganii* is morphologically very similar to *E. lucida,* but smaller in height, and in leaf and flower size. There has been much debate as to its taxonomic status. Bausch (1938) was emphatic in considering *E. milliganii* as a distinct species, but both Dress (1956) and Dickison (1978) queried this view, suggesting it may be an alpine variant of *E. lucida.* More recently, Hill (1991a) favoured separate species status on the grounds that in areas where the two taxa occur sympatrically, they are morphologically distinct with no intermediates present. The two subspecies recognised by Barnes *et al.* (2000) are discussed in sect. 6.4.2.

The mainland Australian *Eucryphia* species have a far less complicated taxonomic history. Charles Moore discovered what was later named *E. moorei* in the mountainous forests of New South Wales. The species was described by Mueller in 1863 in his 'Fragmenta Phytographiae Australiae'. The other two species were described recently from Queensland in a paper by Forster and Hyland (1997) who named one species each: *E. wilkiei* B. Hyland was named after the orchidologist Jack Wilkie, who discovered it in, 1970, and *E. jinksii* P. Forster was named in recognition of the sharp eyes and tree climbing prowess of David Jinks, who found it in 1994 (Forster and Hyland, 1997).

6.2 Relationships to other families and genera

Depending on the relative importance attributed to different taxonomic characters, *Eucryphia* has at various times been placed in at least 13 different families (Bate-Smith *et al.*, 1967), the most common placements being: Saxifragaceae, Cunoniaceae, Rosaceae and Hypericaceae (Dress, 1956). The first time *Eucryphia* was considered to be a separate family, Eucryphiaceae, was in 'Genera Plantarum' (Endlicher 1838). Based on morphological similarities and on anatomical structures, Bausch (1938) was the first to suggest that Eucryphiaceae was allied to the Cunoniaceae, and most taxonomists have agreed (Cronquist, 1988; Cronquist, 1973; Cronquist, 1981; Takhtajan, 1969; Takhtajan, 1996; Thorne, 1983; Thorne, 1992).

Recently this idea of *Eucryphia* was taken further by Dickison and Rutishauser (1990) who, based on anatomical evidence, suggested that *Eucryphia* does not warrant a separate family, but should be included within the Cunoniaceae. A phylogenetic survey of Cunoniaceae using morphological data agreed with this (Hufford and Dickison, 1992), and the possession of a decussate leaf arrangement was considered to be a synapomorphy for the family (Kalkman, 1988; Dickison and Rutishauser, 1990). Another morphological feature which *Eucryphia* shares with other members of the Cunoniaceae is the presence of interpetiolar stipules with colleters (Dickison, 1978). Other evidence supporting a close relationship of *Eucryphia* with Cunoniaceae includes pollen morphology (Hideux and Ferguson, 1976) and the presence of the flavonol azaleatin (Bate-Smith *et al.*, 1967; Jay, 1968). Molecular data based on *rbc*L (Chase *et al.*, 1993; Morgan and Soltis, 1993) and on sequences from 18S rDNA (Soltis and Soltis, 1997) also support *Eucryphia* being nested within the Cunoniaceae.

In their phylogenetic analysis of the Cunoniaceae, Hufford and Dickison (1992) produced a cladogram, based on morphological data, that illustrates the position of *Eucryphia.* According to this study, *Eucryphia* is 'sandwiched' between a basal group of genera (which includes *Gillbeea*) and two other main groups, to which it is sister (Fig. 6.2). One of these groups includes *Ceratopetalum* and *Schizomeria*, and the other *Bauera.* I have therefore selected representatives of these four genera to serve as potential outgroups in the present study, viz. *Gillbeea adenopetalum, G. benthamii, Ceratopetalum apetalum, C. gummiferum, Schizomeria ovata, Bauera rubioides and B. sessiliflora.*

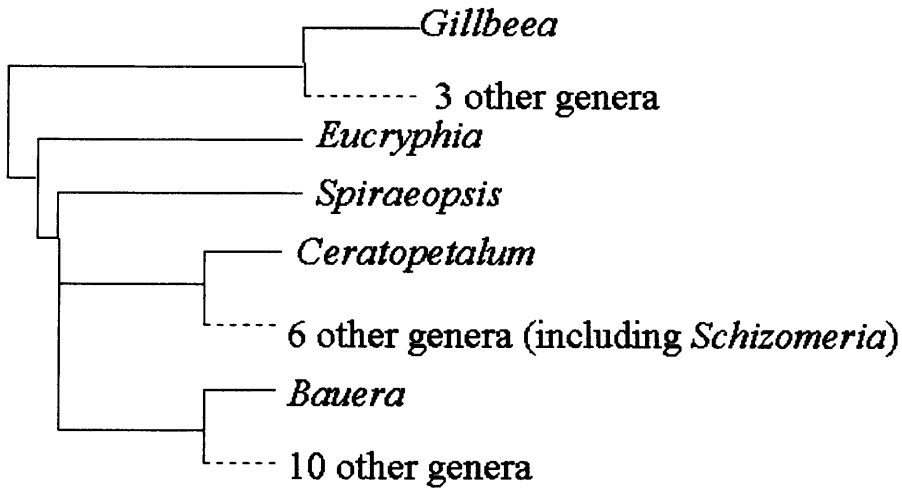

Fig. 6.2. Cladogram to show relationships between *Eucryphia* and other genera of the Cunoniaceae. Modified from Hufford and Dickison (1992) and based on a strict consensus of 47 equally most parsimonious trees.

6.3 Reproductive biology

The flowers of *Eucryphia* are actinomorphic, about 5 cm across, hermaphrodite and arise singly from the axils of the leaves (Fig. 6.3a). They superficially resemble those of *Hypericum*. The partially imbricate calyx of four sepals forms a false operculum which falls away to reveal four imbricate, white petals and numerous stamens, all of which are attached below the superior ovary. The stamens are differentiated into a distinct filament and anther. The anthers contain a connective and four locules. Each microsporangium has a well developed endothecium with banded thickenings. Before dehiscence the partition between the two locules of the same theca break down to form a single chamber. The dehiscence is latrorse (Dickison, 1978).

The gynoecium is composed of 4-14 (occasionally 18) multi-ovulate carpels which are united proximally to form a compound ovary but distally free, each with a slender style. Placentation is axile. The locules contain a few pendulous ovules that are attached at their inner edge. The ovary is densely pubescent in some species.

The flower- and leaf-buds are often conspicuously resinous. The buds are protected by this sticky resin (Fig. 6.3b). The resin is particularly conspicuous in the two Tasmanian species, *E. lucida* and *E. milliganii* (Forster and Hyland, 1997). The exudates are thought

to originate in the colleters (mucilagenous hairs) that are present in all *Eucryphia* species and which are positioned on the inner surface of the stipules that protect the developing buds (Dickison, 1978).

Fig. 6.3a. Flower of *Eucryphia milliganii* **6.3 b.** Resinous buds of *Eucryphia lucida* [Picture taken from Missouri botanical gardens website]

In *E. lucida*, it is known that the abundance of flowering varies both between and within localities. Flowering generally does not occur until the plants reach 75 years of age, and maximum flowering, and therefore nectar production, occurs in trees between 175–- 210 years old. Where *E. lucida* is not shaded it produces far more flowers. Wet springs and autumns are thought to increase flowering (Australian National University website, 1998).

The flowers of *Eucryphia* are pollinated by insects that visit in pursuit of the nectar which is present in nectaries at the base of the filaments. Honey obtained from *Eucryphia* flowers both in Chile and in Tasmania is distinctive and highly prized (Hoffman, 1997; Australian National University website, 1998).

Thus far, no work has been published on the breeding system of *Eucryphia* species, although presumably in view of their importance to insects as a source of nectar, they experience a considerable amount of outcrossing. But it is not known whether self-

incompatibility is present in the genus. Obviously this attribute has important implications regarding the initial establishment of *Eucryphia* populations.

In terms of possible dispersal at the sporophyte stage of the life-cycle, the fruit of *Eucryphia* is a woody or leathery, septicidal capsule which, at maturity, splits into 5-12 separate, boat-shaped, valves that are attached basally to the central column (Fig 6.4). Each valve contains a few, oblong, winged seeds.

Fig. 6.4. Leathery capsules of *Eucryphia lucida*
[Picture taken from Missouri botanical gardens website]

Radford *et al.* (1974) listed many South American-Australasian genera which, due to their attractive fruit, they thought were capable of long-distance dispersal by birds. They then stated "There are several taxa however, which, because of large, dry or heavy cones, fruits, or seeds, are not logically explained by bird-carriage or other forms of long-distance dispersal. Among these are *Araucaria*, *Eucryphia*, *Laurelia*, *Nothofagus*, and the four genera of Proteaceae."

It certainly seems unlikely that the large, dry capsule of *Eucryphia* would be an effective diaspore over long distances. The winged seeds, however, although presumably unattractive to birds, might travel some distance in air currents. This would be dependent

on factors such as the height of release and wind speed. Regrettably there are no data on these points, nor on whether the seeds can survive in sea-water (see Chapter 2 for a general review of dispersal).

6.4 Distribution and ecology

Schematic distribution maps for *Eucryphia* species are shown in Figs 6.5a - 6.5b. Two species of *Eucryphia* occur in Chile, three in Australia and two in Tasmania.

6.4.1 Chilean species (Fig. 6.5a)

Of the Chilean species, *E. cordifolia* is a dominant plant in the temperate broad-leaved evergreen forests which occur along the very wet zone at an altitude of 700 m between approximately 38° S and the island of Chiloe at around 42° S (Hoffmann, 1997; Rodriguez *et al.*, 1983; Veblen *et al.*, 1983). *E. cordifolia* also has a more limited distribution at lower altitudes (below 400 m) in the Andean Cordillera from approximately 39° 30′ - 41° 30′. Generally the range of *E. cordifolia* is up mountains to the edge of glaciers.

E. glutinosa is found between an altitude of 250 and 900 m from the Province of Linares to the Province of Malleco (about 36° S – 39° S); it is particularly abundant at an altitude of 700 m (Rodriguez et al., 1983). *E. glutinosa* is much less common than it once was and, although formerly it was widely distributed in the Andean foothills, it is currently only found abundantly in the Chillan region in the centre of Chile (Rodriguez, Melica pers. comm.).

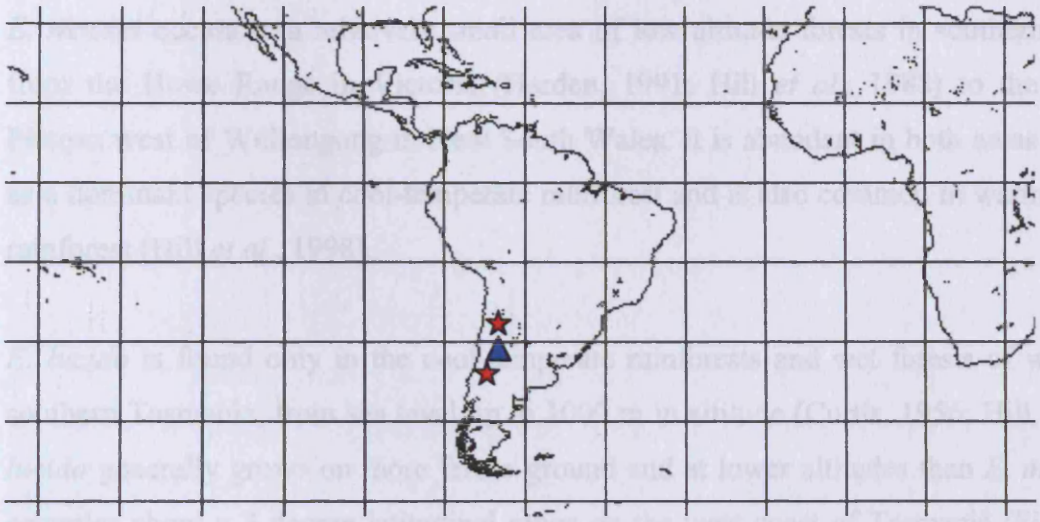

Fig. 6.5a. Map of South America showing the approximate locations of the Chilean species of *Eucryphia.* Distribution of *E. cordifolia* ★ and of *E. glutinosa* ▲ .

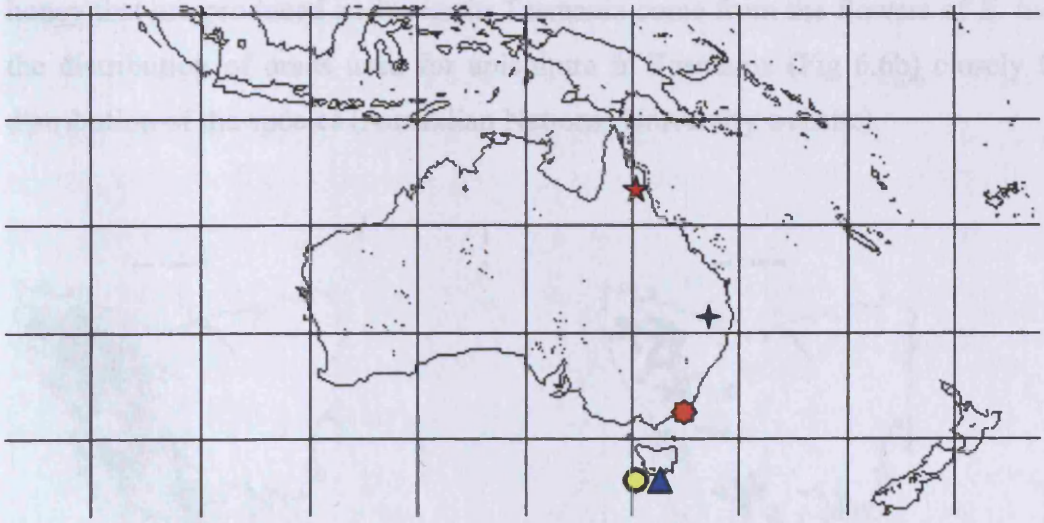

Fig. 6.5b. Map of Australia showing the approximate locations of *E. wilkiei* ★, *E. jinksii* ✦, *E. moorei* ●, *E. lucida* ○, and *E. milliganii* ▲ .

*6.4.2 Australasian species (*Fig. 6.5b*)*

Of the Australasian species of *Eucryphia,* three occur in temperate regions: *E. moorei, E. milliganii* and *E. lucida*; and two are found in subtropical (*E. jinksii*) and tropical (*E. wilkiei*)areas.

E. moorei occurs in a relatively small area of low altitude forests in southern Australia, from the Howe Range in Victoria (Harden, 1991; Hill *et al.*, 1988) to the Woronora Plateau west of Wollongong in New South Wales. It is abundant in both areas and grows as a dominant species in cool-temperate rainforest and is also common in warm-temperate rainforest (Hill *et al.*, 1998).

E. lucida is found only in the cool-temperate rainforests and wet forests of western and southern Tasmania, from sea level up to 1000 m in altitude (Curtis, 1956; Hill, 1991a). *E. lucida* generally grows on more fertile ground and at lower altitudes than *E. milliganii*. It occupies about a 3 degree latitudinal range on the west coast of Tasmania (Fig. 6.6a). In the wetter forests that receive 1000 – 2000 mm per annum *E. lucida* grows as an understory tree (Australian National University website). Although able to grow at 1000 m, only 3 % of all the records come from above an altitude of 800 m (Read and Busby, 1990). An interesting commercial point related to this is that 70% of the 1000 tonnes of honey that are produced each year in Tasmania come from the flowers of *E. lucida*. Thus the distribution of areas used for apiculture in Tasmania (Fig 6.6b) closely follows the distribution of the species (Australian National University website).

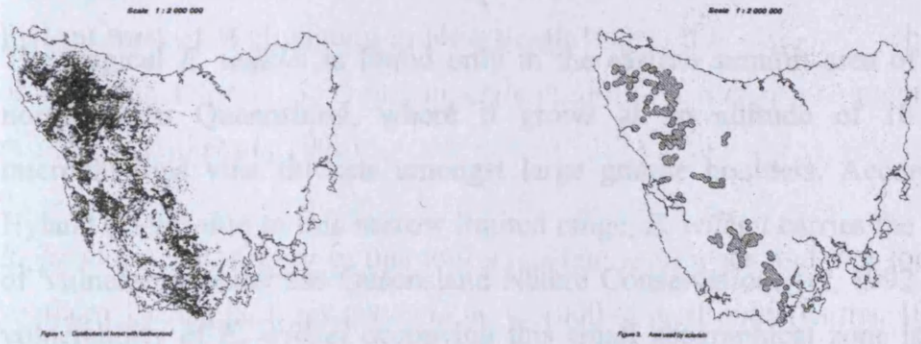

Fig. 6.6. a Distribution of *E. lucida* in Tasmania; **b** Areas used for apiculture.

The distribution of *E. milliganii* overlaps that of *E. lucida* but is more restricted. It is generally confined to areas of land that are more exposed, less fertile and at higher altitudes. Recently Barnes *et al.* (2000) subdivided *E. milliganii* into two subspecies on the grounds that there are distinct northern and southern Tasmanian phenotypes. The southern *E. milliganii* plants differ from those of the north in that they have ovate to

elliptic leaves with a pubescent abaxial surface and dense marginal trichomes (rather than the oblong, sparingly pubescent to glabrous leaves found in the north). Barnes *et al.* (2000) recognised the southern variant as subspecies *pubescens*.

E. milliganii subs. *milliganii* is found in the west and south-west of Tasmania although it very rarely exceeds 43°S. The main habitat for this subspecies is in high altitude (above 800 m), high rainfall alpine shrubbery and rainforest. The geology consists of quartzite, sandstone and dolerite. Occasionally this subspecies is found in lowland, riparian, cool-temperate rainforest, in tea-tree scrub and sometimes it is found with *E. lucida* (Barnes *et al.*, 2000).

E. milliganii subsp. *pubescens* is found over a wide range of altitudes on sites that are mostly south of 42° 30′S. Again, it grows on soils composed of quartz, sandstone and dolerite and is common in the same habitat of high altitude, high rainfall, alpine shrubbery and rainforest. Also similarly to the sister subspecies it is occasionally found in lowland, riparian, cool-temperate rainforest or tea-tree scrub, often near *E. lucida* (Barnes *et al.*, 2000).

The tropical *E. wilkiei* is found only in the eastern summit area of Mt Bartle Frere, in north-eastern Queensland, where it grows at an altitude of 10 - 1400 m in the microphyllous vine thickets amongst large granite boulders. According to Forster and Hyland (1996) due to this narrow limited range, *E. wilkiei* carries the conservation coding of 'vulnerable' under the Queensland Nature Conservation Act, 1992 (Jessup, 1994). The vulnerability of *E. wilkiei* occupying this small geographical zone is exacerbated by the fact that the national park in which it occurs is subject to sporadic damage from cyclones and storms (Forster and Hyland, 1997). *E. wilkiei* seems to like the cyclone damage as it regenerates in the areas where it has occurred (Barnes, pers. comm.).

Only one population of the subtropical *E. jinksii* is known and it occurs in south-eastern Queensland, near Springbrook, in the MacPherson range (Forster and Hyland, 1997). It occurs at an altitude of 770-800 m on rhyolite rocks and amongst dense vegetation termed by the authors 'complex notophyll vine forest' (Forster and Hyland, 1997). This population of *E. jinksii* consists of several hundred mature individuals in an area of about two

hectares, but is concentrated in a narrow band spanning a steep and rocky river. Although fruit set appears to be abundant, there is no evidence of recent seedling recruitment (Forster and Hyland, 1997). Although some individuals occur on private land, most live within Springbrook National Park. This species is very susceptible to changes in water catchment structure such as could occur on the private land above the zone where the main population lives. Therefore under the Queensland Nature Conservation Act, 1992, the species has been categorised as 'endangered' (Forster and Hyland, 1996).

6.5 Evolution in *Eucryphia*

6.5.1 Fossil record

Eucryphia fossils have been found over much of Tasmania, particularly in the north-west, and in Lake Bungarby, which is in south-eastern New South Wales, Australia (Table 6.2 and Fig. 6.7). This work has been carried out at the university of Tasmania, primarily by Hill (1991a) and Barnes and Jordan (2000). The oldest *Eucryphia* fossils known are about 60 million years old and were found in Lake Bungarby. Five leaf compressions were assigned to *Eucryphia*, and described as the fossil taxon *E. falcata*. They have compound leaves whose leaflets have serrated margins.

Partially complete *Eucryphia* leaves and leaflets have been recovered from Middle – Upper Eocene sediments at Loch Aber in north-eastern Tasmania, (Hill, 1991a; Barnes and Jordan, 2000). These resemble *E. milliganii* and *E. lucida* but have a greater trichome density and have been assigned to *E. aberensis*. Again this species has compound leaves but in this case the margins of the leaflets are either serrate or entire. Recently, additional material has been found from this location as well as further specimens from a new locality of slightly more recent age, viz. Little Rapid River in Early Oligocene deposits. A capsule was also found from this locality but it was not attached to the *E. aberensis* leaf macrofossils. Because it was poorly preserved it was not given species status (Barnes and Jordan, 2000).

Table 6.2. Accepted records of *Eucryphia* macrofossils from south-eastern Australia, modified from Barnes and Jordan (2000). N.B. Three separate fossiliferous deposits occur at Regatta Point: Early Eocene sediments overlain by glacial outwash that contains clay clasts of Early Pleistocene and Early-Middle Pleistocene age.

Species name	Macrofossil	Geological age	Geographical distribution
E. falcata R.S. Hill	Leaf compressions	Late-Paleocene	Lake Bungarby
E. microstoma R.S. Hill	Leaf compressions	Early-Eocene	Regatta Point
E. aberensis R.S. Hill	Mummified leaf fragments	Middle to Late- Eocene	Loch Aber
E. aberensis R.S. Hill	Mummified leaf fragments	Early-Oligocene	Little Rapid River
E. sp. 'LRR1' R.W. Barnes and G.J. Jord.	Incomplete mummified capsule	Early-Oligocene	Little Rapid River
E. mucronata R.W. Barnes and G.J. Jord.	Leaf compression	Early-Oligocene	Wilsons Creek, central Tasmania
E. reticulata R.W. Barnes and G.J. Jord.	Mummified capsule	Early-Oligocene	Lea River
E. leaensis R.W. Barnes and G.J. Jord.	Mummified leaf fragment	Early-Oligocene	Lea River
E. sp. 'Leven' R.W. Barnes and G.J. Jord.	Dispersed cuticle	Early-Oligocene	Leven River
E. lucida	Mummified leaves	Early-Pleistocene	Regatta Point
E. milliganii ssp. *milliganii*	Mummified leaves	Early-Pleistocene	Regatta Point
E. sp. (Jordan *et al.*, 1995)	Dispersed cuticle	Early-Middle Pleistocene	Regatta Point
E. lucida	Mummified leaves	Early-Pleistocene	Regency Formation
E. milliganii ssp. *milliganii*	Mummified leaves	Early-Pleistocene	Regency Formation
E. sp. (Jordan *et al.*, 1991)	Dispersed cuticle	Late-Pleistocene	Melaleuca Inlet
E. lucida	Mummified leaves	Late-Pleistocene	Pieman Dam

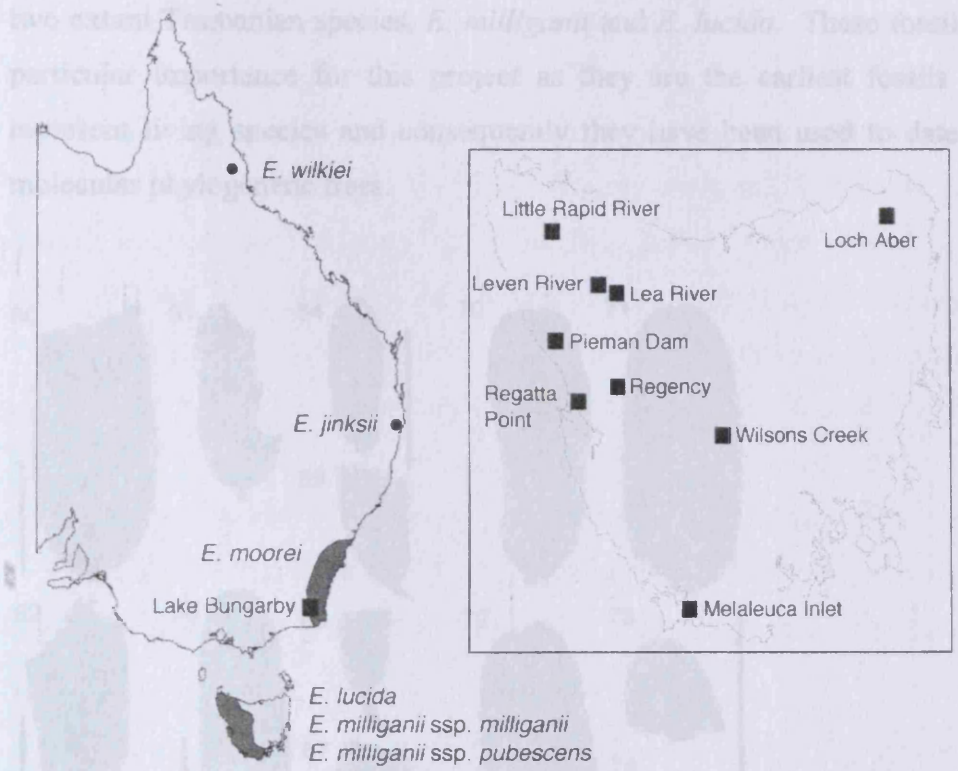

Fig. 6.7. Fossil and present-day distributions of *Eucryphia* in the eastern mainland of Australia and in Tasmania (in box). The fossil locations are shown as squares and the extant locations are shown as shaded areas where the distribution is widespread and as circles where the distribution is discreet (From Barnes and Jordan, 2000).

The taxon *E. microstoma* has been disregarded as a species of *Eucryphia* (Barnes *et al.* 2001). The oldest fossil capsule (Early Oligocene) of *Eucryphia* has been described as *E. reticulata*, from Lea River in Tasmania (Barnes and Jordan, 2000). The capsule is similar to those found in the extant South American species *E. cordifolia* and *E. glutinosa* in that it is large with many valves.

At the same site where the capsule was found, a leaf macrofossil was also found. Due to the highly falcate nature of this leaf, it is thought to be a leaflet from a compound leaf. Although it is likely that the leaflet is derived from the same plant as the capsule, as no organic connection was present, the leaflet was described as a new species, *E. leaensis* (Barnes and Jordan, 2000).

More recent macrofossils have been found from Early Pleistocene sediments at Regatta Point in western Tasmania. Work by Barnes and Jordan (2000) has shown that, on the basis of foliar hair distribution patterns and density, these fossils are conspecific with the

two extant Tasmanian species, *E. milliganii* and *E. lucida*. These fossils (Fig. 6.8) are of particular importance for this project as they are the earliest fossils to be found that represent living species and consequently they have been used to date the nodes in the molecular phylogenetic trees.

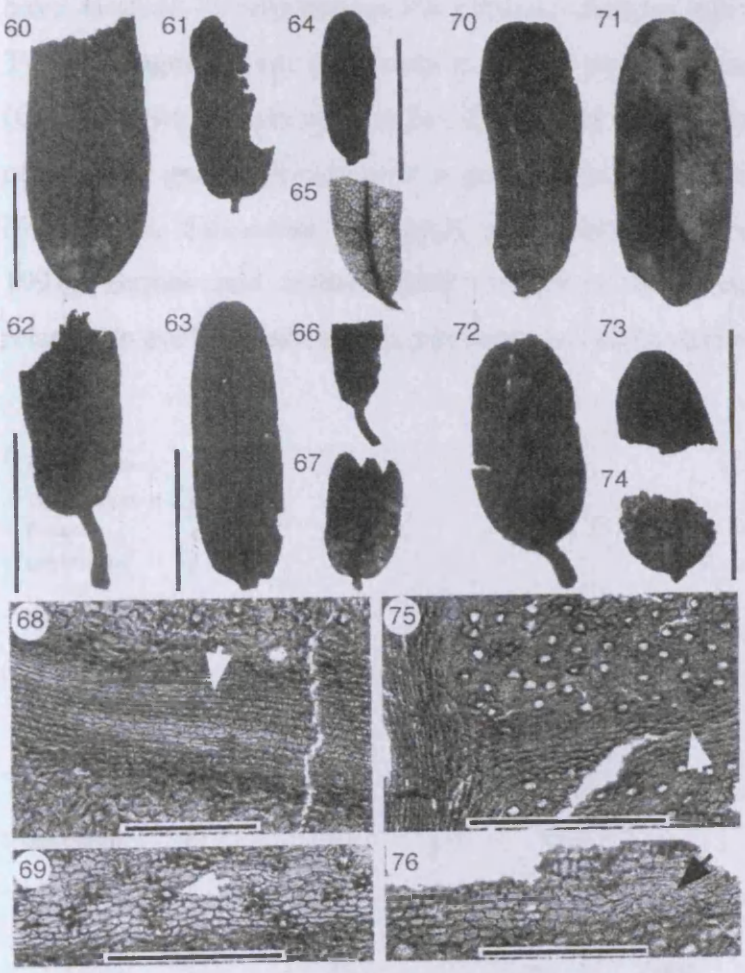

Fig 6.8. Fossil *E. lucida* and *E. milliganii* leaves and abaxial cuticle from Early-Pleistocene sediments at Regatta Point. **60-67**, *E. lucida* macrofossils. **68-69**, *E. lucida* abaxial surfaces (note glabrous midrib and densely arranged trichome respectively). **70-74**, *E. milliganii* ssp. *milliganii* macrofossils. **75-76**, Abaxial surfaces (note glabrous midrib and single trichome base respectively). (From Barnes and Jordan, 2000).

6.5.2 Some evolutionary trends in Eucryphia

The fossil record allows an insight into probable routes of morphological evolution within the genus. The earliest, and therefore possibly ancestral, leaves are compound and the leaflets have serrate margins (Taylor and Hill, 1996). More simple leaves are thought to have evolved in response to the climatic changes which occurred during the Cenozoic. These changes include a decrease in rainfall and temperature and an increase in seasonality (Quilty, 1994; Barnes and Jordan, 2000). The same types of modification are observed in other plant genera which have a good fossil record in south-eastern Australia, such as *Nothofagus*, *Casuarina*, *Acmopyle* and *Dacrycarpus* (Hill, 1994; Hill and Carpenter, 1991). Barnes and Jordan (2000) reviewed the evidence from *Eucryphia* fossil data relating to evolutionary trends, and identified three unrelated ones (Fig. 6.9) as follows.

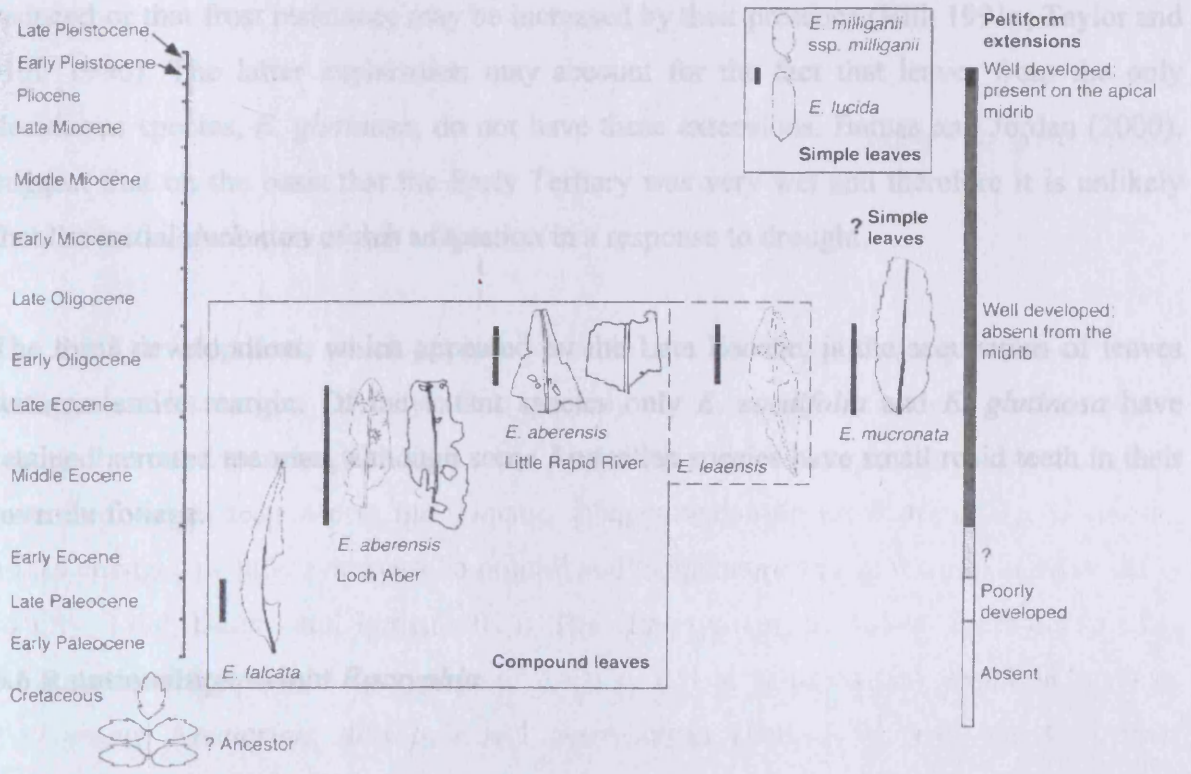

Fig 6.9. Evolutionary trends in leaf morphology in *Eucryphia* as indicated by the fossil record. From Barnes and Jordan (2000).

The first and oldest development is a reduction in the number of leaflets within the compound leaf. This is thought to have occurred during the Cretaceous or Paleocene period. This reduction in leaflet number is also observed in other genera of Cunoniaceae (Barnes and Jordan, 2000).

The next trend that occurred was the development of peltiform waxy extensions on the abaxial surface of the leaf. Where these extensions are present, each epidermal cell has a single, shield-shaped, waxy protrusion that may be up to 35 µm in length. *En masse* these extensions make it hard to see the stomata, and they give *Eucryphia* leaves their characteristic white and waxy appearance (Hill, 1991a). They may be found on the leaf lamina, the midrib or on the veins (Taylor and Hill, 1996). Such cuticular extensions are thought to have developed during the Early or Middle Eocene. The ecological significance of these extensions is not clear but it is thought that transpirational water loss may be reduced or that frost resistance may be increased by their presence (Hill, 1991a; Taylor and Hill, 1996). The latter explanation may account for the fact that leaves from the only deciduous species, *E. glutinosa*, do not have these extensions. Barnes and Jordan (2000), suggest that on the basis that the Early Tertiary was very wet and therefore it is unlikely that the initial evolution of this adaptation in a response to drought.

The third development, which appeared by the Late Eocene, is the acquisition of leaves with an entire margin. Of the extant species only *E. cordifolia* and *E. glutinosa* have retained serrated margins, although some Australian species have small rosid teeth in their juvenile foliage.

6.6 Relationships within *Eucryphia*

6.6.1 Morphological evidence

Taylor and Hill (1996) conducted a phylogenetic analysis of six extant species of *Eucryphia* (not *E. jinksii*) and in a second analysis included three fossil species, *E. aberensis*, *E. microstoma* and *E. falcata*. Their data consisted of 27 morphological characters. The leaf characters examined were: leaf and petiole shape, leaf structure, lamina margin, leaf areoles, venation, mucilage cells, cuticular extensions, and trichomes. In the case of the living species, wood and reproductive characters were also analysed.

Wood characters assessed were features relating to the perforation plates, vessel elements, crystals and rays. The reproductive structures examined were the petals, sepals, ovary and pollen grains (Taylor and Hill, 1996). The characters were polarised with reference to an outgroup, *Gillbeea*, a closely related genus in the Cunoniaceae.

The morphological data for living species ultimately gave rise to one optimal cladogram with a length of 57 (Fig. 6.10). This shows that the Australian species constitute a monophyletic group to which the two South American species *E. cordifolia* and *E. glutinosa* are basal and paraphyletic. All Australasian species have entire-margined leaves, derived leaf apices (mucronate, rounded or obtuse), helical thickening in the xylem, prismatic crystals in the wood parenchyma and unseriate to locally biseriate wood rays (Taylor and Hill, 1996).

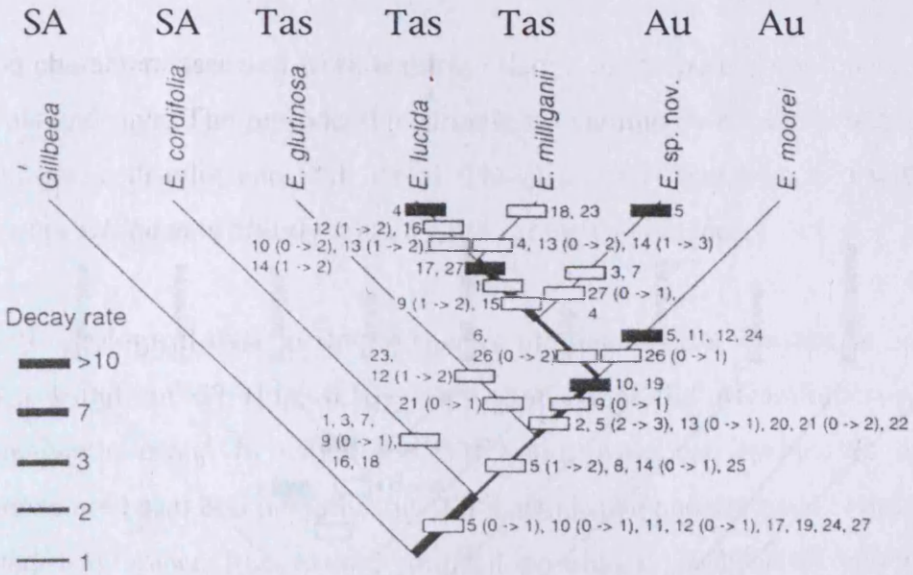

Fig. 6.10. Consensus of extant *Eucryphia* species based on morphological data, figure from Taylor and Hill, (1996). *E.* sp. nov = *E. wilkiei*. *E. jinksii* was unknown at the time. The numbers represent morphological characters which define the groups; details of which are unimportant here.

Within the monophyletic Australian group, the Tasmanian species form a monophyletic group, united by the presence of cuticular extensions on both the lamina midrib and secondary veins of the abaxial leaf surface and by the presence of a glabrous petiole. *E. wilkiei* (referred to as "sp. nov." by Taylor and Hill (1996)) is basal to the Tasmanian species, forming a clade with them by virtue of its unthickened and well-developed pollen muri. *E. moorei* appears to have undergone a disproportionate number of character-state reversals which suggests that this species requires further investigation (Taylor and Hill, 1996). The position of *E. cordifolia* is basal to all the other species whereas that of *E. glutinosa* is less certain. The authors suggest that this could be related to its deciduous nature which might cloud the true relationships. It seems that further work is necessary to resolve this.

A strict consensus tree of the two equally parsimonious cladograms derived from data on both extant and extinct species is shown in Fig. 6.11. The study showed that in this analysis *E. glutinosa* rather than *E. cordifolia* is in a basal position to the rest of the genus. Of the species from Australia, *E. moorei* is the least derived.

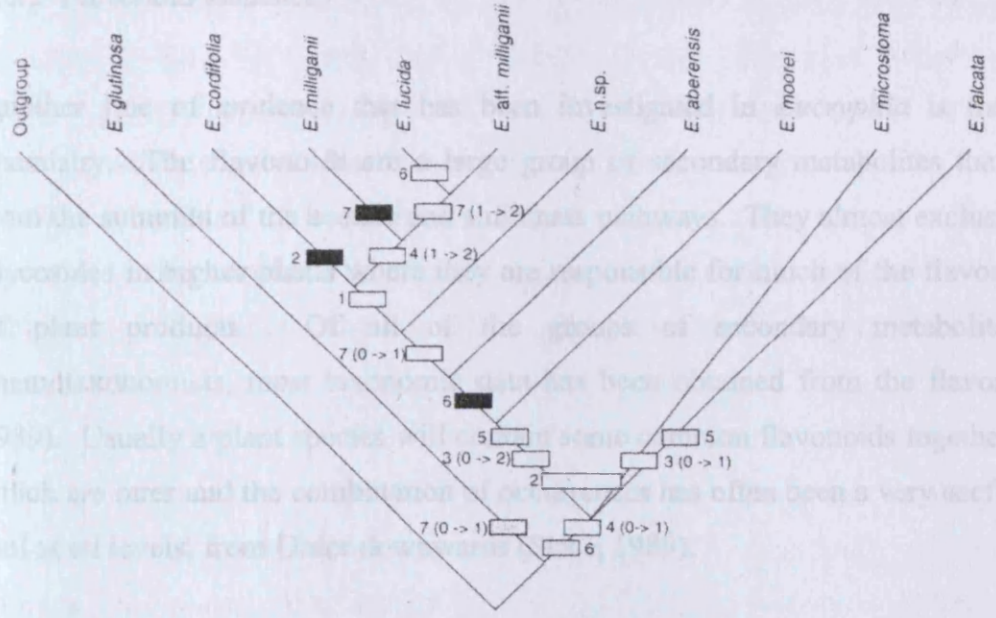

Fig. 6.11. Cladogram of living and fossil *Eucryphia* species based on morphologcal data. *E.* n. sp. = *E. wilkiei*. *E. jinksii* was unknown at the time. From Taylor and Hill (1996).

The two studies differ markedly in their treatment of the two South American species, and Barnes (pers. comm.) has questioned the validity of some of the morphological characters used in the survey. He performed a second analysis, which included *E. jinksii,* but arrived at a similar result, the only difference being that *E. willkiei* and *E. moorei* were transposed in his new tree.

Other morphological work that warrants discussion is that by Forster and Hyland (1996) who described and commented on *E. jinksii*, the species not included in the Taylor and Hill (1996) study. On morphological evidence, *E. jinksii* is thought to be more closely related to *E. wilkiei* than to the other Australian species based on similarities in the foliage. The biggest difference between the two species is that *E. wilkiei* is a shrub that grows only to 3-4 m high whereas *E. jinksii* is a large tree which grows to 25 m high (*E. wilkiei* lives in much harsher conditions and at high altitude compared to *E. jinksii* which is one of the main reasons of height differences). *E. wilkiei* also has smaller stipules (4-5.5 mm long compared to 12-15 mm long), shorter petioles (0.3-5 mm long as opposed to 5-15 mm long), slightly recurved leaf margins compared with the flat margins of *E. jinksii*, and has fewer pairs of lateral veins in the leaflets (12-20 instead of 22-30); (Forster and Hyland, 1996).

6.6.2 Flavonoid evidence

Another line of evidence that has been investigated in *Eucryphia* is their flavonoid chemistry. The flavonoids are a large group of secondary metabolites that are derived from the subunits of the acetate and shikimate pathways. They almost exclusively exist as glycosides in higher plants where they are responsible for much of the flavour and colour of plant products. Of all of the groups of secondary metabolites used by chemotaxonomists, most taxonomic data has been obtained from the flavonoids (Stace, 1989). Usually a plant species will contain some common flavonoids together with others which are rarer and the combination of occurrences has often been a very useful taxonomic tool at all levels, from Order downwards (Stace, 1989).

The flavonoid chemistry of *Eucryphia* tissue was first studied by Bate-Smith *et al.* (1967). They studied the five species of *Eucryphia* which were known at the time: *E. cordifolia*, *E. glutinosa*, *E. lucida*, *E. milliganii* and *E. moorei.* They found seven flavonol glycosides

and one dihydroflavonol glycoside. They also found two flavonol methyl ethers as aglycones in these five species. In a subsequent study, Wollenweber *et al.* (2000) analysed the flavonoids that are found in the exudates of the leaves and buds of all seven *Eucryphia* species. They identified 28 compounds in all (Table 6.3), although no attempt was made to look at infra-specific variation: only one accession was used for each species.

Of these compounds, the South American species, *E. glutinosa,* has two flavones, one of which it shares with two Australian and one Tasmanian species. The other Chilean species, *E. cordifolia,* contains only a trace amount of one flavone; neither South American species contains flavonols.

The two exudate-rich Tasmanian species (*E. lucida* and *E. milliganii*) have many flavonoids and share seven that are not recorded for any other taxon of *Eucryphia.* Four further flavonoids are unique to *E. lucida* and two others are produced only by *E. milliganii.* The two species from South Australia (*E. jinksii* and *E. moorei*) also have many diverse flavonoids, four of which are present only in *E. jinksii* and one is unique to *E. moorei.* The only species capable of hydroxylating the 8-position is *E. jinksii.*

The South Australian and the Tasmanian species are able to methylate the same hydroxyl groups of the flavonoids. Interestingly, the northern Australian species *E. wilkiei* is the only species which contains flavonoid glycosides in its exudates. Finally the two Australian species, *E. wilkiei* and *E. jinksii,* and the Tasmanian *E. lucida* all share an unidentified flavonoid aglycone.

Table 6.3. Flavonoids found in *Eucryphia* species; Ome = methyl ether moiety, + = present, (+) = trace amount present. Modified from Wollenweber *et al.* (2000). Blank spaces indicate absence of a compound.

Eucryphia species / Flavonoids	*cordifolia* Chile	*glutinosa* Chile	*lucida* Tasm.	*Milliganii* Tasm.	*Jinksii* Aust.	*Moorei* Aust.	*Wilkiei* Aust.
Apigenin (Ap)			+		(+)	(+)	
Ap-7-Ome			+	+		+	
Ap-4'-Ome					+	+	
Ap-7,4'-Ome	(+)	+	+	+	(+)	+	
Isoscutellarein-8,4'–Ome					+		
Isoscut-7,8,4'–Ome					(+)		
Luteolin (Lu)			+		+		
Lu-7-Ome			+		+		
Lu-3'-Ome			+				
Lu- 7,3'-Ome		+	+	+	+		
Lu- 7,4'-Ome				+			
Lu- 7,3',4'-Ome			+	+			
Lu-6,7-Ome						(+)	
Kaempferol-3-Ome			+	+			
K-4'-methyl ether					+		
K-3,7-Ome			+				
K-3,4'-Ome			+				
K-3,7,4'-Ome			+	+			
Quercetin-3-Ome			+				
Q-3,7-Ome				+		(+)	
Q-3,3'-Ome			+	(+)			
Q-3,4'-Ome				+			
Q-7,3'-Ome					+		
Q-3,7,4'-Ome			+	+			
Q-3,3',4'-Ome			+	+			
Q-3,7,3',4'-Ome			(+)	+			
Naringenin-7-Ome						+	
Unknown flavonoid aglycone			+		+		(+)
Flavonoid glycosides							+

After a phenetic analysis (Fig. 6.12), Wollenweber *et al.* (2000) concluded that: a) *E. wilkiei* does not appear to be closely related to any other taxon; and b) with the exception of *E. wilkiei,* the remaining taxa cluster along geographical lines. Thus there is a South American group and an Australian group, and the latter is further subdivided into a Tasmanian group and a mainland Australian group. A cladistic analysis of the same data (Wollenweber *et al.*, 2000) yielded six equally parsimonious trees, of which the strict consensus was "largely uninformative, except that it indicated that *E. lucida* and *E. milliganii* were sister taxa".

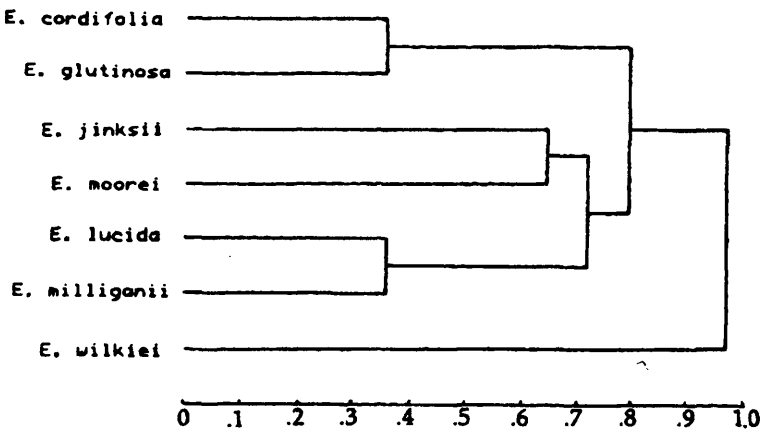

Fig. 6.12 UPGMA phenogram of *Eucryphia* species based on their flavonoid profiles. From Wollenweber *et al.* (2000).

This flavonoid study, therefore, to a large extent, supports the conclusions of the morphological study of living species by Taylor and Hill (1996), in particular in highlighting the separation of the South American from the Australasian members of the genus.

6.7 Molecular analyses: materials and methods

6.7.1 Plant material

Table 6.4 lists the collection details of the species used in the study. Leaves were collected from living specimens and dried in bags of silica gel prior to DNA extraction.

Table 6.4. Collection details of accessions of *Eucryphia* used in the present study. Voucher specimens from the Royal Botanical Garden Edinburgh material are lodged therein. Voucher specimens sent from Colin Forde at Seaforde Gardens are in his private herbarium. Voucher specimens for material collected Richard Barnes is in his private herbarium.

E. cordifolia
ex Royal Botanic Garden Edinburgh, accession, 19839125.
ex Seaforde Gardens, Co. Down.
Chile: Región VIII, Dpto Talcahueno, Concepcion, Parque Hualpén; 36° 48´S, 73° 12´W; elev. 40 m. 9/9/1999, *M.R.J. Clokie.*
Chile: Región X, Puyehue National Park, Casilla 1337, Osorno, 14/9/1999, *M.R.J. Clokie.*

E. glutinosa
Chile: Región VIII [Biobío], *Gardner and Knees S 4666*; (ex hort. Royal Botanic Garden Edinburgh, accession, 19900558).
ex Seaforde Gardens, Co.Down.

E. jinksii
ex hort. Canberra, accession 9613504.

E. lucida
Tasmania, *Edinburgh New Zealand and Australia Expedition 881,* (ex hort. Royal Botanic Garden Edinburgh, accession, 19980982.
ex R. Barnes

E. milliganii subsp. *milliganii*
Clear Hill Road, Lake Pedder, south-west Tasmania), R. Barnes
Serpentine Dam, Lake Pedder, south-west Tasmania), R. Barnes

E. milliganii subsp. *pubescens*
Mt. Eliza, south-west Tasmania) - R. Barnes
Moonlight Flats, near Mt La Perouse, southern Tasmania, R.Barnes

E. moorei
ex Royal Botanic Garden Edinburgh, accession, 19921837.
Ex R. Barnes

E. wilkiei
ex Seaforde Gardens
ex R. Barnes

Bauera rubioides
ex R. Barnes

Bauera sessiliflora
Grampians, Victoria, R. Barnes

Ceratopetalum apetalum
ex R. Barnes

**Ceratopetalum gummiferum*
ex R. Barnes

**Gillbeea adenopetala*
ex R. Barnes

·*Geissois benthamiana*
ex R. Barnes

·*Schizomeria ovata*
ex R. Barnes

*Sequenced for *rpo*A and *trn*L – F only
· Sequenced for *rpo*A only

6.7.2 Methods

DNA sequences from six gene regions were obtained following the methods described in Chapter 4. Three of the sequences were from the chloroplast (*rpo*A, *trn*L-F and *trn*H-K) and three were from the nuclear genome (ITS-1, ITS-2, *G3pdh*-1 and *G3pdh*-2). Sequences were aligned using Megalign, then checked by eye. Full details are available on request.

6.7.3 Analyses

Phylogenies were derived from the *rpo*A, *trn*H-K, *trn*L-F, ITS and *G3pdh*-1 and *G3pdh*-2 sequences using the branch and bound parsimony method described in Chapter 4. The outgroups were selected on the basis of results from analyses of both morphological and *rbc*L data (Hufford and Dickison, 1992; Morgan and Soltis, 1993). The choice was narrowed accordingly to *Bauera rubioides*, *Bauera sessiliflora*, *Ceratopetalum apetalum*, *Ceratopetalum gummiferum*, *Gillbeea adenopetala*, *Geissois benthamaniana* and *Schizomeria ovata*. Of these only *Ceratopetalum apetalum* and *Bauera sessiliflora* could be amplified and sequenced for all six genes. However, *C. apetalum* in all cases but one

(*G3pdh* locus 2, where *B. sessiliflora* was closer) proved to be one of the closest taxa to *Eucryphia* (Figs 6.13 and 6.14). This was on the basis of sequences aligned by Megalign and constructed in PAUP as described in the Methods chapter. When the number of mutations specific each to each taxa are counted *Ceratopetalum* has the fewest number so consequently it was chosen as the outgroup.

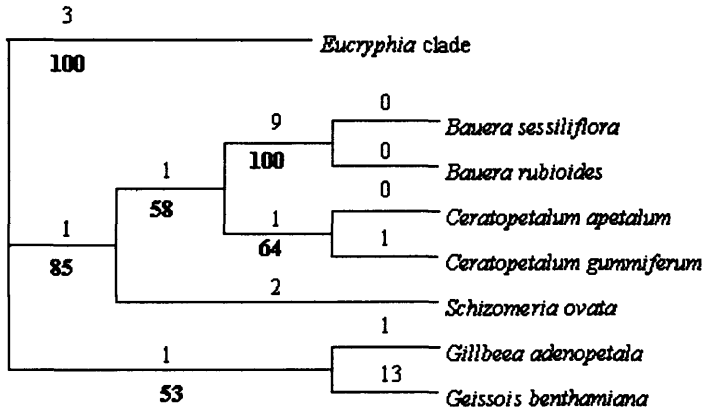

Fig. 6.13. Cladogram of *Eucyphia* and potential outgroup taxa based on *rpo*A sequences. Numbers of mutations are given above the clades, and bootstrap values are below in bold.

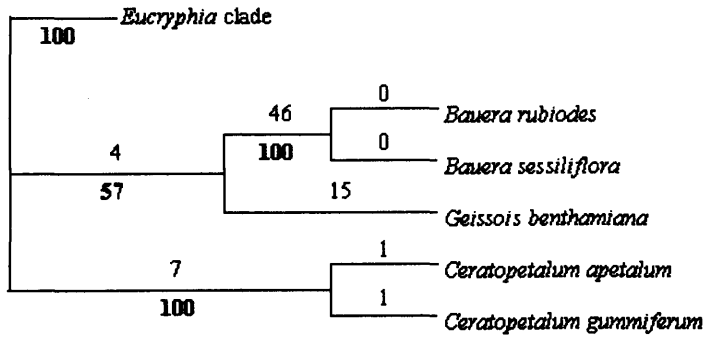

Fig. 6.14. Cladogram of *Eucryphia* and potential outgroup taxa based on *trn*L-F sequences. Numbers of mutations are given above the clades, and bootstrap values are below in bold.

6.7.4 Phylogenies

The individual phylogenies are given and discussed for all genes starting with the chloroplast sequences *rpo*A, *trn*L –F, and *trn*H-K and followed by the nuclear sequences ITS and *G3pdh*. Combined trees for all chloroplast genes, all nuclear genes and all genes together are also given. In all cases, where only one most parsimonious tree was found,

this is the tree given in the results, complete with bootstrap values and number of mutations occurring on each branch. Where more than one tree was found, the strict consensus tree with bootstrap values is shown. An example of a most parsimonious tree is also given because this gives an indication of the number of mutations present on each branch.

6.7.5 Tables of statistics relating to the data sets

The properties of the sequences are given in tables, following the phylogenetic tree for each individual gene. It is important to make the point that these data refer simply to the variation and properties of the *Eucryphia* data set and do not include the outgroup in the calculations.

Information is given for the range and mean length of sequence. The number of insertions and deletions is also given along with their size and whether or not they are informative. It is discussed in the text where and to which species these insertions or deletions occur. The number and percentage of constant, autapomorphic and informative sites is shown. This is further expanded on by showing how many of these point mutations are transitions and which constitute transversions. The average number of steps character is shown. The range of sequence divergence is also shown. Finally the GC content is given. This information also allows comparisons to be made regarding the inherent usefulness of different gene sequences.

6.7.6 Graphs to show positions of variability

As there are relatively few insertions or deletions in these data sets, they are discussed along with the other properties appertaining to the gene sequences. Much of the structure however within the phylogenetic trees is derived from point mutations, the distribution of which within each sequence is shown in a graph.

6.8 Molecular phylogenies

6.8.1 rpoA phylogeny

The *rpo*A sequences produced a single most parsimonious tree of length 9 (Fig. 6.15). There was no homoplasy (CI = 1, RI = 1, RC = 1, HI = 0). Although clearly showing that *Eucryphia* is monophyletic, unfortunately there was very little other resolution apart from a clade containing the two Tasmanian species, *E. lucida* and *E. milliganii*, and a larger clade containing all species except for *E. cordifolia* and *E. wilkiei*. Both of these clades however had moderate bootstrap support of 62% and 63% respectively. There was no intraspecific variation.

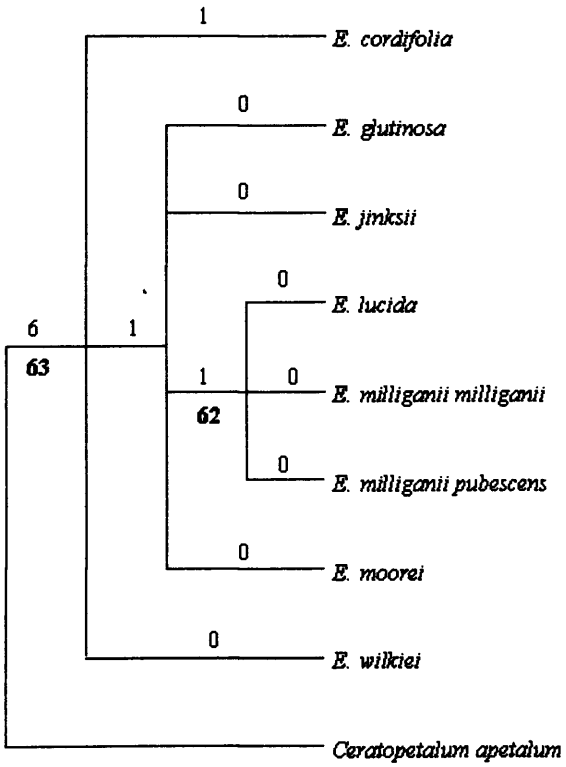

Fig. 6.15. Phylogeny for *rpo*A in *Eucryphia*. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

6.8.2 rpoA variation

Statistics relating to the properties of the *rpo*A data are given in Table 6.5. From this information it is clear why the *rpo*A tree is fairly unresolved. Of the three (0.7%) variable bases two (0.44%) are informative and one (0.22%) is autapomorphic. The low rate of change is reflected in the low average number of steps per character (0.007). All three point mutations are transitions, which is consistent with the fact that this region of DNA is in an exon. There are also no insertions or deletions. There is no sequence divergence between some of the species of *Eucryphia* for this gene and the maximum divergence is a fairly low 0.67% (Gillham, 1994). The GC content (37.1%) is typical for a chloroplast gene (Gillham, 1994).

Table 6.5. Statistics relating to *rpo*A sequence data in *Eucryphia*

Length range (bp)	450
Length mean (bp)	450
Number of characters (aligned length + indels)	450
Number of indels	0
Size of indels (bp)	N/A
G+C content mean %	37.1
Sequence divergence %	0 – 0.67
Number of constant sites (%)	447 (99.3)
Number of autapomorphic sites (%)	1 (0.22)
Number of informative sites (%)	2 (0.44)
Transitions	3
Tranversions	0
Percentage transitions	100
Average number of steps per character	0.007

The pattern of variation is illustrated in Fig. 6.16, where it can be seen that the three variable bases change once each in the ingroup to produce the most parsimonious tree. The changes occur at discrete- intervals along the sequence and there is no cluster of variation. Although there is twice as much variation within the *Eucryphia* and *Ceratopetalum* data set, the overall level is still low. There is little correlation in the pattern of variation in the two plots.

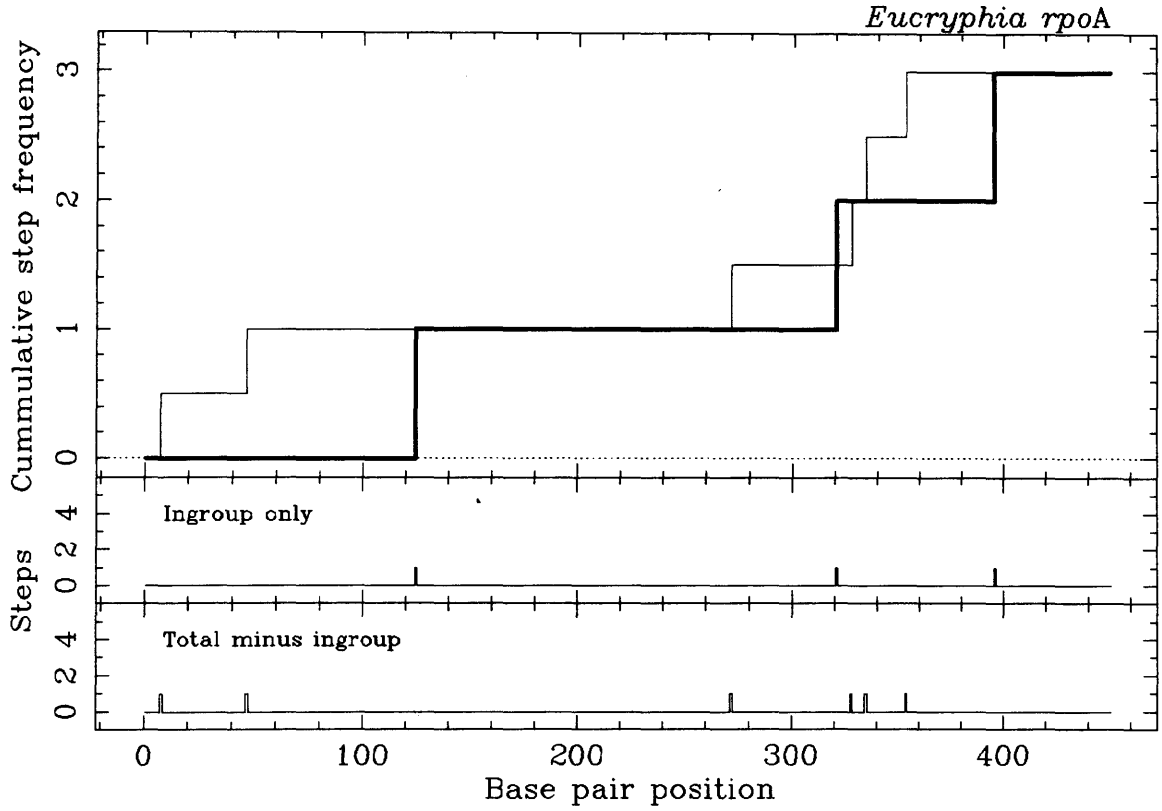

Fig. 6.16. Graph to illustrate the behaviour of *rpo*A in *Eucryphia* and its outgroup *Ceratopetalum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Ceratopetalum*) from the ingroup (*Eucryphia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

6.8.3 TrnL–F phylogeny

The *trn*L-F sequences produced a single most parsimonious tree of length 26 (Fig. 6.17). There was very little homoplasy (CI = 0.962, RI = 0.923, RC = 0.868, HI = 0.038). *Eucryphia* is clearly monophyletic. Within *Eucryphia* there are three well supported clades. The two Chilean species form a clade which is supported by a high bootstrap value of 81%. This indicates the presence of a single trans-Antarctic disjunction within *Eucryphia*. Although *E. wilkiei* is included as being basal to these species, the bootstrap support is relatively low, 57%. There is a clade that contains the two Tasmanian species *E. lucida* and *E. milliganii* (with good bootstrap support of 76%). There is also a clade consisting of the two accessions of *E. milliganii* ssp. *milliganii*, with one mutation holding them together. *E. milliganii* ssp. *pubescens* differs from ssp. *milliganii* by one or two base pairs. The Australian species *E. jinksii* and *E. moorei* form a clade which is moderately supported with a bootstrap value of 66%.

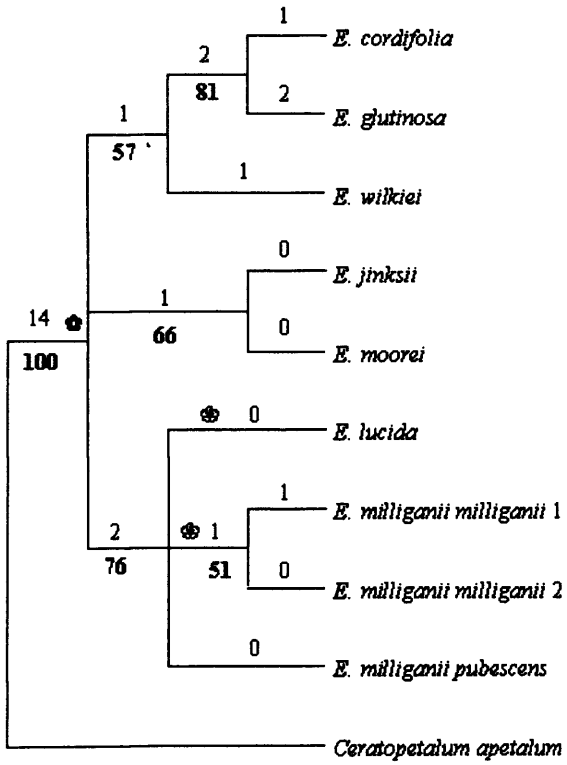

Fig. 6.17. Phylogeny for *trn*L-F. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold. ❀ designates the position on the tree of the 2.6 MYO fossils and ✿ designates the position of the 60 MYO fossil.

6.8.4 TrnL-F variation

Statistics relating to the properties of the *trn*L-F data are given in Table 6.6. In this data set 10 (1.08%) of the bases are variable, and of these more than half (0.65%) are informative. The average number of steps per character is a moderate 0.013. There is also one insertion unique to *E. cordifolia*, of two base pairs. Interestingly, even though this region of DNA comprises an intron and an intergenic spacer region, all of the mutations are transitions. The sequence divergence is as low as 0 (there are no differences in this region between *E. jinksii* and *E. moorei*) and the highest value is 0.76%. The GC content is fairly low but again at 0.345 it is typical for the chloroplast genome.

Table 6.6. Statistics relating to *trn*L-F sequence data

Length range (bp)	921 – 923
Length mean (bp)	921.2
Number of characters (aligned length + indels)	924
Number of indels (informative)	1
Size of indels (bp) [uninformative]	[2]
G+C content mean %	34.5
Sequence divergence %	0 – 0.76
Number of constant sites (%)	914 (98.92)
Number of autapomorphic sites (%)	4 (0.43)
Number of informative sites (%)	6 (0.65)
Transitions	10
Transversions	0
Percentage transversions	100
Average number of steps per character	0.013

The pattern of variation is illustrated in Fig. 6.18. When the ingroup alone is considered only 10 of their positions change in the entire sequence. The last peak which is observed, as has been previously discussed, represents the coding for an insertion. This actually occurs in *E. cordifolia* at position 367 – 369 bp. Of the point mutations one changes twice within the ingroup data and the rest once. In the data set that includes the outgroup *Ceratopetalum*, there are two main regions of variability, the first is between positions 200 and 500 bp and the second between positions 800 and 900 bp. The intron sequence ranges from 1 – 606 bp and the IGS spans from 607 – 924 bp. As can be seen therefore, variation within *Eucryphia* itself is divided more or less equally between the intron and the IGS. There is also a fairly good correlation between the areas which are variable within the ingroup, and those which vary with the outgroup species.

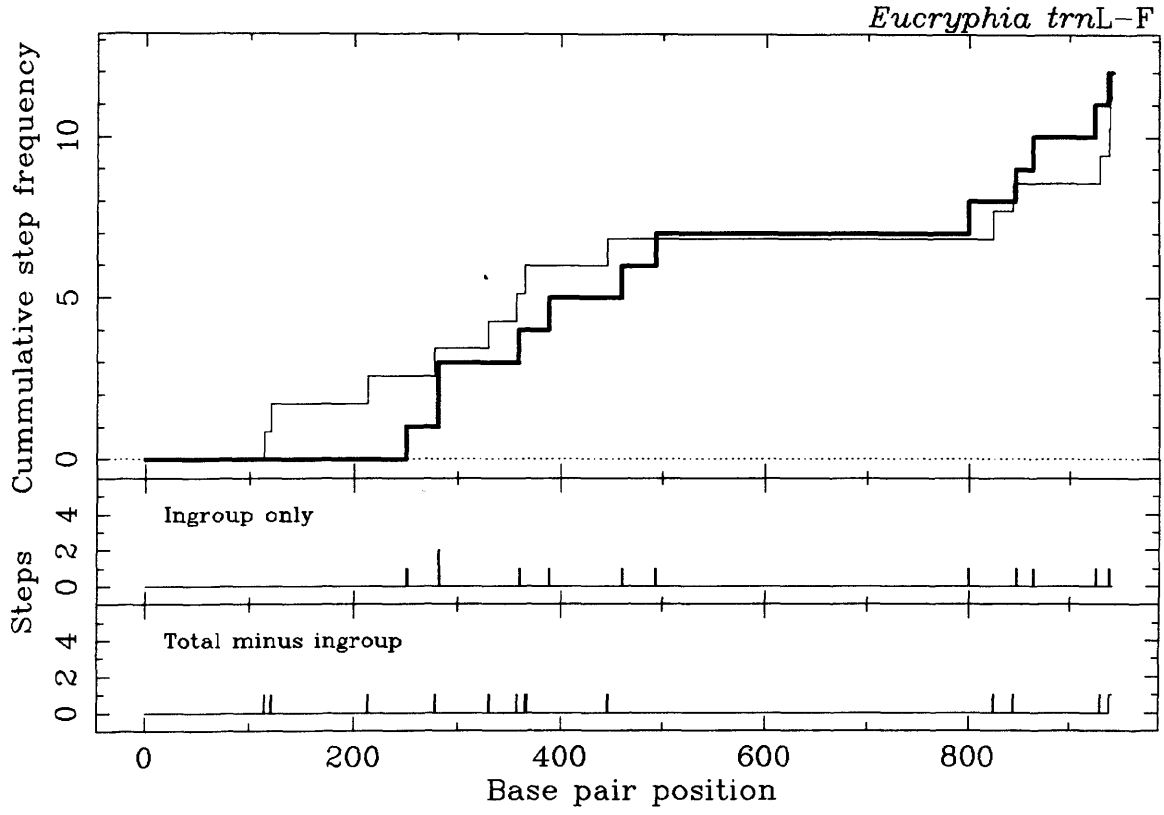

Fig. 6.18. Graph to illustrate the behaviour of *trn*L-F in *Eucryphia* and its outgroup *Ceratopetalum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Ceratopetalum*) from the ingroup (*Eucryphia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

6.8.5 TrnH-K phylogeny

The first thing to say about the *trn*H – K data set is that it consists on average of 657 bp of sequence data which was sequenced in a forwards direction, and on average 450 bp of sequence data which was obtained by sequencing in the reverse direction. Unfortunately in *Eucryphia* and in the outgroup species the sequences are not contiguous but because they are part of the same region, they have been grouped together and analysed as one entity.

The *trn*H-K sequences produced a single most parsimonious tree of length 46 (Fig. 6.19). There was no homoplasy (CI = 1, RI= 1, RC = 1, HC = 0). Although the tree is largely unresolved, it does contain a clade composed of the two South American species, *E. cordifolia* and *E. glutinosa*, with 64% bootstrap support. A single trans-Antarctic disjunction is therefore indicated. The mutation that unites the species is a 7bp inserted repeat sequence. This phylogeny also suggests that *E. moorei* is basal to the rest of the genus.

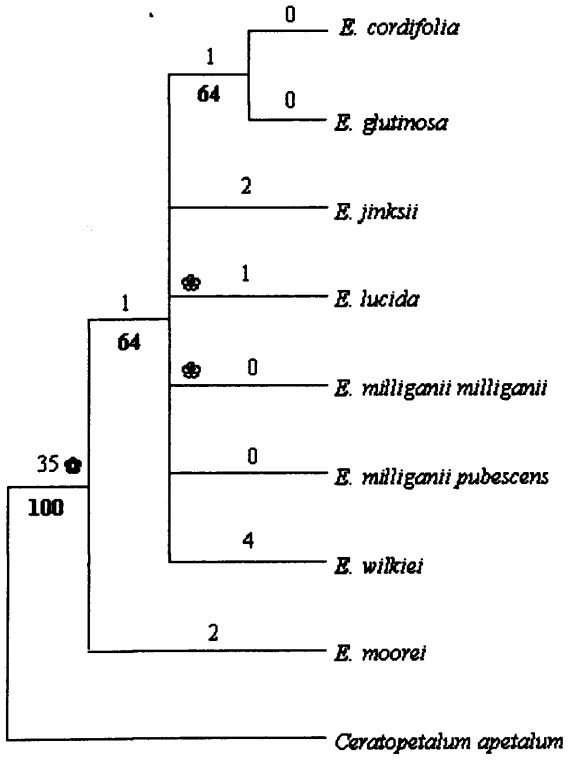

Fig. 6.19. Phylogeny for *trn*H-K. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold. ❀ designates the position on the tree of the 2.6 MYO fossils and ✿ designates the position of the 60 MYO fossil.

6.8.6 TrnH-K variation

Statistics relating to the properties of the *trn*H-K data are given in Table 6.7. From this table, it is clear why the *trn*H-K tree is fairly unresolved. Of the 9 (0.74%) variable bases, none is shared between more than one species. Again there is a very low average number of steps per character (0.009). Of these autapomorphic point mutations, two-thirds are transitions and a third transversions. The sequence divergence ranges from 0 to 0.0305. The 7 base pair insertion mutation that unites the two Chilean species occurs at 117 – 123 and the non-informative mutation unique to *E. wilkiei* is found at 1195 bp. The GC content at 32.1%, again is a typically low value as expected for the chloroplast.

Table 6.7. Statistics relating to *trn*H-K sequence data

Length range (bp)	1209 – 1216
Length mean (bp)	1211.24
Number of characters (aligned length + indels)	1219
Number of indels (informative)	2 (1)
Size of indels (bp) [uninformative]	7 [1]
G+C content mean %	32.1
Sequence divergence %	0 – 0.305
Number of constant sites (%)	1202 (99.26)
Number of autapomorphic sites (%)	9 (0.74)
Number of informative sites (%)	0 (0)
Transitions	6
Transitions	3
Percentage transversions	66.66
Average number of steps per character	0.009

The pattern of variation is illustrated in Fig. 6.20. The *trn*H-K graph has only 9 variable positions. Remembering that this region is actually the two ends of *trn*H-K, the variable bases generally are near the 3' ends of the two sections. The two insertions again, are coded for at the end of the base pair chart, and their positions mentioned above. When the outgroup *Ceratopetalum* is compared the variation occurs in similar areas to where it occurs within *Eucryphia*. The cumulative plot shows the fairly good correlation between the two data plots.

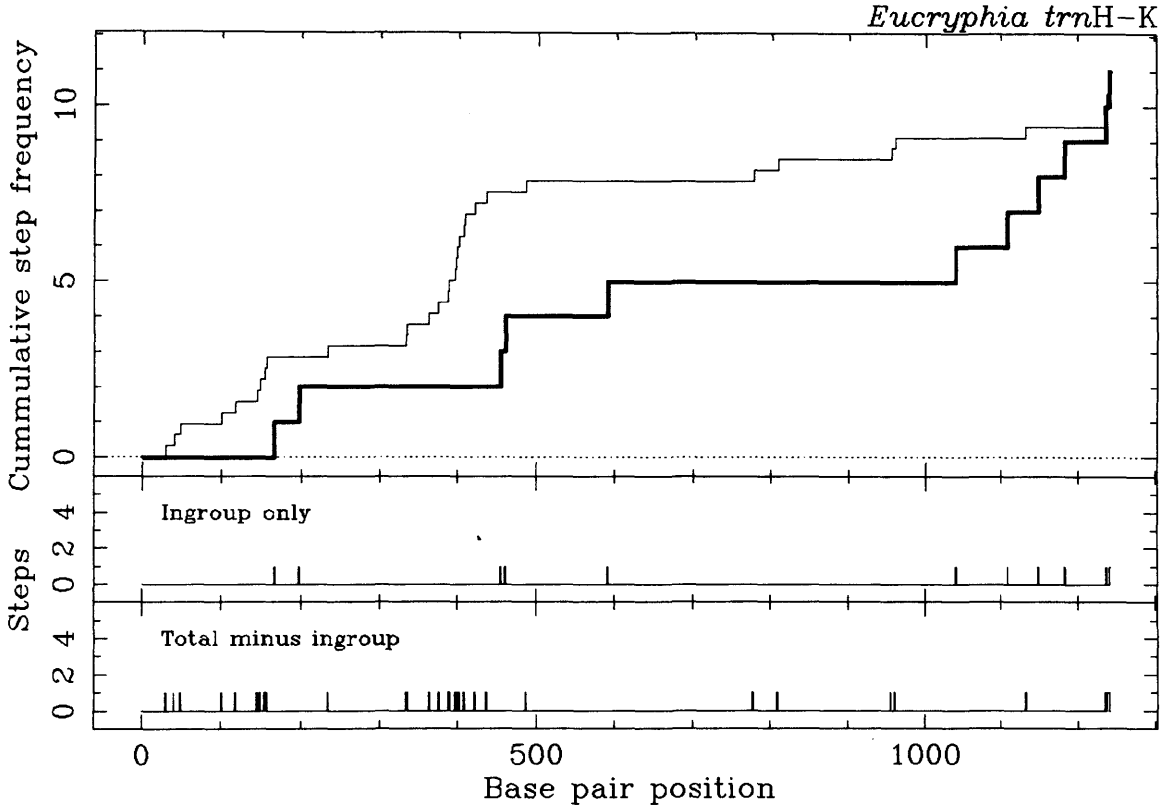

Fig. 6.20. Graph to illustrate the behaviour of *trn*H-K in *Eucryphia* and its outgroup *Ceratopetalum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Ceratopetalum*) from the ingroup (*Eucryphia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

6.8.7 Combined rpoA, trnL –F and trn H-K phylogeny

A phylogeny based on all the chloroplast sequences was obtained by analysing *rpo*A, *trn*L-F and *trn*H-K as a single data matrix. A single most parsimonious tree of length 83 was obtained with very little homoplasy (CI = 0.964, RI = 0.85, RC = 0.82, HC = 0.036). The overall topology of the tree (Fig. 6.21) is very similar to that for the *trn*L-F tree but with improved bootstrap values on some of the groupings. The bootstrap support for the Chilean species is now convincingly high (87%). Similarly, the support for the Tasmanian group is now high (88%). *E. wilkiei* is associated with the Chilean species and *E. jinksii* and *E. moorei* are associated with each other but only with moderate levels of bootstrap support (60% and 51% respectively). *E. milliganii* shows infraspecific variation such that it is not fully differentiated from *E. lucida.*

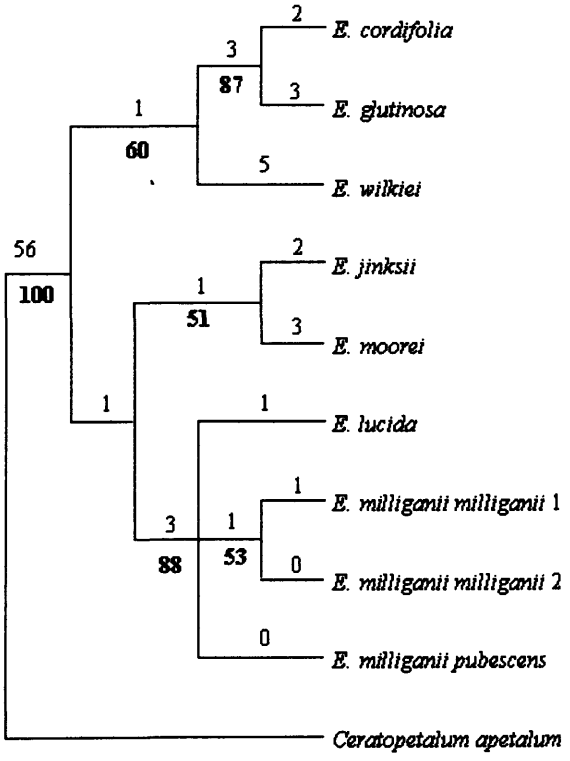

Fig. 6.21. Phylogeny for *rpo*A, *trn*L-F and *trn*H-K. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

6.8.8 ITS phylogeny

It was not possible to align *Eucryphia* sequences with the outgroup species *Ceratopetalum apetalum* (despite the 5.8S coding region). On the basis that the group containing *E. wilkiei* and *E. jinksii* is basal to the other species for other nuclear genes, *G3pdh* locus 1 and 2, these two species were used to route the ITS phylogeny. ITS sequence data produced one shortest tree of length 69 (Fig. 6.22). Homoplasy was extremely low (CI = 0.97, RI = 0.97, RC = 0.94, HI = 0.029). The tree shows firstly that the two South American species, *E. cordifolia* and *E. glutinosa*, form a well-supported clade (99% bootstrap value). A single trans-Antarctic disjunction is therefore indicated. Secondly, although slight intraspecific variation occurs within *E. milliganii*, the accessions form a monophyletic group with 100% bootstrap support. Thirdly, the Australian species *E. moorei* is sister to the clade containing the Chilean and the Tasmanian species. The other two Australian species *E. jinksii* and *E. wilkiei* are unresolved and sister to all of the other species.

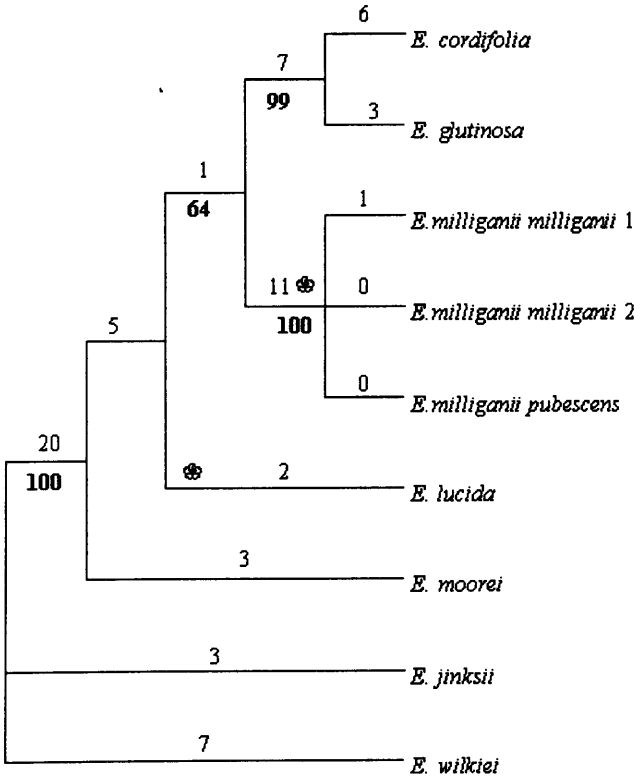

Fig. 6.22. Phylogeny for ITS. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold ❀ designates the two positions on the tree of the 2 MYO fossils.

6.8.9 ITS variation

Statistics relating to the properties of the ITS data are given in Table 6.8. In the ITS data set, 756 (93%) of the sites are constant, and of the remaining variable sites, 21 (2.6%) are autapomorphic while 36 (4.4%) are informative. This relatively high amount of variability is reflected in the average number of steps per character being 0.083. The transition to transversion ratio is 1.375. There is no sequence divergence between one accession of *E. milliganii milliganii* and *E. pubescens*, but aside from this the divergence ranges from 0.123% to 4.605%.

Of the eight insertions or deletions, four are informative, one of which is variable in length. The informative indels are as follows: a 4 bp deletion in *E. jinksii* and *E. wilkiei* at positions 198 – 201, a further 4 bp deletion in the same species at 241 – 244. The variable deletion is between 468 and 469 bp (in *E. jinksii* and *E. wilkiei*) or between 469 and 472 bp (in *E. moorei* and the two Tasmanian species). These two more northerly Australian species *E. jinksii* and *E. wilkie* share a final 1 bp deletion at 518 bp. The autapomorphic indels are as follows: two 1 bp deletions in *E. cordifolia* at 66 and 92bp, a 1 bp in *E. moorei* at 264, and a 3 bp insertion in *E. glutinosa* at 100 – 103. The GC content is a fairly high 59.9%.

Table 6.8. Statistics relating to ITS sequence data

Length range (bp)	804 – 820
Length mean (bp)	813
Number of characters (aligned length + indels)	825
Number of indels (informative)	8 (4)
Size of indels (bp) [uninformative]	1, 2-4, ,4,4 [3,1,1,1]
G+C content mean %	59.9
Sequence divergence %	0.123 – 4.605
Number of constant sites (%)	756 (93)
Number of autapomorphic sites (%)	21 (2.6)
Number of informative sites (%)	36 (4.4)
Transitions	33
Transversions	24
Percentage transitions	57.89
Average number of steps per character	0.083

The pattern of variation is illustrated in Fig. 6.23. This shows that only five of the 57 variable bases (8.8%) are found in the 5.8S rRNA gene. This gene spans 19% of the region, so by chance alone, we would expect to find 11 ± 3 variable bases within it. Furthermore, none of the eight insertions or deletions is found within this gene. There are three main clusters where variation occurs, the first is between 30 and 130 bp, the second is towards the end of the first spacer between around 200 and 260 bp. The final main cluster is at the beginning of the second spacer region, particularly between bases 440 – 490 bp. This can be seen clearly in the cumulative plot.

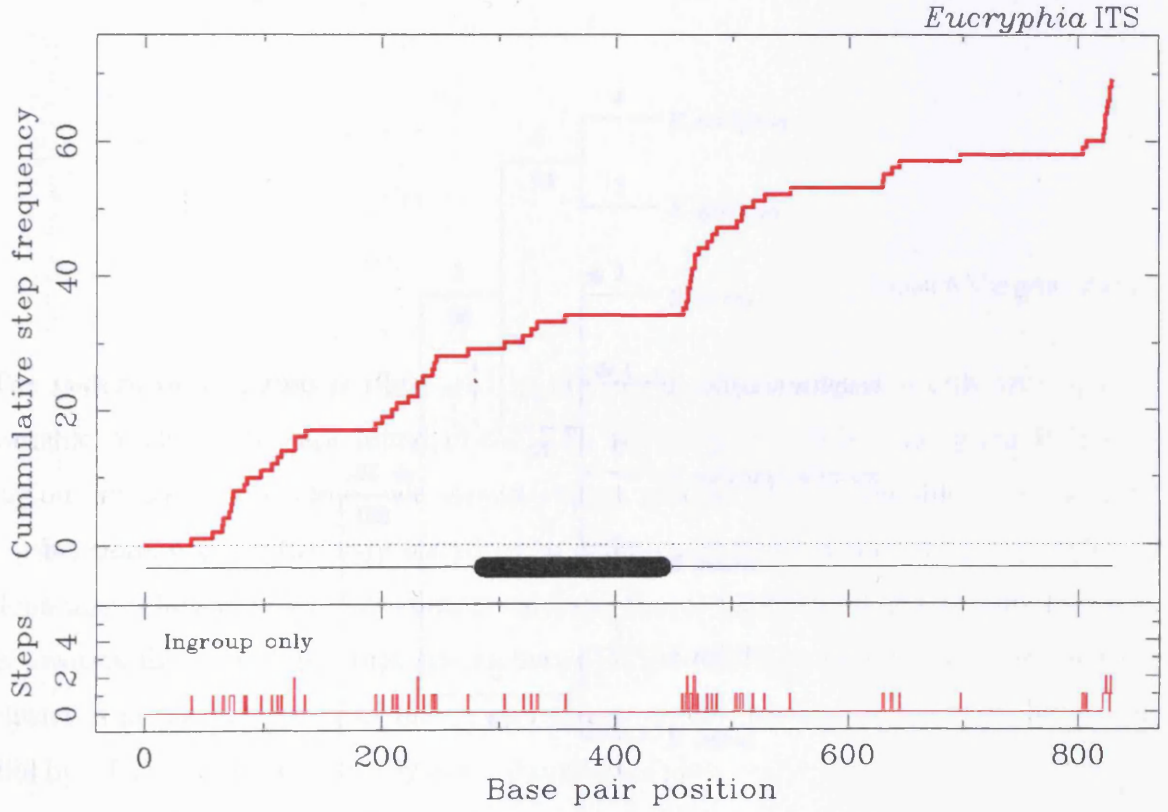

Fig. 6.23. Graph to illustrate the behaviour of the ITS in *Eucryphia*. ITS spacer 1 (fine black line) is from 1 – 287. The 5.8S rRNA gene (indicated by a thick black block) is from 288 – 439 bp. The spacer 2 is shown again as a thin line from 440 – 831 bp. The bottom panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these data.

6.8.10 G3pdh locus 1 phylogeny

The *G3pdh* locus 1 data set produced one shortest tree of length 140 (6.24). There was a small amount of homoplasy (CI = 0.97, RI = 0.79, RC = 0.77 and HC = 0.029). *Eucryphia* is clearly monophyletic with 100% bootstrap support for the clade. This tree shows that within *Eucryphia*, a major division is between the Australasian and the Chilean species (96% bootstrap support). Again this indicates a single trans-Antarctic disjunction. Both Tasmanian species are in a group with the southern Australian species *E. moorei*, this has 63% bootstrap support. The two more northerly Australian species *E. jinksii* and *E. wilkiei* are united by two mutations and a bootstrap value of 61%.

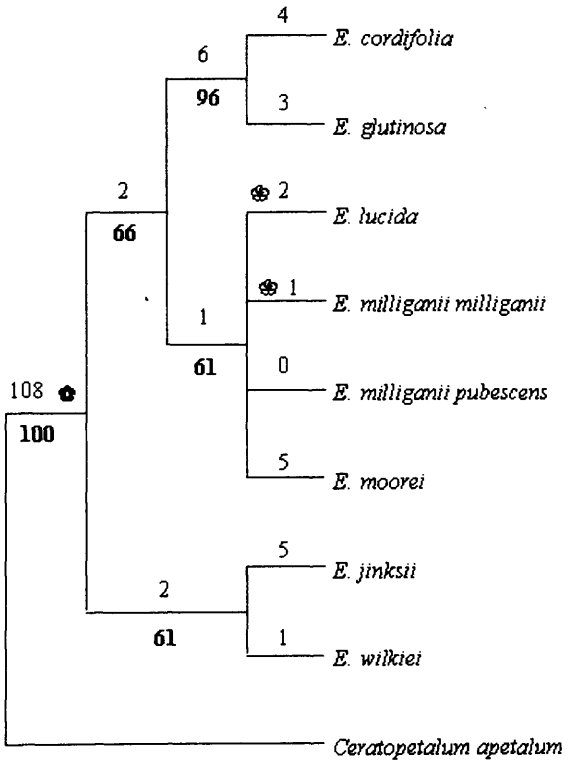

Fig. 6.24. Phylogeny for *G3pdh* locus 1. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold. ❀ designates the position on the tree of the 2.6 MYO fossils and ✿ designates the position of the 60 MYO fossil.

6.8.11 G3pdh locus 1 variation

Statistics relating to the properties of the *G3pdh* locus 1 data set are given in Table 6.9. In this data set 26 (2.7%) of the 987 bases are variable. Of these, 18 (1.83%) are autapomorphies and 8 (0.81%) are informative. This translates as an average of 0.052 steps per character. The transition to transversion ratio is 1:1. The sequence divergence is between 0.10 – 1.53%. Of the four insertions or deletions, two are informative (a 2 bp deletion between 40 and 42 bp for *E. glutinosa* and *E. cordifolia*, and a further 2 base pair deletion for the same species at position 223 – 225. The two autapomorphic indels are a 2 bp deletion for *E. jinksii* at position 223 and a 1 bp insertion in *E. glutinosa* at 593. The GC content is a fairly low 36.6%.

Table 6.9. Statistics relating to *g3pdh locus 1* sequence data

Length range (bp)	984 – 988
Length mean (bp)	986.88
Number of characters (aligned length + indels)	992
Number of indels (informative)	4 (2)
Size of indels (bp) [uninformative]	2,2, [2, 1]
G+C content mean %	36.6
Sequence divergence %	0.10 – 1.53
Number of constant sites (%)	961 (97.37)
Number of autapomorphic sites (%)	18 (1.83)
Number of informative sites (%)	8 (0.81)
Transitions	13
Transversions	13
Percentage transitions	50
Average number of steps per character	0.052

The pattern of variation is illustrated in Fig. 6.25. In both data sets, it is apparent that there is a strong correlation between the position of the introns, and the variable bases. This can be clearly seen by observing the flat gradient of the cumulative plots over the corresponding exons. Furthermore, all four insertions and deletions are found in the introns, the informative indels are found in intron A and B, and the autapomorphic ones in intron B and C.

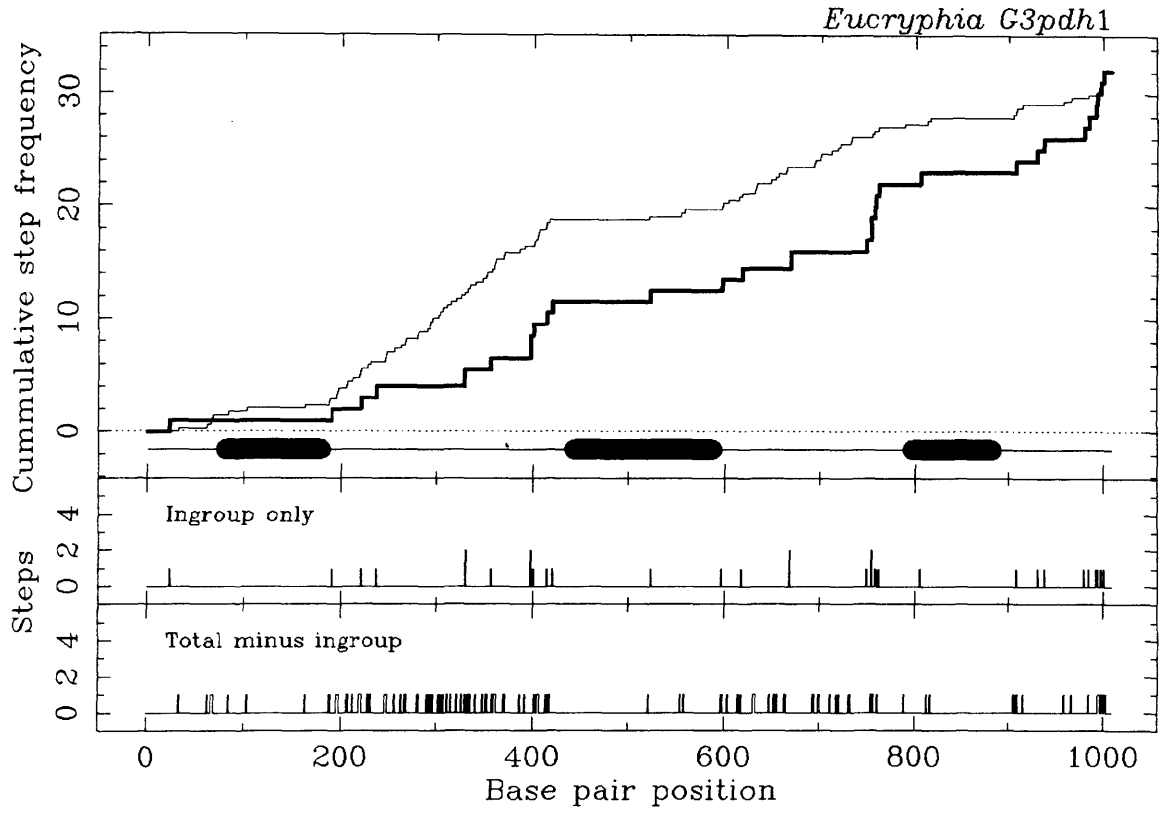

Fig. 6.25 Graph to illustrate the behaviour of *G3pdh* locus 1 in *Eucryphia.* Boundaries between introns and exons are shown between the normal and cumulative plots. The introns (A – D) are shown as fine black line and exons B – D are indicated by thick black blocks. The exact positions are: intron A (1 – 79bp), exon B (80 – 177 bp), intron B (178 – 441 bp), exon C (442 – 584 bp), intron C (585 – 794 bp), exon D (795 – 878 bp) and intron D (879 – 989 bp). The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Ceratopetalum*) from the ingroup (*Eucryphia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. These data produce four equally most parsimonious trees and include two positions which are invoked more than once in just one of these trees, such bases are plotted in blue. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

6.8.12 G3pdh locus 2 phylogeny

The *G3pdh* locus 2 data set produced six equally most parsimonious trees of length 117 (Fig. 6.25a and b). There is a small amount of homoplasy in the consensus tree (CI = 0.97, RI = 0.84, RC = 0.81 and HC = 0.034). *Eucryphia* is clearly monophyletic. The two Chilean species are held together by a bootstrap of 77%. This indicates a single trans-Antarctic translocation event. As is the case for the *G3pdh* locus 1, the Tasmanian species are placed in a group with the southern Australian *E. moorei* which are here supported by a bootstrap value of 94%. Although it is surprising that *E. milliganii* subsp. *milliganii* is more closely linked to *E. lucida* than to *E. milliganii* subsp. *pubescens*, the support holding them together is only 65%. In this tree, *E. jinksii* and *E. wilkiei* are not placed in any groups within *Eucryphia*. There are 2 bp of infra-specific variation found between the two accessions of *E. moorei*. A single base pair of infra-specific variation is also found within the two *E. wilkiei* accessions.

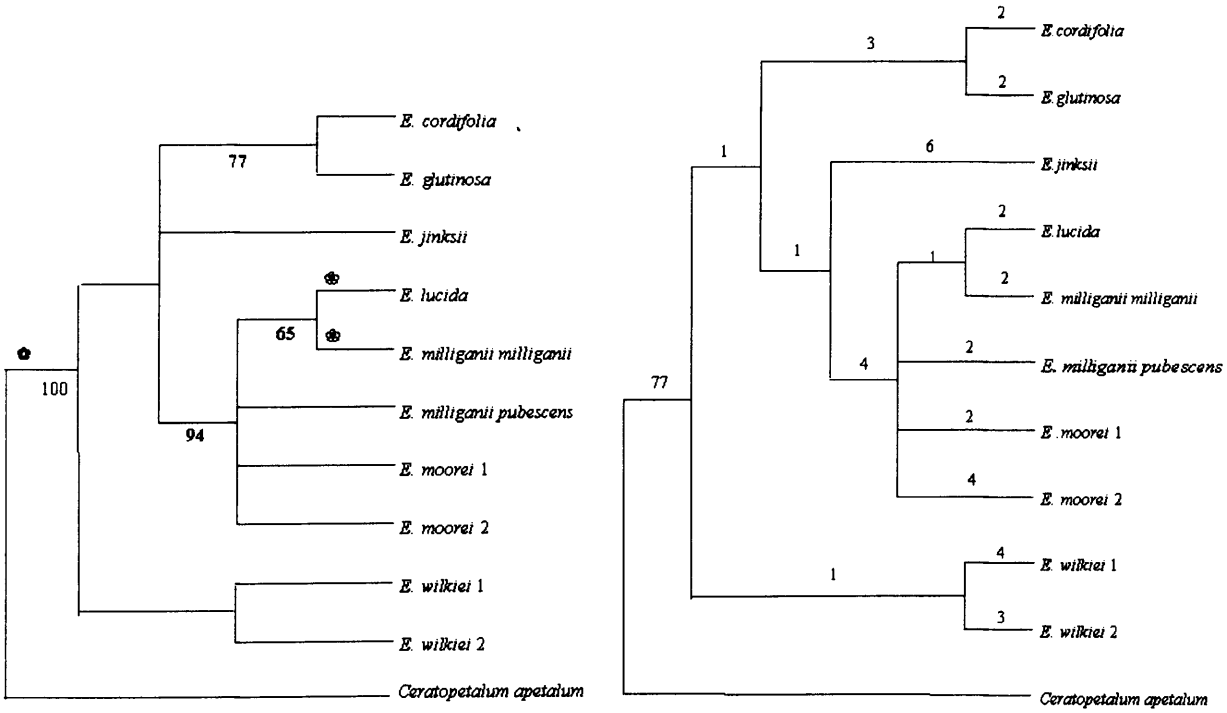

Fig. 6.25a. consensus tree for *G3pdh* locus 2 . **Fig. 6.25b.** example of shortest tree.

For the consensus tree (Fig. 6.25a) bootstrap values where they occur are below in bold and ❀ designates the position on the tree of the 2.6 MYO fossils and ✿ designates the position of the 60 MYO fossils. For the example of the shortest tree (Fig. 6.25b) the numbers of mutations are given above the clades.

6.8.13 G3pdh locus 2 phylogeny variation

Statistics relating to the properties of the *G3pdh* locus 2 data are given in Table 6.10. In this data set of the 37 (4.55 %) base pairs which are variable, 32 (3.93% are autapomorphic) and 5 (0.61%) are informative. Interestingly, 14 of these point mutations are transitions and 23 are transversions, giving a ratio of 0.61. The average number of steps per character is 0.047. The sequence divergence between species ranges from 0.049% and 1.718%. There are only two indels in this data set, both of which are uninformative. They are a 10 bp deletion at 582 – 592 in *E. glutinosa* and a 4 bp insertion in *E. jinksii* at 644 – 648. The GC content is 40.2%.

Table 6.10. Statistics relating to *G3pdh* 2 sequence data

Length range (bp)	805 – 819
Length mean (bp)	814.25
Number of characters (aligned length + indels)	821
Number of indels (informative)	2
Size of indels (bp) [uninformative]	[10,2]
G+C content mean %	40.2
Sequence divergence %	0.0491 – 1.718
Number of constant sites (%)	777 (95.45)
Number of autapomorphic sites (%)	32 (3.93)
Number of informative sites (%)	5 (0.61)
Transitions	14
Transversions	23
Percentage transitions	37.84
Average number of steps per character	0.047

The pattern of variation is illustrated in Fig. 6.26. These graphs show that as in the case for the *G3pdh* locus 1, there is a strong correlation between the position of introns and the variable bases. This may help understand the high proportion of transversions, as point mutations occur predominantly in the introns. Again, the two indels occur in introns, both in intron C. Although there are two equally most parsimonious trees for the ingroup-only data set, they both contain base changes in identical positions, thus only the topology of the trees differs slightly.

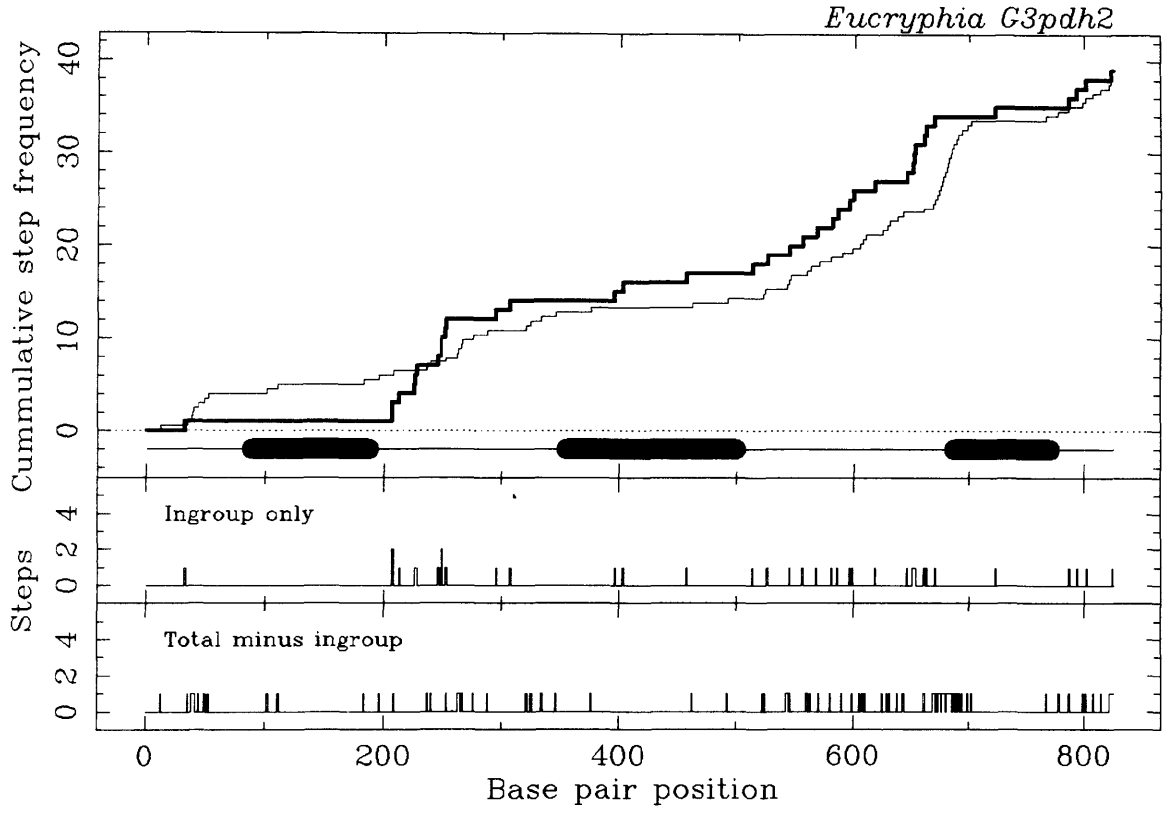

Fig. 6.26 Graph to illustrate the behaviour of *G3pdh* locus 2 in *Eucryphia.* Boundaries between introns and exons are shown between the normal and cumulative plots. The introns (A – D) are shown as fine black line and exons B – D are indicated by thick black blocks. The exact positions are: intron A (1 – 88bp), exon B (89 – 186 bp), intron B (187 – 352 bp), exon C (353 – 495 bp), intron C (496 – 684 bp), exon D (685 – 768 bp) and intron D (769 – 819 bp). The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Ceratopetalum*) from the ingroup (*Eucryphia*), plotted in green. These data produce four equally most parsimonious trees and include two positions which are invoked more than once in just one of these trees, such bases are plotted in blue. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

6.8.14 Combined phylogeny for G3pdh loci 1 and 2

The combined *G3pdh* locus 1 and locus 2 data set produced one most parsimonious tree with 258 steps (Fig. 6.27). There is a small amount of homoplasy (CI = 0.97, RI = 0.87, RC = 0.81, HC = 0.035). This tree shows 100% support for the monophyly of *Ecuryphia.* There is now 98% bootstrap support for the group containing the two Chilean species, therefore convincingly suggesting a single trans-Antarctic disjunction. The group containing the Tasmanian species (*E. lucida* and *E. milliganii*) and the Australian species (*E. moorei*) is now supported by a bootstrap value of 95%. Within this group there is 100% bootstrap support for the two different accessions of *E. moorei.* Although *E. milliganii* ssp. *milliganii* is in a clade with *E. lucida*, the bootstrap support is a fairly low 62%. Similarly the positioning of *E. jinksii* as a sister species to *E. wilkiei* has a low bootstrap support of 54%. The two different *E. wilkiei* accessions have 81% bootstrap support.

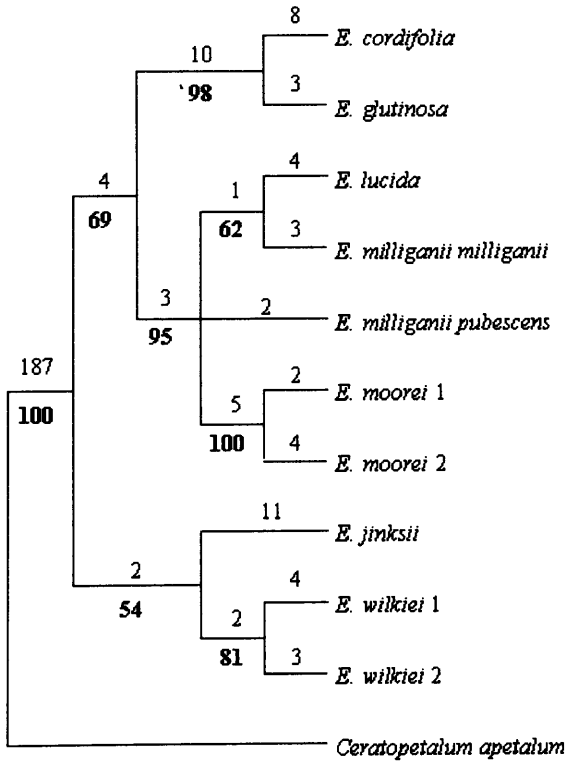

Fig. 6.27. Phylogeny for *G3pdh* locus 1 and 2 combined. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

6.8.15 Comparison of locus 1 and locus 2

In order to establish the position of introns and exons in the *G3pdh* sequence, ideally the individual mRNAs would be sequenced. Fortunately this work has been done for cassava and therefore by translating the *G3pdh* nucleotide sequences in *Eucryphia*, and comparing them to the amino acid sequence for cassava it was possible to work out the position of introns and exons in both of the *Eucryphia G3pdh* loci. The primers were designed to amplify the region from the right-hand side of exon A to the very beginning of exon E (Fig. 6.27).

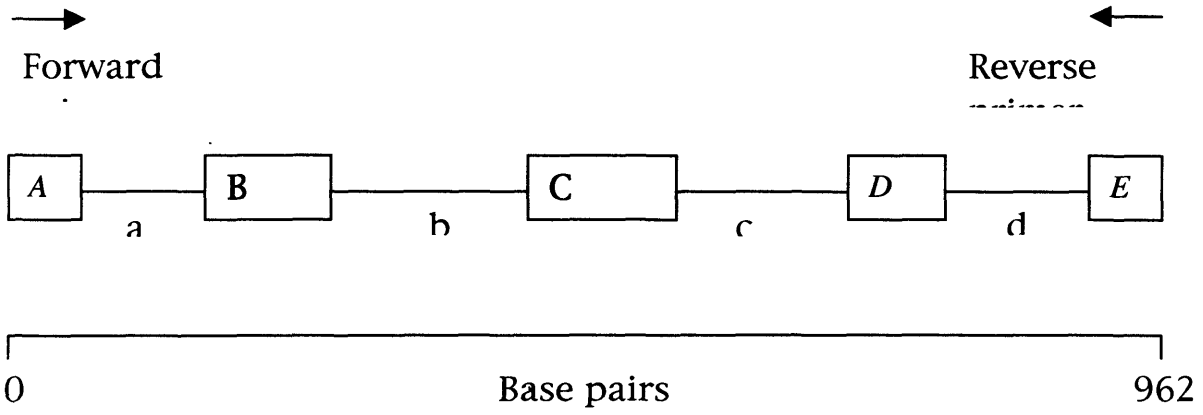

Fig. 6.28. Structure of the *G3pdh* region in *Manihot*. Uppercase and lower case designate exons and introns respectively, Redrawn from Olsen and Schaal, (1999).

The sequences for locus 1 and locus 2 were aligned for all species. A number of interesting observations can be made. In both loci three complete exons are present. The sizes of the introns and exons for the two loci are given in Table 6.11.

A study of the two *G3pdh* loci yields a number of interesting observations. To start with intra-locus variation, in locus 1 there is no variation within exon B in any of the species of *Eucryphia*. This is not the case for the other two exons (C and D) which display 1 and 2 base pairs of variation respectively. In locus 2, there is also no variation within exon B. Similarly to G3pdh 1, there are variable positions in exons C +D (4 and 1 respectively).

The exons from the two loci differ from each other chiefly in the nucleotide which codes for the third position of the amino acid, in other words mutations are synonymous. Exon B differs in locus 1 from locus 2 by nine nucleotides, but the resulting amino acids are entirely constant. In exon C, of the 26 variable base pairs, 21 are at the third position and therefore the other five cause the production of different amino acids so 43 of the 48 amino acids are consistent. In the final exon (exon D), of the 14 variable nucleotides, only 4 are not at the third position and therefore they code for four different amino acids (out of 27).

Table 6.11. The lengths of the introns and exons in locus 1 and locus 2 of *G3pdh* gene

	Locus 1	Locus 2
Intron A	79	88
Exon B	98	98
Intron B	264	166
Exon C	143	143
Intron C	210	189
Exon D	84	84
Intron D	111	51

The main interesting observation that can be made from this table is that in both *G3pdh* loci, the same number of exons and introns are present. No intron has been lost as is sometimes the case in the evolution of different loci of the same gene (Charlesworth *et al.* 1998). Locus 1 was 170 bp on average longer than locus 2. The introns are simply longer in locus 1 than they are in locus 2, (apart from intron A). Intron B is 96 base pairs longer in locus 1 than it is for locus 2. Intron C is a further 34 base pairs longer in locus 1, and intron D another 60 bp.

Interestingly, when the amino acid sequence coded for by these loci are compared to cassava, the completely conserved exon B is also identical. Where the amino acids in exon C differ from cassava, three of the locus 1 amino acids agree with the cassava sequence and for the remaining two differences, the amino acids in locus 2 agree with the cassava sequence. There is also one amino acid toward the end of this exon where cassava has a T

and both of the *Eucryphia* loci have an S. Similarly, of the four differences in amino acids in exon D, two agree with locus 1 and two with locus 2.

The introns are generally variable between the two loci and in particular the intron B has 96 bp worth of deletions in locus 2 compared to locus 1. In intron C there are only 34 extra base pairs of deletions in locus 2. The final intron however again contains substantial deletions in locus 2 (60 bp).

There are a number of conclusions from this section of work. The first is that the intra-locus sequence of the exons within *Eucryphia G3pdh* loci, is very conserved. The statistics for inter-locus comparisons are as follows: exon B has undergone a high level of synonymous mutation (9%) but no nonsynonymous mutations, exon C has 15% synonymous mutations and 3% non-synonymous mutations and exon D has 12% synonymous mutations and 5% non-synonymous mutations. Neither locus is closer than the other to the single G3pdh locus in Cassava. There is a limited degree of homology between the introns of the different loci, intron B and intron D contain many more deletions than intron C.

6.8.16 Combined ITS and G3pdh phylogeny

The combined nuclear data produce two equally most parsimonious trees of length 331 (Fig. 6.29a and b). Again there is a small level of homoplasy (CI = 0.96, RI = 0.91, RC = 0.87 and HI = 0.045). The ITS data were added to the matrix containing the two *G3pdh* loci. Because there is no outgroup for the ITS data set, *Ceratopetalum* was simply coded as question marks. Therefore the combined data set has been polarised using *Ceratopetalum* data from the *G3pdh* loci. There is no conflict between the ITS tree and the *G3pdh* trees and combining these data results in the strengthening of many bootstrap values (Fig. 6.29). The Chilean species again form a group with 98% bootstrap support. There is now 80% bootstrap support for the Chilean group being sister to the group which contains the Tasmanian species and the South Australian *E. moorei*. Within this group, there is now 100% bootstrap support for the two different subspecies of *E. milliganii*. There is a further strengthening of the bootstrap support (of 64%) for a clade of the more northerly Australian species, *E. wilkiei* and *E. jinksii*.

This tree shows a strong correlation with geography. The northern Australian species, *E. jinksii* and *E. wilkiei*, form a relatively well supported group sister to the rest of the genus. The latter is split into two clades, one containing the two South American species, *E. cordifolia* and *E. glutinosa*, and another containing the Tasmanian (*E. lucida*, *E. milliganii*) and neighbouring South Australian *E. moorei*.

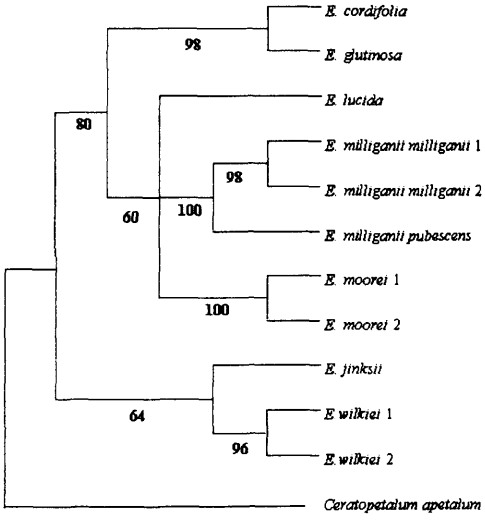

Fig. 6.29a. consensus tree for combined nuclear data

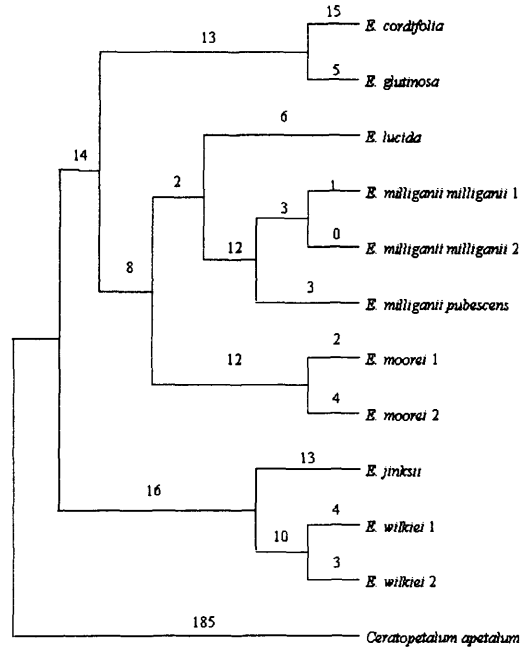

Fig. 6.29b. example of shortest tree for combined nuclear data

For the consensus tree (Fig. 6.29a) bootstrap values where they occur are below in bold. For the example of the consensus tree (Fig. 6.29b) the numbers of mutations are given above the clades.

6.8.17 Combined rpoA, trnL – F, trnH-K, ITS and G3pdh phylogeny – the total combined tree for all molecular data obtained

The combined chloroplast and nuclear data produce two equally most parsimonious trees of length 417. Again there is a small amount of homoplasy (CI = 0.950, RI = 0.890, RC = 0.845 and HI = 0.050). Again there is an undisputed 100% bootstrap support for the monophyly of *Eucryphia.* There is now also 100% bootstrap support for the clade of two Chilean species, *E. cordifolia* and *E. glutinosa.* There is 80% bootstrap support for the relationship of the Chilean species to the Tasmanian and southern Australian species *E. moorei.* The bootstrap support for the Tasmanian group is 92%, furthermore, the two supspecies of *E. milliganii* have 99% bootstrap support and there is 100% support for *E. milliganii ssp. pubescens* being a sister species to these. There is 63% bootstrap support for the inclusion of southern Australian species *E. moorei* in a group with the Tasmanian species. This suggests that they share a common ancestor. The two more northerly Australian species are unresolved and sister to all of the other species.

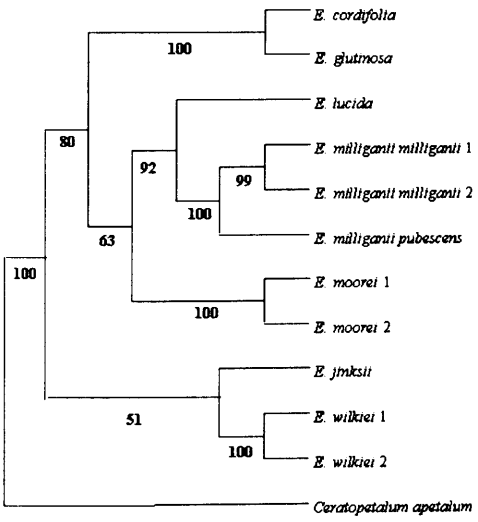

Fig. 6.30a. consensus tree for total combined data

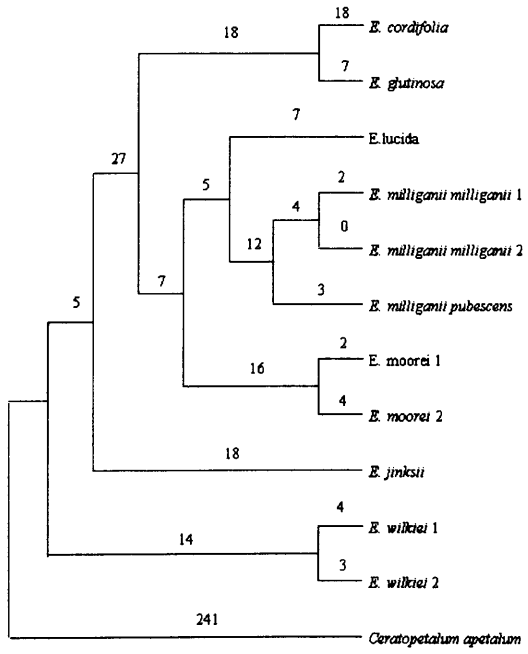

Fig. 6.30b. example of shortest tree for total combined data

For the consensus tree (Fig. 6.29a) bootstrap values where they occur are below in bold. For the example of the consensus tree (Fig. 6.29b) the numbers of mutations are given above the clades.

6.9 Sequence characteristics of the *Eucryphia* data set

Having displayed the sequence characters of the genes when discussed in turn, it is now useful for comparative purposes to see these data in one table (6.12). In such a fashion therefore these properties will be briefly considered for each gene region. It is then possible to state which sequences had more useful attributes than others.

Table 6.12. The sequence characteristics of the *trn*L-F, *trn*H-K, *rpo*A, ITS 1+2, *G3pdh* locus 1, *G3pdh* locus 2 and *G3pdh* locus 1+2 for 7 species of *Eucryphia*.

Parameter	*rpo*A	*trn* L-F	*trn* H-K	ITS 1 + 2	*g3pdh* locus 1	*g3pdh* locus 2
Length range bp	450	921-923	1209 – 1216	804 – 820	984 – 988	805 - 819
Length mean bp	450	921.2	1211.24	813	986.88	814.25
No characters (aligned length + indels)	450	924	1219	825	992	821
No. of indels (informative)	0	1 (0)	2 (1)	8 (4)	4 (2)	2 (0)
Size of indels, bp	N/A	[2bp]	7bp [1 bp]	1 x (1bp) 2 x (4bp) 1 x (2-4bp) [3 x (1 bp) 1 x (3bp)]	2 x (2bp) [1 x (2bp) 1 x (1bp)]	[1 x (10 bp) 1 x (2 bp)]
informative [uninformative]						
Number of constant sites (%)	447 (99.3)	909 (98.9)	1202 (99.3)	756 (93)	961 (97.4)	777 (95.5)
No. of autapomorphic sites (%)	1 (0.22)	4 (0.43)	9 (0.74)	21 (2.6)	18 (1.83)	32 (3.93)
No. of informative sites (%)	2 (0.44)	6 (0.65)	0 (0%)	36 (4.4)	8 (0.81)	5 (0.61)
Transitions	3	10	6	33	13	14
Transversions	0	0	3	24	13	23
Percentage transitions	100%	100%	67%	58%	50%	38%
Average no. of steps per character	0.007	0.013	0.009	0.083	0.052	0.047
Sequence divergence (%)	0 – 0.667	0 – 0.76	0 – 0.305	0.123– 4.605	0.10 – 1.53	0.49 – 1.72
G+C content mean, %	37.5	34.5	32.1	59.9	36.6	40.2
Number of trees found	1 (1)	1 (1)	1 (1)	1 (1)	1 (4)	6 (2)

6.9.1 Comparison of the properties of the data from genes sequenced for Eucryphia

The mean length of regions sequenced for *Eucryphia* ranged from 450-1219 base pairs in length. This is relevant in estimating the number of point mutations that are likely to stem from a certain proportion of variable bases. When discussing the utility of the different regions sequenced, it is useful to compare the proportion of constant bases, informative variable bases and autapomorphic bases. The indels are also an important factor. These factors are discussed in turn.

One important factor in the utility of gene regions to resolve species relationships is the proportion of variable bases, this ranged from 0.66% in the case of *rpo*A through to 7% in the ITS region. In constructing phylogenetic trees, the most useful type of variation is that of derived characters which are shared between two or more species, in other words the informative sites. This ranged from none in the *trn*H-K region to 4.4% in the ITS region. Because they were both relatively long regions a relatively resolved tree could be obtained from the 0.81% and 0.61% informative bases in the *G3pdh* 1 and 2 respectively.

Another useful factor in constructing phylogenetic trees, is the presence of informative insertions or deletions. Three regions which were sequenced did not have any informative indels (*rpo*A, *trn*L-F and *G3pdh* 2). One was present in the *trn*H-K region, two in the *G3pdh* locus 1, and four in the ITS region.

It is clear therefore that both the percentage variable informative bases and the number of informative indels are clear indicators to the resolution within a phylogenetic tree. Although the *rpo*A region has a lower number of variable bases than the *trn*H-K region, a higher proportion of these bases are informative as all of the variability observed in the *trn*H-K region is autapomorphic. Both the trees from the *rpo*A region and from the *trn*H-K region are fairly unresolved. However the tree for the *trn*H-K region does show the separation between the Chilean and the Australasian species in the *trn*H-K region due the presence of an informative indel. The *trn*L-F tree with its 0.65% of its 924 bases informative, is the most resolved of all chloroplast trees. From this discussion the importance of considering all parameters associated with the trees has been highlighted.

Other factors which need to be taken into account are the limitations of the different types of DNA. These are discussed in chapter 2. Although in this case be ITS data set has the

highest proportion of informative bases and the largest number of indels, one practical problem is that it is not possible to align these data with an outgroup. The *G3pdh* gene regions have both produced consistent well resolved trees, as has the chloroplast gene *trn*L-F.

The proportion of transitions is generally fairly high, particularly in the chloroplast regions where *trn*H-K has 67% transitions and in the other two regions 100% of all of the changes are transitions. Of all of the mutations in the ITS region 58% are transitions. The situation is different for the *G3pdh* genes where in the case of *G3pdh* 1, 50% of all mutations are transitions but only 38% in the case of *G3pdh* 2.

The average number of steps per character is again an indication of the resolution of the tree with the two most unresolved trees, the *rpo*A tree and the *trn*H-K tree having a value which is almost ten times smaller than for the other gene regions.

The sequence divergence gives a measure of the minimum and maximum distance between the different species, in all three chloroplast trees at least two species were identical and in all of the nuclear genes there is some difference between species, even if the difference is just autapomorphic.

The GC content for the data sets ranged from 32.1% in *trn*H-K, all of the chloroplast regions were around this value, the highest being *rpo*A at 37.1%. The ITS region had a fairly high GC content of 59%. The two *G3pdh* loci again had a fairly low value of 36.6 and 40.2% respectively for locus 1 and 2.

A final point to make concerns the number of most parsimonious trees. These are given both for the phylogenetic analysis of the *Eucryphia* data alone, the number is given in brackets when the outgroup *Ceratopetalum* is included. In most cases this value is one but in the case of the *G3pdh* genes, in locus 1, when the *Eucryphia* are analysed alone, four equally most parsimonious trees are found. Similarly for the *G3pdh* locus 2 region, six trees are found when the outgroup is included in the analysis and two are found when the *Eucryphia* data set alone is analysed.

6.10 Molecular clocks: dating the phylogenetic split

6.10.1 Relative rate tests

The first test to be established is whether there is rate heterogeneity across the different branches of the tree. If this is the case then it is possible to proceed with applying a molecular clock. The rate heterogeneity test advocated by Tajima (1993) was performed as described in chapter 4. It involves two components: m_A which is the number of sites where ingroup A has one nucleotide and ingroup B and the outgroup share a different nucleotide; and m_B, which is defined analogously with the alternative nucleotide shared with the outgroup . The quantity $(m_A - m_B)^2 / (m_A + m_B)$ is distributed approximating to a χ^2 distribution with one degree of freedom. Tables 6.13 - 6.18 show the m_A and m_B values and the value obtained for $(m_A - m_B)^2 / (m_A + m_B)$, for the different DNA sequences.

Where both subspecies of *E. milliganii* are identical, a single heading for the species is sufficient. In all tables, Glut = *E. glutinosa*, Cord = *E. cordifolia*, Jink = *E. jinksii*, Luc = *E. lucida*, Mill mill = *E. milliganii*, Moor = *E. moorei*, Wilk = *E. wilkiei.*
This Tajima test statistic follows a chi-squared distribution. The critical value at 1 degree of freedom and 99% significance level is 3.84.

Table 6.13. M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Eucryphia rpoA*

	Glut	Cord	Jink	Luc	Mill	Moor	Wilk
Glut	X	1,1	0,0	0,1	0,1	0,0	1,0
Cord	0.00	X	1,1	1,2	1,2	1,1	1,0
Jink	0.00	0.00	X	0,1	0,1	0,0	1,0
Luc	1.00	0.33	1.00	X	0,0	1,0	2,0
Mill	1.00	0.33	1.00	0.00	X	1,0	2,0
Moor	0.00	0.00	0.00	1.00	1.00	X	1,0
Wilk	1.00	1.00	1.00	2.00	2.00	1.00	X

All of the values in this table pass the test, on these grounds therefore the molecular clock can be applied.

Table 6.14. M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Eucryphia trn*L-F

	Glut	Cord	Jink	Luc	Mill Mill1	Mill Mill2	Mill Pub	Moor	Wilk
Glut	X	2,1	5,1	5,2	4,3	4,2	5,2	5,1	4,1
Cord	0.33	X	4,1	4,2	3,3	3,2	4,2	4,1	3,1
Jink	2.67	1.80	X	1,2	1,4	1,3	1,2	0,0	1,2
Luc	1.29	0.67	0.33	X	0,2	0,1	0,0	2,1	2,2
Mill Mill1	0.14	0.00	1.80	2.00	X	1,0	2,0	4,1	4,2
Mill Mill2	0.67	0.20	1.00	1.00	1.00	X	1,0	3,1	3,2
Mill Pub	1.29	0.67	0.33	0.00	2.00	1.00	X	2,1	2,2
Moor	2.67	1.80	0.00	0.33	1.80	1.00	0.33	X	1,2
Wilk	1.80	1.00	0.33	0.00	0.67	0.20	0.00	0.33	X

All of the values in this table pass the test, on these grounds therefore the molecular clock can be applied.

Table 6.15. M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Eucryphia trn*H-K

	Glut	Cord	Jink	Luc	Mill	Moor	Wilk
Glut	X	0,0	1,2	1,1	1,0	2,2	0,4
Cord	0.00	X	1,2	1,1	1,0	2,2	0,4
Jink	0.33	0.33	X	2,1	2,0	3,2	2,4
Luc	0.00	0.00	0.33	X	1,0	2,2	1,4
Mill	1.00	1.00	2.00	1.00	X	1,2	0,4
Moor	0.00	0.00	0.20	0.00	0.33	X	2,5
Wilk	4.00	4.00	0.67	1.80	4.00	1.29	X

In some cases, *E. wilkiei* just fails the test at the 1% level. It passes at the 5% level. Therefore it is acceptable to carry on with the application of a molecular clock.

Table 6.16. M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Eucryphia* ITS

	Glut	Jink	Luc	Mill Mill1	Mill Pub	Moor	Wilk
Glut	X	2,28	3,11	3,18	3,17	2,15	2,30
Jink	22.53	X	21,3	20,8	20,7	19,5	3,6
Luc	4.57	13.50	X	2,11	3,10	3,8	3,23
Mill Mill1	10.71	5.14	6.23	X	1,0	11,9	8,22
Mill Pub	9.80	6.26	3.77	1.00	X	10,9	7,22
Moor	9.94	8.17	2.27	0.20	0.05	X	5,21
Wilk	24.50	1.00	15.38	6.53	7.76	9.85	X

This table shows that there is really no rate consistency in the ITS data set. The worst offender is *E. wilkiei* with values of 24.5 when compared to *E. glutinosa* and 15.38 when compared to *E. lucida*. The test is failed by practically every comparison, the only acceptable values being those between *E. moorei* and the Tasmanian species *E. lucida* and *E. milliganii*. It is therefore not acceptable to apply a molecular clock to this data set.

Table 6.17. M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Eucryphia G3pdh* locus 1

	Glut	Cord	Jink	Luc	Mill Mill1	Mill Pub	Moor	Wilk
Glut	X	2,0	7,6	4,3	4,2	4,1	4,5	6,3
Cord	2.00	X	7,7	4,4	4,3	4,2	4,6	6,4
Jink	0.08	0.00	X	4,4	6,5	6,4	6,8	4,2
Luc	0.14	0.00	0.00	X	2,1	2,0	2,3	3,1
Mill Mill1	0.17	0.14	0.09	0.33	X	1,0	1,4	4,3
Mill Pub	0.20	0.67	0.10	2.00	1.00	X	0,4	3,3
Moor	0.11	0.40	0.29	0.20	1.80	4.00	X	6,4
Wilk	1.00	0.40	0.67	1.00	0.14	0.00	0.40	X

This table shows that all of the values except one passed the relative rate test at the 1% level. The failing species are *E. moorei* and *E. milliganii* subspecies *pubescens*. This test only just fails however, and passes at the 5% level. Therefore it is appropriate to apply a molecular clock to this data set.

Table 6.18. *M_A, M_B above the diagonal and $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for* for *Eucryphia G3pdh* locus 2

	Glut	Cord	Jink	Luc	Mill Mill	Mill Pub	Moor 1	Moor 2	Wilk1	Wilk2
Glut	X	4,6	6,5	3,0	2,4	2,3	2,2	2,5	7,3	7,3
Cord	0.40	X	5,6	6,5	5,9	4,8	5,6	5,10	6,5	6,4
Jink	0.09	0.09	X	6,4	5,8	5,8	5,6	5,9	6,4	6,3
Luc	3.00	0.09	0.40	X	0,5	0,5	0,3	0,6	5,5	5,4
Mill Mil	0.67	1.14	0.64	5.00	X	2,2	4,2	4,5	9,4	9,3
Mill Pub	0.20	1.33	0.64	5.00	0.00	X	4,2	4,5	9,4	9,3
Moor1	0.00	0.09	0.09	3.00	0.67	0.67	X	2,5	7,4	7,3
Moor2	1.29	1.67	1.14	6.00	0.11	0.11	1.29	X	10,4	10,3
Wilk	1.60	0.09	0.40	0.00	1.92	1.92	0.82	2.57	X	4,3
Wilk2	1.60	0.40	1.00	0.11	3.00	3.00	1.60	3.77	0.14	X

This table shows that most species comparisons are below the critical number. The Tasmanian species *E. lucida* does however appear to be evolving at a rate which is slightly out of synchrony with the other species. The values are slightly too high when compared both to the other Tasmanian species *E. milliganii*, and to the Australian *E. moorei*. Therefore I would conclude that a molecular clock can be applied, but caution must be taken.

Calibrating the clock

The molecular clocks were calibrated using the oldest *Eucryphia* fossil known, the 60 million years old specimen *E. falcata* described by Hill (1991). Because this is the oldest known *Eucryphia* fossil, and not assignable to any extant species, for the purposes of calibration it was placed at the root of the tree (shown in all phylogenetic trees with ✿).

Obviously it is incredibly useful to be able to have more than one estimate of divergence, and in *Eucryphia* we are lucky as other fossils are also known. Therefore for a second estimate of age, the earliest dates known from fossils of extant species, *E. lucida* and *E. milliganii* (both 2.6 million year old, Fig. 6.8) were used for a measure of internal calibration within the *Eucryphia* genus. These fossils are marked on all of the phylogenetic trees with ❀.

In order to calibrate a molecular clock, an average, minimum, and maximum rate of mutation K, was calculated. Firstly, the number of mutations over the tree subtended by the fossil was counted. K was then calibrated as follows:-

$$K = \frac{\text{number of mutations}}{2\,(\text{age of fossil} \times \text{length of sequence})}$$

For the *rpo*A data set no clock was used as the data was too invariable for a clock to be accurately applied. In the case of the ITS data, although it would have been possible to calibrate the rate of evolution using the more recent fossil data, the test is not applied due to the lack of rate constancy.

Table 6.19 shows the average, minimum and maximum rate in the number of changes per site per million years which have occurred over the four different gene regions that passed the rate constancy test. The information is given for the two different calibrations.

Table 6.19. Molecular clock calibration table for four gene regions

Gene region	*trn*H-K	*trn*H-K	*trn*L-F	*trn*L-F	*G3pdh*1	*G3pdh*1	*G3pdh*2	*G3pdh*2
Fossil age (MYO)	60	2.6	60	2.6	60	2.6	60	2.6
Length	1219	1219	924	924	989	989	819	819
Av. no. mutations	37.57	2.40	17	1.00	114.63	3	84.67	3.00
Min. mutations	36	1	15	0	111	1	81	3
Max. mutations	41	5	19	2	119	6	87	3
Rate (K) $x10^{-4}$	2.57	3.79	1.53	2.08	9.66	5.83	8.62	7.04
Min. K	2.46	1.58	1.35	0	9.35	1.94	8.24	7.04
Max. K	2.80	7.88	1.71	0.416	10.03	11.67	8.85	7.04

From the information on the rates of evolution for the different genes, the timing of the South American – Australasian disjunction can be dated. This is calculated by counting the average number of mutations between the South American species and the others in the sister clade. The date of the split was then calculated as follows:-

$$\text{Age of disjunction} = \frac{\text{(average number of mutations / sequence length)}}{K}$$

This date was calculated using the all three rates given in Table 6.19. The average number of mutations between the species in South America and those in Australasia, and the corresponding estimated divergence times are given in Table 6.20. Obviously the maximum divergence time is calculated using the minimum rate and vice versa.

Table 6.20. Divergence times for the South American – Australasian disjunction, for the four gene regions

Gene region	***trn*H-K**	***trn*H-K**	***trn*L-F**	***trn*L-F**	***G3pdh*1**	***G3pdh*1**	***G3pdh*2**	***G3pdh*2**
Fossil age (MYO)	60	2.6	60	2.6	60	2.6	60	2.6
Av. no. of mutations between split	3.6	3.6	6.7	6.7	15.17	15.17	13.25	13.25
Av. divergence time (MYO)	11.49	7.78	47.39	34.86	15.88	26.42	19.04	22.98
Min. divergence time (MYO)	10.54	3.74	42.40	17.43	15.35	13.20	18.77	22.98
Max. divergence time (MYO)	12.00	18.67	53.70	-	16.47	79.38	19.64	22.98

6.10.3 Application of the molecular clock

The first interesting point to make is that for each sequence the two estimates of divergence generally agree quite well with each other. This allows a fairly good degree of confidence in the molecular clock. When we look at the average divergence times we see that they range from 7.091 million years to 47.39 million years. This latter value was obtained for the *trn*L-F clock calculated using the 60 million year old fossil. This clock gives a much longer divergence time than the other seven clocks do. All other average divergence times are much lower than this with an average of 23.23 millions of years.

In order for the distribution to be explained in terms of vicariance, it is necessary to have approximately 40 million years worth of divergence. This is the date from which Australia was isolated from Antarctica and also when the land bridges between South America and Antarctica began to break. The average divergence times suggest that not enough time has elapsed since the separation of the Australasian species from the Chilean species to explain this distribution based on vicariance alone. It seems likely therefore that a dispersal event has occurred. Even allowing for error, in most cases, much less than half of the necessary evolutionary time for a vicariance explanation has occurred. If the minimum times are studied, again the explanation for all of the regions, must be based on dispersal. Even when studying the maximum divergence times only two of the estimates suggests suitable

divergence time whereby a vicariance explanation would be satisfactory (*trn*L-F with the 60 million year old fossil and *G3pdh* locus 1 with the 2.6 million year old fossil).

6.11 Discussion and conclusions

Having now discussed the results from the molecular work in the context of the evolutionary history and morphological data, it is now possible to bring many interesting strands together. It seems unequivocal that the two Chilean species form a single group. Phylogenetic data from the three nuclear genes suggests that they are most closely related to a group which contains the South Australian species *E. moorei* and the two Tasmanian species *E. lucida* and *E. milliganii*. Interestingly, neither the chloroplast nor the nuclear genes suggest that the other two species from Australia, the northern *E. wilkiei* and the central *E. jinksii* are any more related to *E. moorei* than they are to the Chilean species. Given the biogeography of the species, a likely scenario for their distribution is a single dispersal event. It is not possible from the cladograms to establish the directionality of this dispersal event. Because the outgroup genera are most abundant in Australasia than in South America, it is more likely that the direction of was from Australasia to South America.

The close relationship between *E. moorei* and Tasmanian species suggests that an ancestral species radiated both in southern Australia and in Tasmania. It seems likely that this ancestral species also radiated north to form *E. jinksii* and *E. wilkiei*.

From the application of the molecular clock, it seems likely that this dispersal event occurred between 10 and 20 million years ago. There is simply not enough molecular variation for an explanation to be consistent with vicariance. The most likely mechanism for such a dispersal event is in the ocean currents. As mentioned previously, the *Eucryphia* seeds contain no special mechanism to assist with rafting, so they possibly got caught in some form of debris.

CHAPTER 7

THE GENUS GRISELINIA

7.1 Introduction

The seven species of *Griselinia* J. R. Forster and G. Forster are evergreen, dioecious trees, shrubs (occasionally epiphytic). They may grow from 1 to 17 metres tall. The leaves are leathery, and have an alternate to sub-spiral or sub-opposite arrangement. They are simple, with margins that are entire to dentate (Fig. 7.1, 1-4). The flowers are small, actinomorphic with five green-yellow petals (Fig. 7.1, 1; the staminate flowers, always have petals but the female flowers lack petals in some species). The fruits are a 1-seeded fleshy drupe (Fig. 7.1, 3). Two species are found in New Zealand and five occur in Chile (Dillon & Muñoz-Schick, 1993).

The genus has been far less thoroughly studied than has either *Eucryphia* or *Coriaria*. A recent morphological revision was carried out by Dillon & Muñoz-Schick (1993). Although some recent work has been carried out on germination ability (Burrows, 1995), little has been published specifically on anatomy, chemistry or their fossil record. The only reported chromosome counts are for the New Zealand species of *Griselinia*: *G. littoralis* (Wanscher, 1933) and for *G. lucida* (Beuzenberg & Hair, 1959), in both cases n = 18.

Some work has been done on the relationship between *Griselinia* and invertebrates. It is known that the scale insects or coccoids overwinter in the leaves of *Griselinia* (Hodgson *et al.*, 2000). It is also known that caterpillars from two species of the moth genus *Ctenopseustis* are often found on the leaves of *Griselinia*; *Ctenopseustis herana* is found on *G. littoralis* and *Ctenopseustis obliquana* is found on both New Zealand species. (Horticulture and Food Research Institute of New Zealand Ltd., 1999).

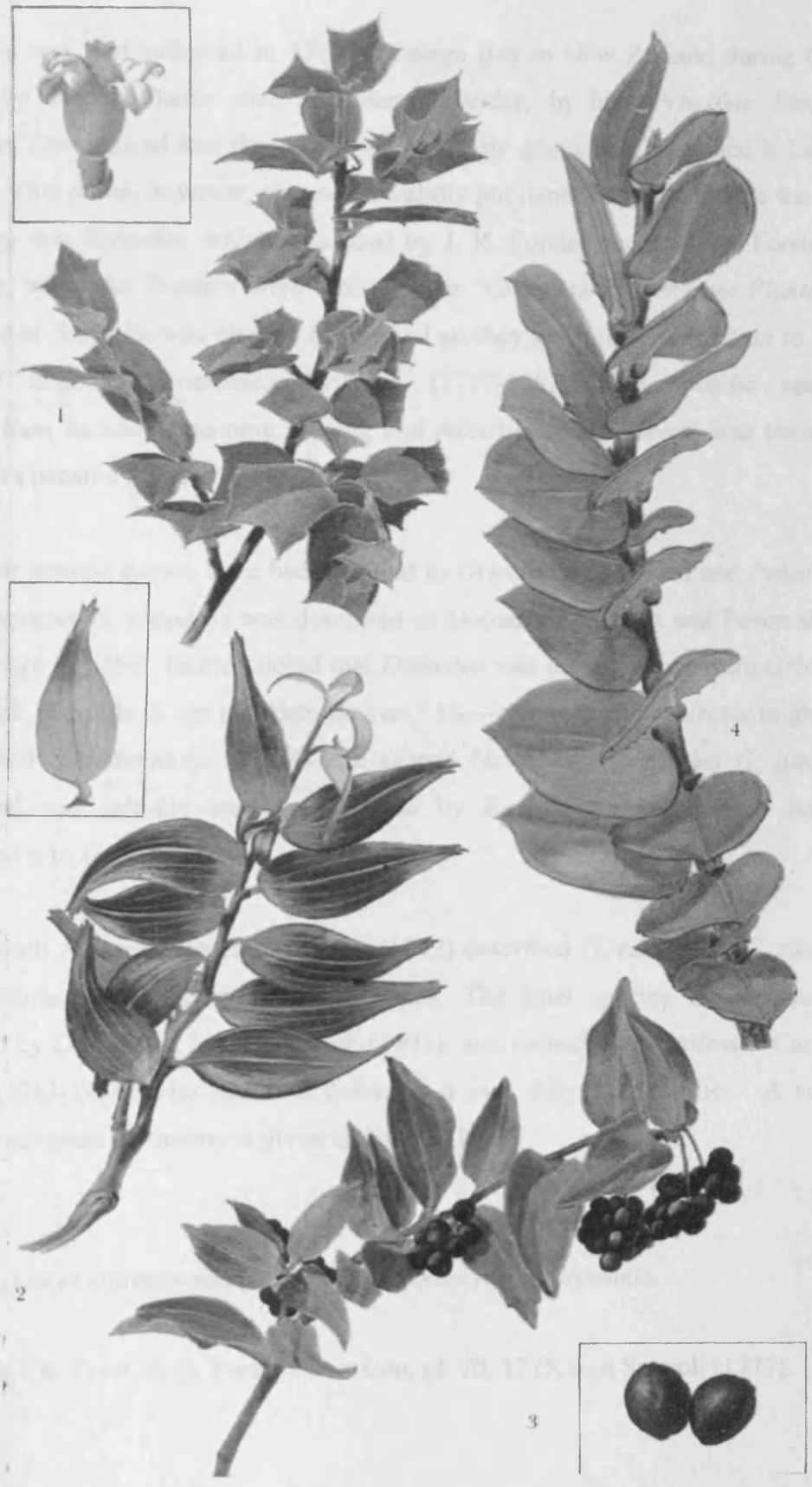

Fig 7.1. 1. *G. jodinifolia* leaves and close up of flower, 2. *G. ruscifolia*, showing the typical three points at the leaf apices, 3. *G. racemosa* showing the typical drupe, 4. *G. scandens*. From Hoffmann, (1997).

Griselinia was first collected in 1769 at Tolaga Bay in New Zealand during Cook's first voyage by Joseph Banks and Solander. Solander, in his "*Primitiae Florae Novae Zelandiae*", recognised that the taxon was in a new genus and he named it *Lissophyllum lucidum.* This name, however, was never validly published. The next name the genus was known by was *Scopolia,* which was used by J. R. Forster and J. G. A. Forster in 1775. However, when the Forsters were writing their "*Characteres Generum Plantarum*" they realised that *Scopolia* was already in use and so they substituted *Griselinia* to honour the Venetian naturalist Francesco Griselini (1717-1783). The specific epithet from *Lissophyllum lucidum* (meaning shining and referring to the leaves) was transferred and the species became the familiar *G. lucida.*

Two other generic names have been applied to *Griselinia*: *Decostea* and *Pukateria.* The Chilean species *G. scandens* was described as *Decostea* by Lopez and Pavon shortly after its discovery in 1794. Baillon noted that *Decostea* was distinguished from *Griselinia* only by the lack of petals in the pistillate flowers. He wrote that this character might be useful in sectional classification. In 1844 the second New Zealand species *G. littoralis,* was discovered and initially named *Pukateria* by Raoul; two years later, however, he transferred it to *Griselinia.*

Of the South American species, Taubert (1892) described *G. racemosa, G. ruscifolia* and *G. jodinifolia,* in his revision of the genus. The final species *G. carlomunozii* was described by Dillon and Muñoz-Schick (1993), and named after Professor Carlos Muñoz Pizarro (1913-1976) who had first collected it over fifty years earlier. A summary of currently accepted taxonomy is given in Table 7.1.

Table 7.1. List of currently accepted names and synonyms in *Griselinia.*

Griselinia J.R. Forst. & G. Forst., Char. Gen. pl. 70. 1775, non Scopoli (1777).

Chilean species

Griselinia carlomunozii M.O. Dillon & Muñoz-Schick, in Brittonia, 45, 4, 265 (1993).

Griselinia jodinifolia (Griseb.) Taubert, in Bot. Jahrb. Syst. 16: 390 (1892).
Decostea jodiniforia Griseb., Syst. Bemerk. 34 (1854).
Vernacular names: yelmo chico, tribillo

Griselinia racemosa (Phil.) Taubert, in Bot. Jahrb. Syst. 16: 390 (1892).
Decostea racemosa Phill., in Linnaea 28: 703(1857).
Vernacular names: lilinquén, yelmo, lamulahuen

Griselinia ruscifolia (D. Clos) Taubert in Bot. Jahrb. Syst. 16: 391 (1892).
Decostea ruscifolia D. Clos, in Gay, Fl. Chil. 8: 395 (1854).
Vernacular name: voqui, lilinquén

Griselinia ruscifolia var. ***itatiaiae*** (Wawra) Taubert, in Bot. Jahrb. Syst. 16: 392 (1892).
Maytenus itatiaiae Wawra, in Osterr. Bot. Z. 32: 38 (1882).
Vernacular name: Erva-de passarinho-dos-andes

Griselinia scandens (Ruiz & Pav.) Taubert, in Bot. Jahrb. Syst. 16: 391 (1892).
Decostea scandens Ruiz and Pav., Syst. Veg. Fl. Peruv. Chil. 259 (1798).
Griselinia alata Ball, in J. Linn. Soc. 22: 163 (1886).
Vernacular name: yelmo

New Zealand species

Griselinia littoralis (Raoul) Raoul, in Choix Pl. Nouv. – Zél. 22 (1846).
Pukateria littoralis Raoul, in Ann. Sci. Nat., Bot. Ser. 3, 2: 120 (1844).
Vernacular names: broadleaf, kapuka, papaumu, papauma

Griselinia lucida (J. R. Forster & G. Forster) G. Forster, Fl. Ins. Austr. 75 (1786).
Scopolia lucida J. R. Forster & G. Forster, Char. Gen. Pl. 140 (1775).
Griselinia lucida var. *macrophylla* J.D. Hooker, Handb. N. Zeal. Fl. 105 (1864).
Vernacular names: broadleaf, puka, fat-belly (New Zealand)

7.2 Relationships to other families and genera

On the basis of morphological, chromosomal and palynological evidence, Griseliniaceae is accepted as a monogeneric family (Thorne, 1992). *Griselinia* has in the past, however, been placed in many different families. Sprengel in 1817 and Agardh in 1823 both assigned it to the Euphorbiaceae. Later, Kunth in 1892 allied it with the Juglandaceae.

Although Reichenbach originally agreed with this designation Reichenbach later reconsidered and in 1837 he placed the genus in the tribe Crozophoreae of the family Euphorbiaceae. Araliaceae was the family that Endlicher decided on in 1850. Two years later, Hooker suggested that *Griselinia* was a member of the Cornaceae tribe Aucubeae. Other families that have been suggested as being related to *Griselinia* are Aquifoliaceae and Celastraceae (Dillon and Muñoz-Schick 1993).

Many scientists still place *Griselinia* in the Cornaceae (Cronquist, 1981; Thorne, 1968). Philipson disagreed with this placement and initially placed the genus closer to either Araliaceae or Escalloniaceae. However, after extensive studies, he decided that *Griselinia* should be placed in the informal group Unitegminae which is a complex of orders and families that possess ovules with a single integument; this group includes the Cornales, Araliales, Escalloniaceae and Pittosporaceae (Philipson, 1967, 1977). Takhtajan (1980) considered the genus *Griselinia* to be a monotypic family *Griseliniaceae* (Wangenh.) Takht., and felt that it was most closely related to *Garrya* (Garryaceae). More recently, Thorne (1992), when accepting the family *Griseliniaceae* put the Hydrangeales as the closest relatives, rather than the Cornales.

Recent *rbc*L evidence indicates a close affinity with some of genera of the Caprifoliaceae and Araliaceae (Chase *et al.*, 1993). Similarly using *rbc*L data, in a more detailed study on the relationships within the Apiales (Apiaceae and Araliaceae), Plunkett *et al.* (1995) showed that the family status for *Griselinia* was deserved and that it is most closely related to a clade containing the Apiales and Pittosporaceae.

7.3 Reproductive biology

All species of *Griselinia* are dioecious. The individual flowers of *Griselinia* are actinomorphic and arranged in racemes or panicles which are either terminal or in the axils of branchlets. The staminate flowers contain no gynoecium and, similarly, no androecium is found in the pistillate flowers (Dillon and Muñoz-Schick, 1993). The male flowers always have five, yellow to greenish-yellow petals that are reflexed at anthesis and quickly caducous. The five anthers are opposite the sepals and alternate with the petals. These staminate flowers also have a campanulate hypanthium (Fig. 7.2). In four species (*G. carlomunozii*, *G. lucida*, *G. racemosa*, *G. ruscifolia*, *G. scandens*) the female flowers do

not develop petals. Pistillate flowers are epigynous and, when petals are present, there are five of them and again they are yellow to greenish. The ovary has three carpels with two locules per carpel, although only one locule contains an ovule. The single stigma contains three terminal styles which are distinct and subulate and either divergent or recurved.

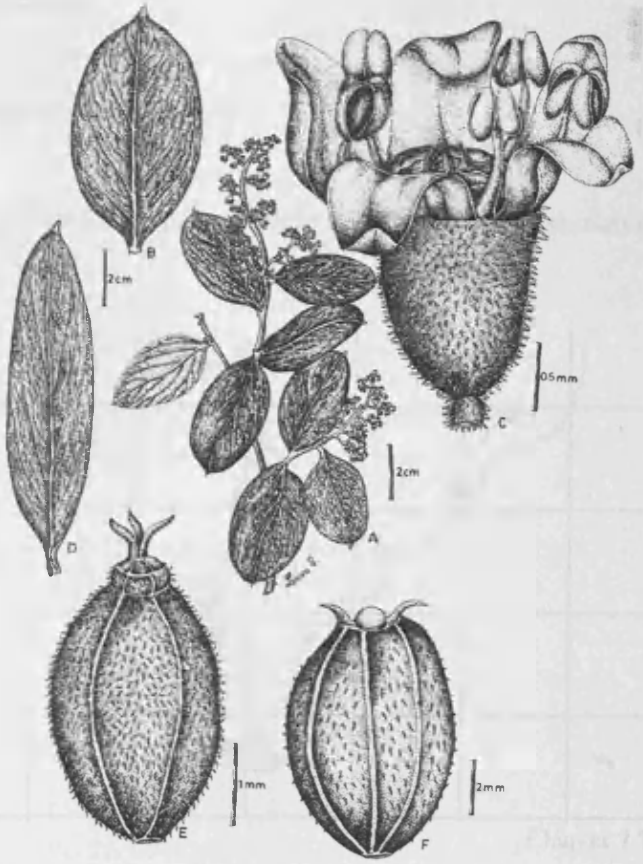

Fig. 7. 2. *Griselinia carlomunozii* **A.** Flowering branch, **B**. Leaf from a staminate individual. **C**. Staminate flower. **D**. Leaf from a pistillate individual. **E**. Ovary. **F**. Fruit, (From Dillon and Muñoz-Schick (1993).

Schlessman *et al.* (1990) has suggested that the dioecious nature in Apiales arose from ancestral andromonioecy, but this has not been confirmed for *Griselinia*.

The inconspicuous, greenish-yellow flowers of *Griselinia* are characteristic of the type that are pollinated by small flies (Johnson, 2000). Pollen is shed from *Griselinia* as single grains. The grains are 3 - colporate and ellipsoidal, with a complete tectum with striate surfaces (Ferguson and Hideux, 1978; Tseng, 1980).

The fruit of *Griselinia* is a fleshy drupe; it is therefore likely that it is are dispersed by animals, probably birds. Work by Burrows (1995) has shown that *G. littoralis* does not germinate unless the fleshy pericarp tissue is removed. This would happen naturally as the seed was passed through the bird's gut. The seeds would naturally be dropped in light conditions and this is reflected by the fact that, in the dark, the germination rate for *Griselinia* seeds is much lower than in the light (Burrows, 1995).

7.4 Distribution and ecology

7.4.1 South American species

Five species occur in South America and their distributions are shown in Fig. 7.3.

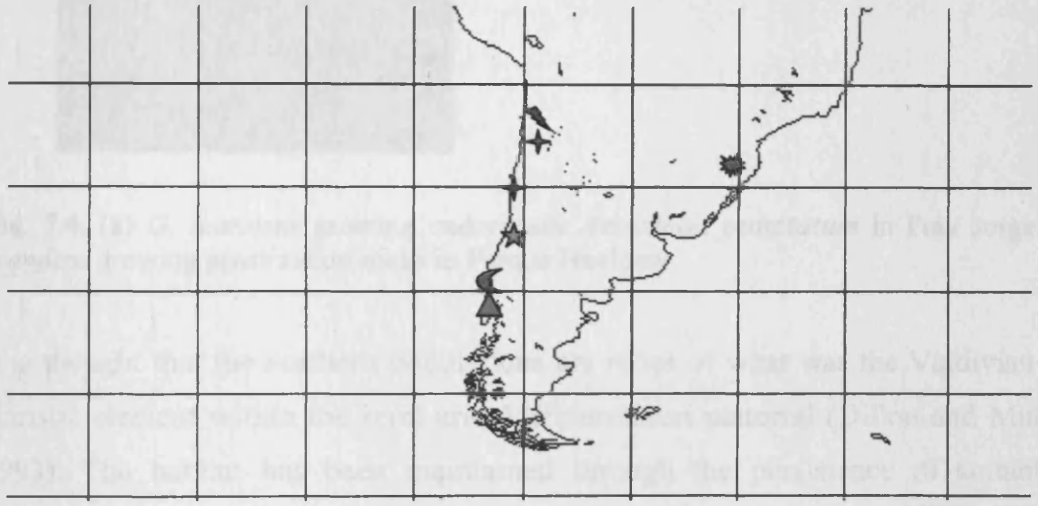

Fig. 7.3. Map of South America showing the approximate distributions of *Griselinia*. *G. carlomunozii*, *G. jodinifolia*, *G. racemosa*, *G. ruscifolia* var. *ruscifolia*, *G. ruscifolia* var. *itatiaiae* and *G. scandens*.

G. carlomunozii is found only in the very north of Chile in the Antofagasta region. It grows on rocky slopes at *ca.* 400 m elevation where it is confined to damp 'islands' within an otherwise arid coastal escarpment, obtaining its moisture from the clouds. This type of vegetation is generally known as *lomas* vegetation and it is highly endemic (Johnston, 1929, Rundel *et al.*, 1991). Other plants found in association with *G. carlomunozii* are: *Polypodium masafuerae, Peperomia doelli, Calceolaria paposana* and *Tillandsia tragophoba* (Dillon and Muñoz-Schick, 1993).

G. scandens is found mainly in central Chile in the Concepción region (ca. 37°S) but a few disjunct populations exist 700 km north of here the northernmost being at Bosque de Fray Jorge (30°30′S; Fig. 7.4a). In the southerly part of its distribution *G. scandens* is found both on the coast, where it grows prostrate on the rocks (Fig. 7.4b), and further inland, where it grows as a tree in close association with other woody shrubs and small trees.

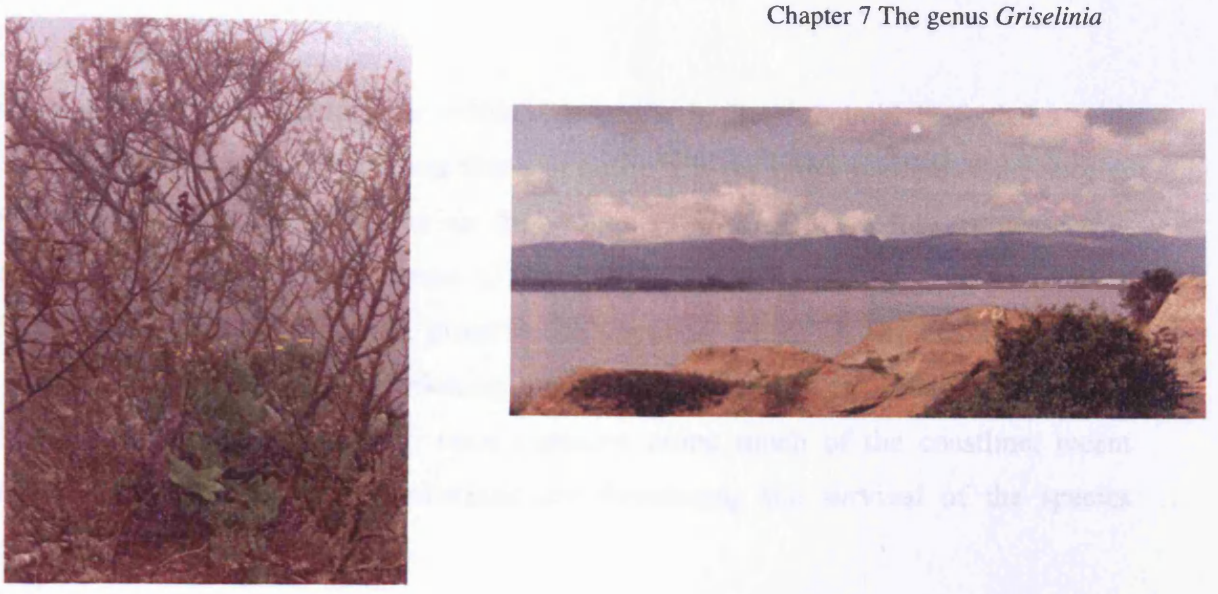

Fig. 7.4. (a) *G. scandens* growing underneath *Aetoxicon punctatum* in Fray Jorge and **(b)** *G. scandens* growing prostrate on rocks in Parque Hualpen

It is thought that the northern populations are relics of what was the Valdivian rain forest floristic element within the semi-arid Mediterranean matorral (Dillon and Muñoz-Schick 1993). The habitat has been maintained through the persistence of suitable climatic conditions; where the Humbolt current hits the coast it causes small areas of the adjacent land to have high levels of precipitation. The northern populations of *G. scandens* are only found in these small pockets of cloud forest, where water vapour condenses on its shiny leaves. *G. scandens* often grows in close association with the olive *Aetoxicon punctatum.* Initially it tends to grow underneath these trees (Fig. 7.4a), they then overtake and outgrow them and invade the area thus preventing regeneration of other plants. Work is currently underway at the university of Santiago to investigate this relationship. Other plants that grow in association with *G. scandens* in the cloud forests are *Myrceugenia correafolia* (Myrtaceae), *Drimys winteri* (Winteraceae), *Berberis chilensis* (Berberidaceae) and *Kageneckia oblonga* (Rosaceae) (pers. obs.; pers. comm. from Luis Jopia Cortés).

Two main theories have been advanced to explain the close relationship between plant populations in the cloud forests of northern Chile and their allopatric counterparts in the south (Arroyo *et al.*, 1996). The first involves range reduction owing to the onset of drier conditions in the Pleistocene, which caused the formerly continuous genepool to become fragmented. The second speculates that there has been a range expansion, with the allopatric populations being generated by dispersal.

Griselinia jodinifolia is distinctive within the genus in its rhomboid leaf blades with spinose teeth, superficially resembling those of holly. It is the most tree-like of the species in South America and the only one on that continent with pistillate flowers possessing petals (Fig. 7.1). It grows in the centre of Chile, from Maule to Osorno, both as a tree to 4m (Fig. 7.5) and as a dwarf shrub, prostrate on sea-cliffs where for large stretches of the coast it is the dominant species blanketing them. In this latter habitat it plays a major role in substrate stabilisation. Although once common along much of the coastline, recent *Eucalyptus* and *Pinus radiata* plantations are threatening the survival of the species (Carloz Pamirez, pers. comm.).

Fig. 7.5. *G. jodinifolia* **a)** growing as a tree and **b)** prostrate form

G. racemosa is found from Cautin to Aisen, and from sea level to an altitude of 1250 m. It grows as a shrub to two metres and has very distinctive leaves that end in between one and three teeth (Fig. 7.6). It is common in moist forests where it is associated with *Drimys* (Winteraceae), *Nothofagus* (Fagaceae), *Podocarpus* (Podocarpaceae), *Fitzroya* (Cupressaceae) and *Weinmannia* (Cunoniaceae).

Fig. 7.6. *G. racemosa* leaf, showing the typical tri-partite tip.

7.4.2 New Zealand species

Two species of *Griselinia* occur in New Zealand; their distribution is shown in Fig. 7.7.

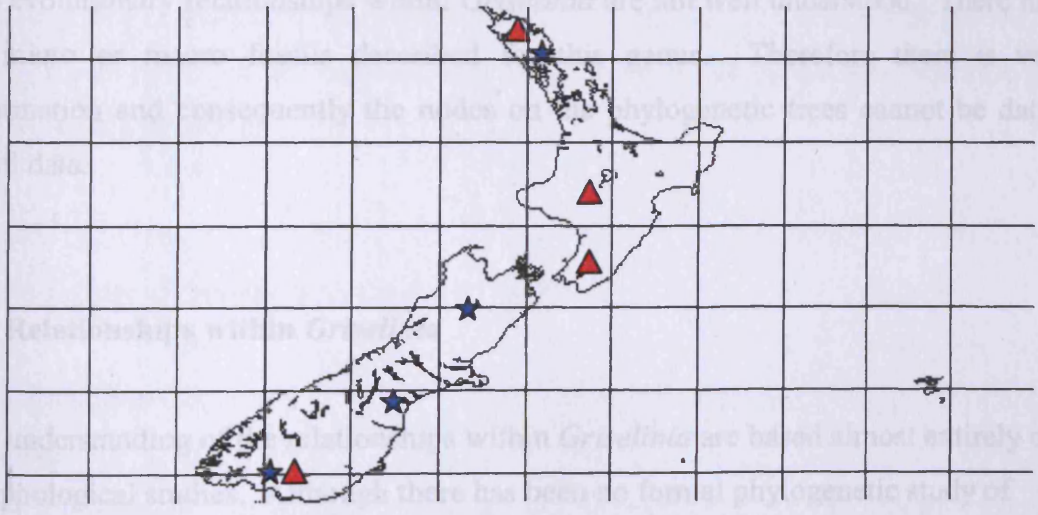

Fig.7.7. Map of New Zealand showing the approximate distributions of *G. littoralis* ★ and *G. lucida* ▲.

G. littoralis is a round-topped tree which grows to 17 m tall with short, gnarled trunks that grow to 1.5 m in diameter. It is found in lowland forest on both the North and South Islands, extending from North Cape in the north to the Foveaux Strait in the south. It is

more plentiful on the North Island than it is on the South Island and ranges in elevation from sea level to 330 m.

G. lucida is a tree which, although it will grow to between 8 – 20 m tall, before maturation it has an initially epiphytic phase. It generally grows as an epiphyte amongst *Astelia* (Liliaceae) or other epiphytes that cover the branches of *Metrosideros robusta* Cunn. (Myrtaceae), it then sends down roots so it can establish itself as an independent tree. Although it superficially resembles *G. littoralis*, it is a much more branched shrub with leaves that are larger and far more leathery and glossy. It also lacks petals on the pistillate flowers. It grows from the lowland and montane forests from North Cape to Foveaux Strait (330 m). In contrast to *G. littoralis*, it is far more plentiful in the North Island than in the South Island where it tends to be replaced by *G. littoralis*.

7.5 Evolution in *Griselinia*

The evolutionary relationships within *Griselinia* are not well understood. There have been no micro or macro fossils described for this genus. Therefore there is very little information and consequently the nodes on the phylogenetic trees cannot be dated using fossil data.

7.6 Relationships within *Griselinia*

Our understanding of the relationships within *Griselinia* are based almost entirely on morphological studies. Although there has been no formal phylogenetic study of *Griselinia*, from a detailed morphological review of the genus Dillon (pers. comm.) suggests that the group has a New Zealand origin and then spread to South America via dispersal. He believes that the radiation of the genus into more arid areas is secondary.

The New Zealand species are thought to be more closely related to each other than to the Chilean species (Dillon & Munoz-Schick, 1993). From the morphological data, it is difficult to determine the relationships between the Chilean species of *Griselinia*. The two species from the south of Chile, *G. racemosa* and *G. ruscifolia*, are morphologically more similar to each other than they are to the central and northern species. The others are all

fairly distinctive in different ways. *G. scandens* has large flat leathery leaves whilst *G. jodinifolia* has the appearance of a fleshy holly.

On the basis of leaf and flower morphology alone, is not possible to ascertain the relationship between the newly described *G. carlomunozii* and the other Chilean species (Dillon and Munoz-Schick, 1993). This species has a combination of characters that is not found in other members of the genus. This includes the small entire leaves with dark green and shiny adaxial surfaces, prominently veined abaxial surfaces, mucronate apices and hispidulous ovaries. The overall morphology suggests a possible relationship with *G. racemosa* which also has non-spinose leaves. However, *G. racemosa* has dentate leaf apices and only three prominent abaxial nerves arising near the base. There are also some morphological similarities between *G. carlomunozii* and *G. ruscifolia* (again *G. ruscifolia* has non-spinulose leaves).

Although it is the geographicaly nearest species to *G. carlomunozii*, morphologically *G. scandens* is very distinct from it (*G. scandens* the winged stems, larger denticulate-spinulose leaves with 2-ranked phyllotaxis and strictly axillary inflorescences and glabrous fruits; the flowers of *G. carlomunozii* may be terminal or axillary and the fruits are hispidulose-puberulent).

7.7 Molecular analyses: materials and methods

7.7.1 Plant material

Table 7.2 lists the collection details of the species used in this study. Leaves were collected from living specimens and dried in bags of silica gel prior to DNA extraction.

Table 7.2. Collection details of accessions of *Griselinia* used in the present study. Voucher specimens for all the Chilean material (including *G. carlumonozii*) are lodged in the natural history museum in Santiago. *G. lucida* is in the private herbarium of Bernard Honey. Collections from the Royal Botanic Garden Edinburgh are lodged therein.

G. carlomunozii

Chile, Región II [Prov. Antofagasta] Quebrada Rinconada, ca. 5 km N of Caleta Paposo (24° 56′ S, 70° 30′ W) 240 – 440 m, *M.O. Dillon & D. Dillon 5864* (**SGO**)

G. jodinifolia

Chile, Región X [Los Lagos], 6/6/1998. *M.F.Gardner & C.N.Page 5054, ex* Royal Botanic Garden Edinburgh (accession 19924311).

Chile: Región X [Prov. Valdivia], c. 1 km E of La Mision (39° 47′ S, 73° 23′ W), 12 m, 12/9/1999. *M.J.R. Clokie* (**SGO**).

G. littoralis

New Zealand, Kaimanawa Ecological Region, *Edinburgh New Zealand and Australia Expedition 153,* ex Royal Botanic Garden Edinburgh (accession 19981045).

G. lucida

Waitakere Hills/Mts just west of Auckland (36° 55′ S, 174° 30′ E), 10/6/2000 *Bernard Honey*

G. racemosa

Chile, Región X [Prov. Osorno], Puyehue National Park, Casilla 1337, Osorno, *M.R.J.* 14/9/1999. *Clokie & José Miguel Lopez Rivera* (**SGO**)

Date collected

Chile, Región X [Prov. Chiloé], Park National Chiloe (42° 37′ S, 74° 06′ W), 7 m, 17/9/1999. *M.R.J. Clokie* (**SGO**)

G. ruscifolia

Ex hort. Hillier Arboretum, Hampshire

Chile, Región X [Prov. Osorno], Puyehue National Park, Casilla 1337, Osorno, 15/9/1999. *M.R.J. Clokie & José Miguel Lopez Rivera* (**SGO**)

G. scandens

Chile, Región IV [Prov. Coquimbo], Fray Jorge, Ellenberg 4669, Ovalle (30° 40′S, 71° 41′W), 22/8/1999. *M.R.J. Clokie & José Miguel Lopez Rivera* (**SGO**)

Chile, Región VII [Prov. Cauqueres], at beach at Pelluhue, (35° 49′S, 72° 35′W) 0 m, 2/9/1999. *M.R.J. Clokie & Patricio Peñoilillo* (**SGO**)

Chile, Región VIII [Prov. Concepción], Parque Hualpén, Dpto. Talcahueno, Concepción (36°47′S, 73° 10′W), 12 m, 9/9/1999. *M.R.J. Clokie & Marcelo Baeza* (**SGO**)

Hedera 'Hibernica' L.

Ex University of Leicester campus. 2/4/2000

Schefflera digitata

Sequences downloaded from EMBL (EMBL:U63188) from work by Mitchell and Wagstaff (1997).

Viburnum davidii

Unlocalised, ex University of Leicester Botanic Garden.

7.7.2 Methods

DNA sequences from four genes were obtained following the methods described in Chapter 4. Three of the sequences were from the chloroplast (*rpo*A, *trn*L-F and *trn*H-K) and the ITS gene was from the nuclear genome.

7.7.3 Analyses

Initially, based on recent molecular work (Chase *et al.*, 1993), two outgroups were chosen for *Griselinia: Hedera helix* (Araliaceae) and *Viburnum davidii* (Caprifoliaceae). In all the chloroplast gene regions that were amplified in terms of the number of mutations, *Viburnum* was the closer of the two outgroups so it was therefore selected as the outgroup for all further analysis. Although the ITS region sequenced well both for *Hedera* and *Viburnum*, it was not possible to align these sequence data with those of *Griselinia.* However a BLAST search revealed that the closest species (in terms of having the fewest

7.8 Molecular phylogenies

7.8.1 rpoA phylogeny

The *rpo*A sequences produced a single most parsimonious tree of length 13 (Fig. 7.8). There was no homoplasy (CI = 1, RI = 1, RC = 1, HI = O). The clearest observation from these data is that the genus *Griselinia* is monophyletic. Within the genus, there is a clade which contains the two New Zealand species *G. littoralis* and *G. lucida* with moderate bootstrap support of 62%.

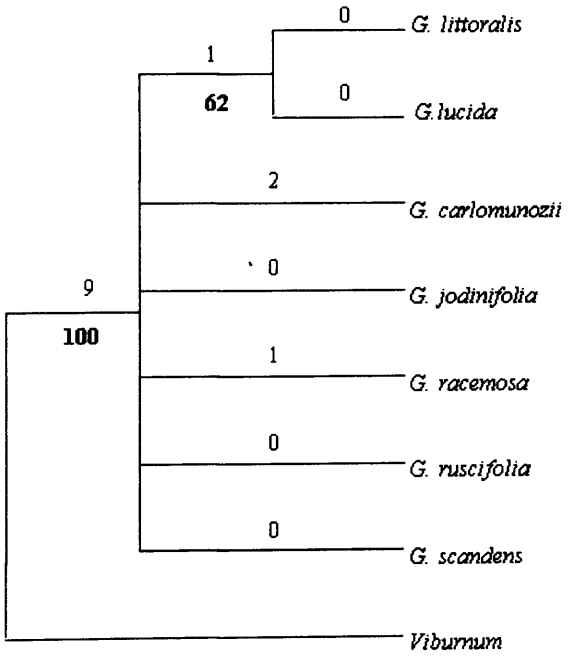

Fig. 7.8. Phylogeny for *rpo*A in *Griselinia*. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

7.8.2 rpoA variation

Statistics relating to the properties of *rpo*A data are given in Table 7.3. From this information it can be seen why the *rpo*A tree is poorly resolved. Of the four (0.91%) variable bases, only one (0.23%) is informative and the other three (0.69%) are autapomorphic. The low rate of change is reflected in the low average number of steps per character of 0.009. Interestingly, half of the point mutations are transitions and half are transversions. One transition and one transversion are observed in the extreme northern species *G. carlomunozii*. The other transversion is found in *G. racemosa* and the remaining transition is shared between the two New Zealand species of *Griselinia*. There are no insertions or deletions in the data set. There is no sequence divergence between some species of *Griselinia* and the maximum divergence is a relatively low 0.70%. The relatively low GC content of 37.4% is typical for the chloroplast genome (Gillham, 1994).

Table 7.3. Statistics relating to *rpo*A sequence data in *Griselinia*

Length range (bp)	429
Length mean (bp)	429
Number of characters (aligned length + indels)	429
Number of indels	None
Size of indels (bp)	N/A
G+C content mean %	37.4
Sequence divergence %	0 – 0.70
Number of constant sites (%)	425 (99.08)
Number of autapomorphic sites (%)	3 (0.69)
Number of informative sites (%)	1 (0.23)
Transitions	2
Transversions	2
Percentage transitions (%)	50
Average number of steps per characters	0.009

The pattern of variation is illustrated in Fig. 7.9, where it can be seen that the four variable bases change once each to produce the most parsimonious tree. Three of the four variable bases are found at the beginning of the sequence at positions 8, 31 and 49. The remaining variable base is at position 315. There is also little variation observed when the ingroup is analysed with the outgroup. This variability is however spread more evenly along the region and does not correlate with the *Griselinia* only variation.

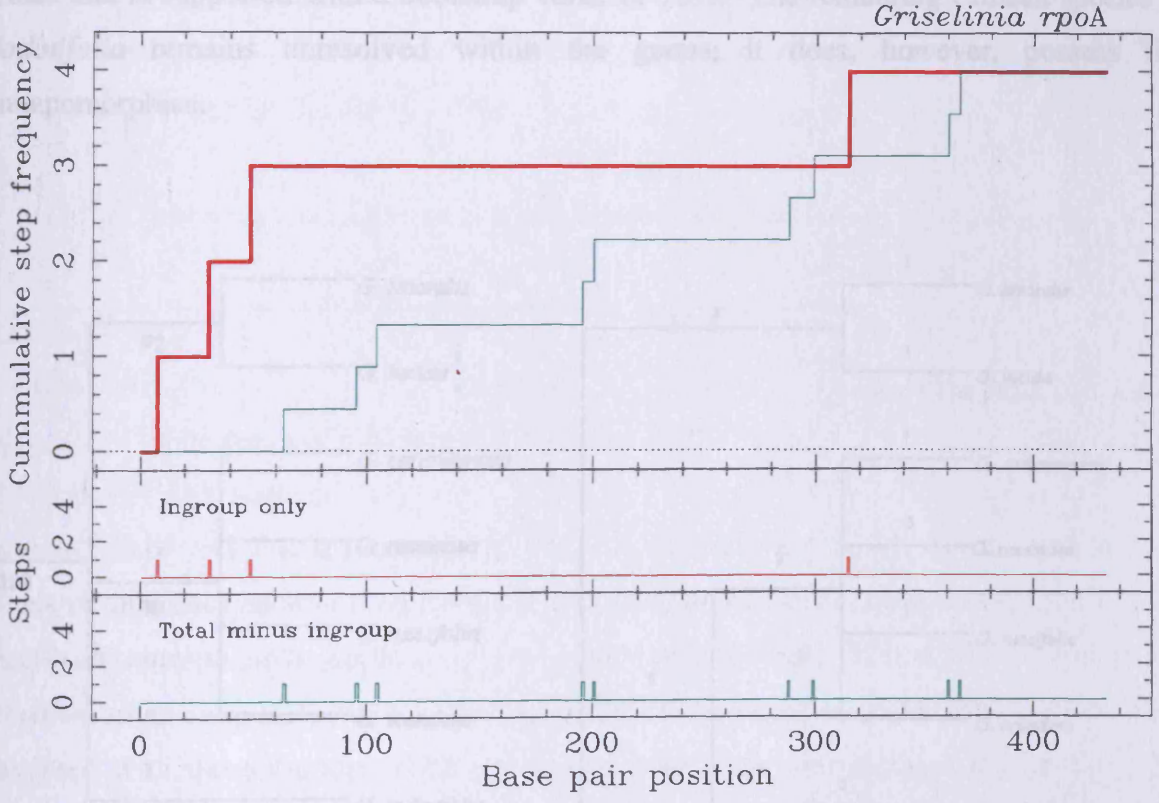

Fig. 7.9. Graph to illustrate the behaviour of *rpo*A in *Griselinia* and its outgroup *Viburnum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Viburnum*) from the ingroup (*Griselinia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

7.8.3 trnL–F phylogeny

The *trn*L-F sequences produced three equally most parsimonious trees of length 84. There is a small amount of homoplasy (CI = 0.988, RI = 0.900, RC = 889, HI = 0.012). The consensus tree and an example of a shortest tree are shown in Fig. 7.10. Within *Griselinia* the two New Zealand species form a well supported clade with a bootstrap value of 92%. This indicates the presence of a single trans-Antarctic disjunction within *Griselinia*. Within the Chilean species, *G. lucida*, *G. littoralis*, *G. racemosa* and *G. ruscifolia* form a clade that is supported with a bootstrap value of 70%. The remaining Chilean species *G. jodinifolia* remains unresolved within the genus; it does, however, possess five autapomorphies.

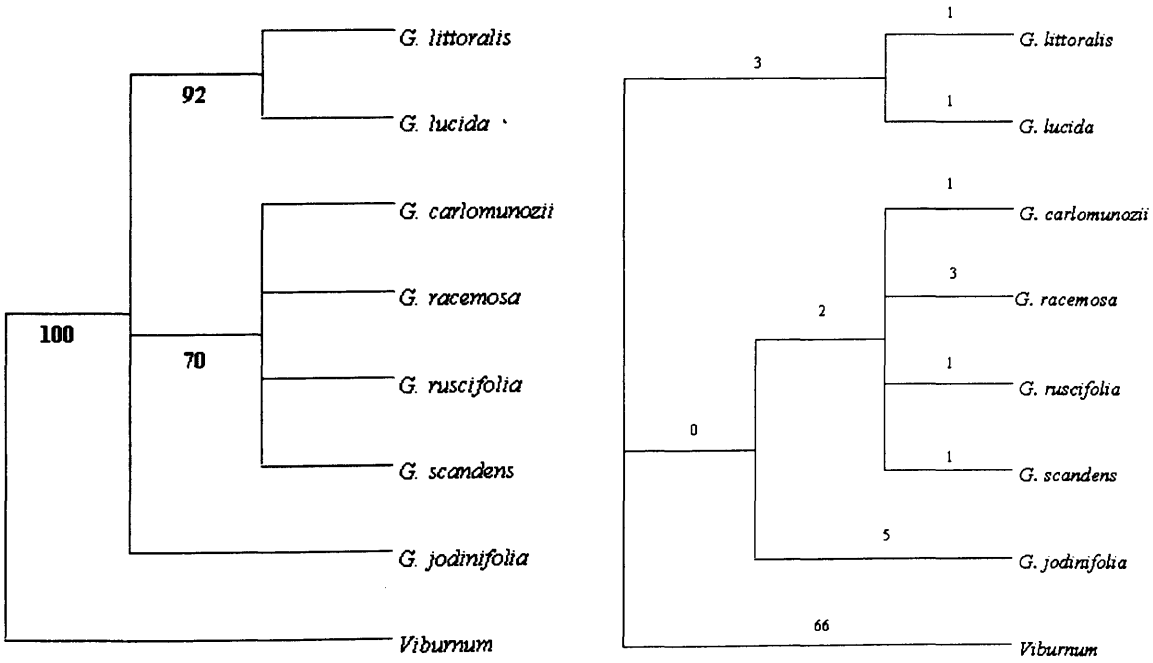

Fig. 7.10 a. Consensus tree for *trnL-F* **b.** Example of shortest tree. For the consensus tree bootstrap values where they occur are below in bold. For the shortest tree the numbers of mutations are given above the clades.

7.8.4 trnL-F variation

Statistics relating to the properties of the *trn*L-F data are given in Table 7.4. In this data set 15 (1.63%) of the bases are variable, but of these only 3 (0.32%) are informative. The average number of steps per character is 0.018. The two New Zealand species share a 15 base pair deletion and the one base pair insertion is shared both by the New Zealand species and the Chilean *G. jodinifolia*. Contrary to the case in *Eucryphia*, in the point mutations, there are a similar number of transversions and transitions (ratio of 0.875). The sequence divergence ranges from 0.216-0.97. The low GC content of 34.8 is typical for the chloroplast region (Gillham, 1994).

Table 7.4. Statistics relating to *trn*L-F sequence data in *Griselinia*.

Length range (bp)	913 – 928
Length mean (bp)	923.14
Number of characters (aligned length + indels)	938
Number of indels (informative)	2 (2)
Size of indels (bp) [uninformative]	15, 1
G+C content mean %	34.8
Sequence divergence %	0.216 – 0.97
Number of constant sites (%)	908 (98.37)
Number of autapomorphic sites (%)	12 (1.3)
Number of informative sites (%)	3 (0.32)
Transitions	7
Transversions	8
Percentage transitions	46.66
Average number of steps per characters	0.018

The pattern of variability is illustrated in Fig. 7.11. Interestingly, although three equally most parsimonious trees are obtained when a phylogenetic analysis using parsimony is performed on the genus data set with the outgroup, when it is analysed without the outgroup only one most parsimonious tree is obtained. Of the three trees obtained when the data are analysed with *Viburnum*, the bases change at the same position for all three trees. When the ingroup alone is considered 15 of the base pairs in the *Griselinia* sequence data change. The last two peaks in the genus data set represent the coding of insertions/deletions. The 15 base pair deletion that the two New Zealand species share occurs at the position 594. The one base pair deletion occurs at position 883. In both data sets, all base pairs that change do so only once. The intron sequence ranges from 1 - 584 of this aligned *trn*L-F data set. The changes are distributed fairly evenly between the intron and the IGS. There is a region between around position 450 and 700 where no variability occurs within the ingroup and relatively little variability between the in-and the outgroup. Generally, there is a fairly good correlation between the areas which are variable just within the ingroup, and those which vary with the outgroup species.

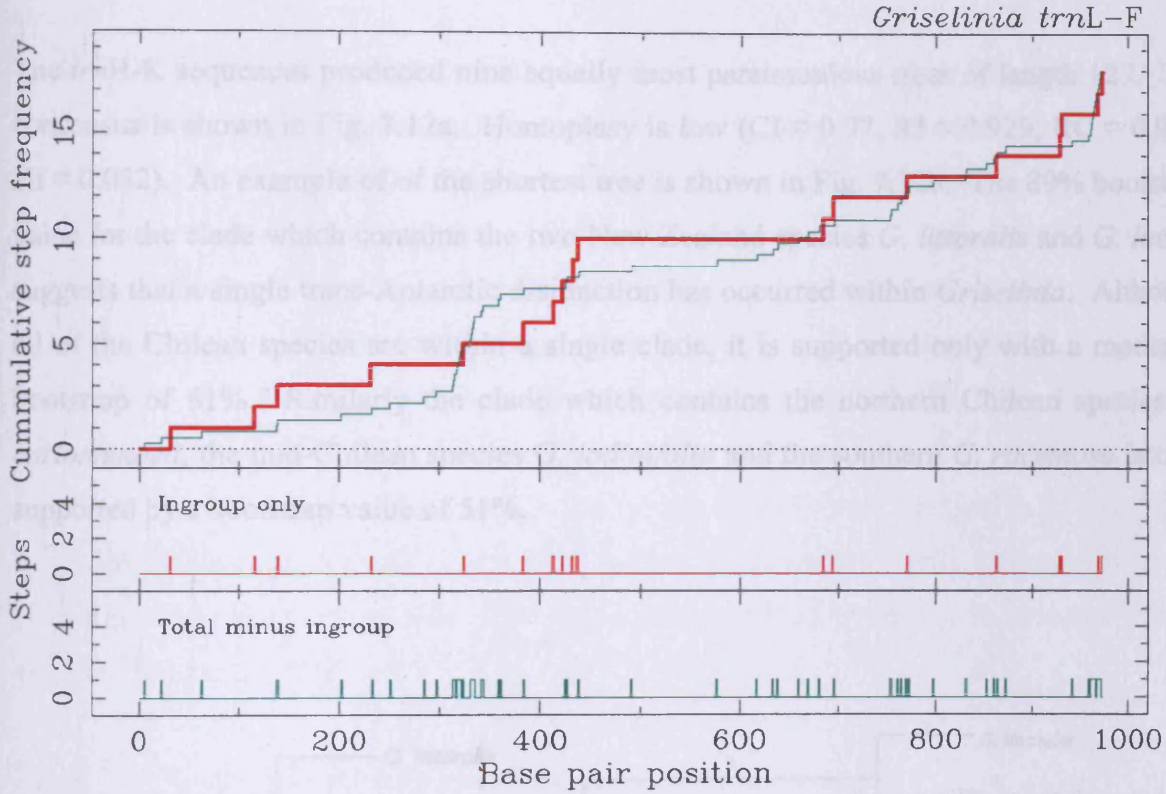

Fig. 7.11. Graph to illustrate the behaviour of *trn*L-F in *Griselinia* and its outgroup *Viburnum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Viburnum*) from the ingroup (*Griselinia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

*7.8.5 trn*H-K phylogeny

The *trn*H-K sequences produced nine equally most parsimonious trees of length 127. The consensus is shown in Fig. 7.12a. Homoplasy is low (CI = 0.97, RI = 0.929, RC = 0.900, HI = 0.032). An example of of the shortest tree is shown in Fig. 7.12b. The 89% bootstrap value for the clade which contains the two New Zealand species *G. littoralis* and *G. lucida* suggests that a single trans-Antarctic disjunction has occurred within *Griselinia*. Although all of the Chilean species are within a single clade, it is supported only with a moderate bootstrap of 61%. Similarly the clade which contains the northern Chilean species *G. carlomunozii*, the mid-Chilean species *G. jodinifolia* and the southern *G. racemosa* is only supported by a bootstrap value of 51%.

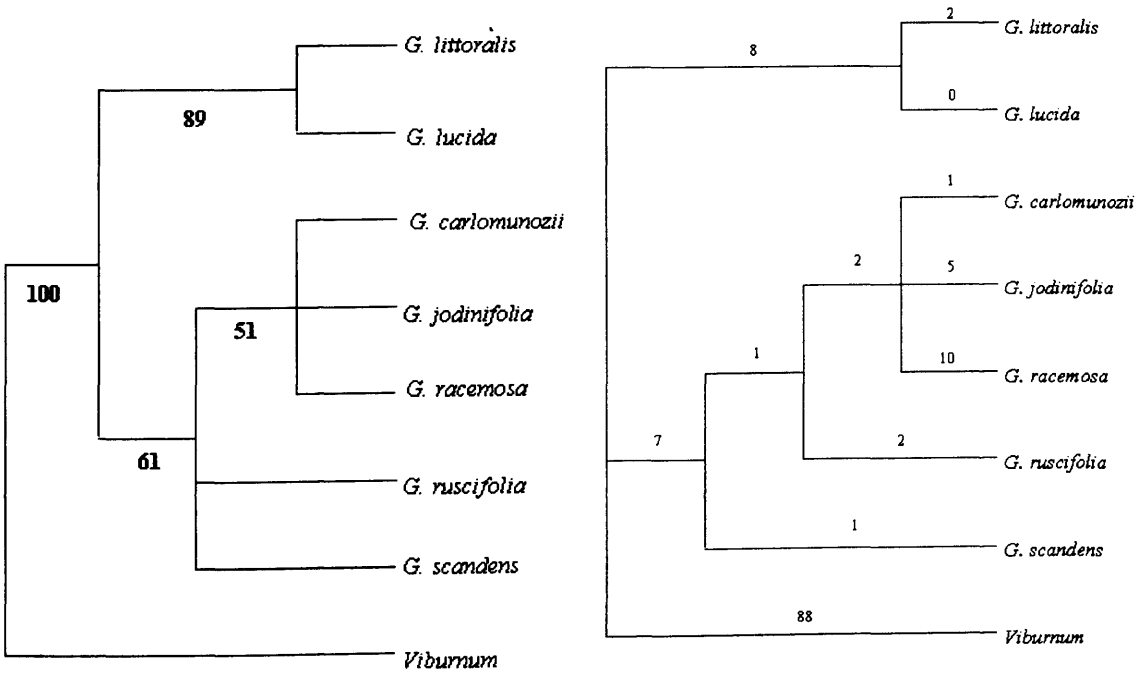

Fig. 7.12a. Consensus tree for *trnH-K* **b.** Example of shortest tree. For the consensus tree bootstrap values where they occur are below in bold. For the shortest tree the numbers of mutations are given above the clades.

7.8.6 TrnH-K variation

Statistics relating to the properties of the *trn*H-K data are given in Table 7.5. This is a very interesting and unusual data set in two ways. Firstly the transition to transversion ratio is a very low 0.37. Secondly there are a relatively high number of variable sites (3.67%). Unlike the situation in *Eucryphia*, a larger proportion of these sites are informative (1.14%). These observations are reflected in the higher number of steps per character (0.044). The sequence divergence ranges from 0.14-2.9%. The species which is the most different is *G. racemosa* which between base pair number 333 and 356 has 10 autapomorphies. All but two of the synapomorphies present in this data set are accounted for by the two New Zealand species. In the remaining two synapomorphies, a T is shared by the two Chilean, *G. ruscifolia* and *G. scandens* and an A is shared by *G. carlomunozii*, *G. jodinifolia* and *G. racemosa*. The GC content at 36.6 is typical for the chloroplast genome (Gillham, 1994).

Table 7.5. Statistics relating to *trn*H-K sequence data in *Griselinia*

Length range (bp)	702 – 708
Length mean (bp)	703.43
Number of characters (aligned length + indels)	708
Number of indels (informative)	2
Size of indels (bp) [uninformative]	1(6)
G+C content mean %	36.6
Sequence divergence %	0.14 – 2.90
Number of constant sites (%)	682 (96.33)
Number of autapomorphic sites (%)	16 (2.26)
Number of informative sites (%)	10 (1.41)
Transitions	7
Transversions	19
Percentage transitions	0.34
Average number of steps per characters	0.044

The pattern of variation is illustrated in Fig. 7.13. In the ingroup only data set, there are two main regions of variation; at 0 – 160 bp, and at around the 300 base pair position. In a very large stretch of the gene region, from around 400 base pairs until the end of the sequence, there is very little variation. This correlates well with the data set which contains both the ingroup and the outgroup as there is also very little variation in this region of the gene. Similarly, there is a significant amount of variation found and beginning of the gene region.

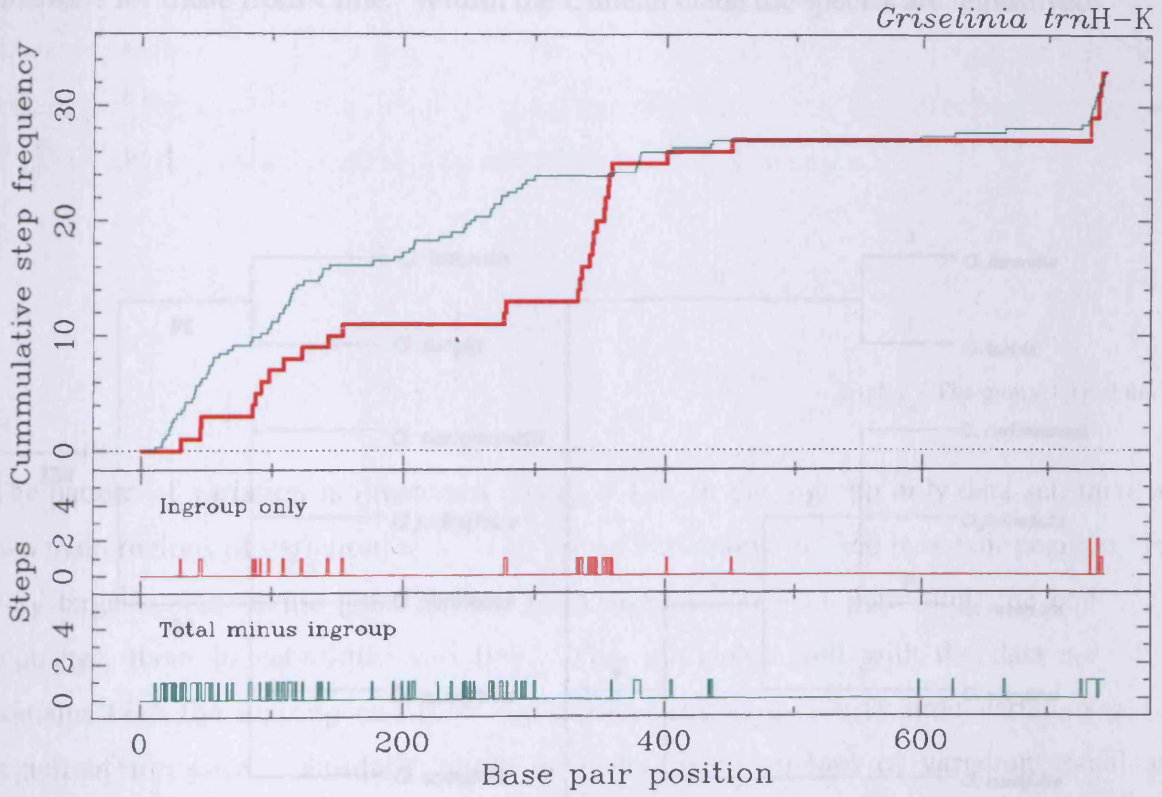

Fig. 7.13. Graph to illustrate the behaviour of *trn*H-K in *Griselinia* and its outgroup *Viburnum*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Viburnum*) from the ingroup (*Griselinia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

7.8.7 Combined rpoA, trnL-F and trnH-K phylogenies

A combined analysis of the three chloroplast data sets found fourteen most equally parsimonious trees of length 226. The consensus tree is shown in Fig. 7.12.a and an example of the shortest tree in Fig. 7.12b. It is worth mentioning that when these data sets are combined, there is more homoplasy than when they are treated individually (CI = 0.84, RI = 0.57, RC = 0.48, HI = 0.16). What is interesting however is the high bootstrap support for the two clades within *Griselinia*: 98% support for the New Zealand species and 82% for those from Chile. Within the Chilean clade the species are unresolved.

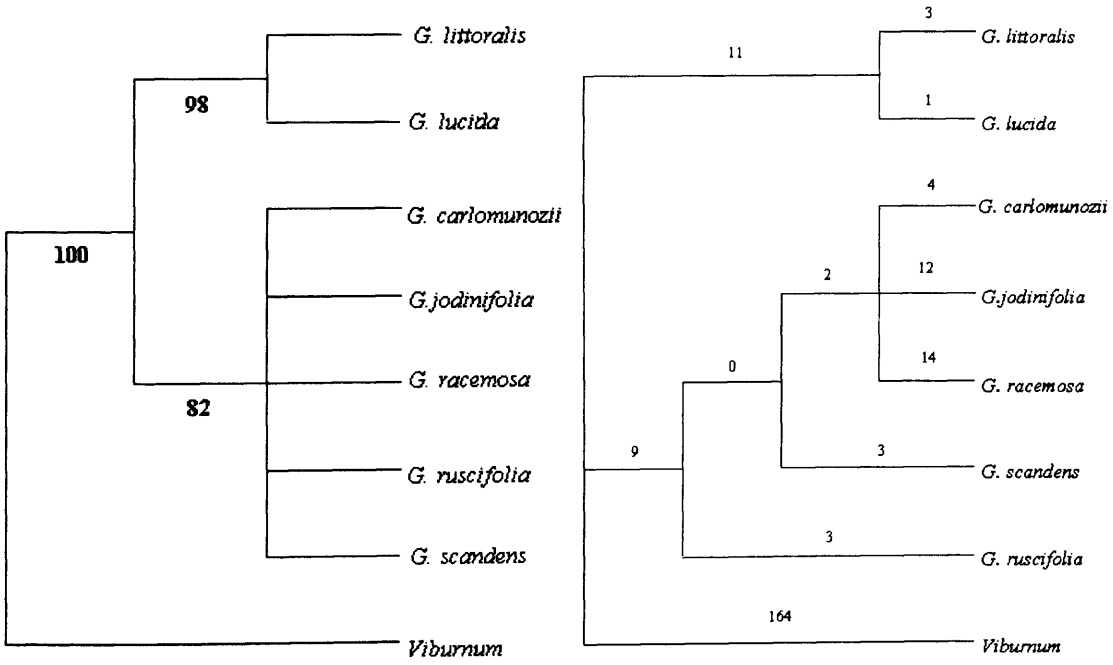

Fig. 7.14 a. Consensus tree for combined chloroplast data **b.** Example of shortest tree
For the consensus tree bootstrap values where they occur are below in bold. For the shortest tree the numbers of mutations are given above the clades.

7.8.8 ITS phylogeny

The ITS sequences produced a single most parsimonious tree of length 100 (Fig. 7.15). Although some homoplasy is present (CI = 0.83, RI = 0.65, RC = 0.54, HI = 0.017), the tree has a well resolved structure. Three well supported clades are present. The two most northerly Chilean species *G. carlomunozii* and *G. scandens* are in a clade which is supported by a bootstrap value of 95%. The two southerly Chilean species *G. racemosa* and *G. ruscifolia* are in a clade supported by a bootstrap value of 79%. The central Chilean species *G. jodinifolia* is a sister to the two New Zealand species, although bootstrap support is a fairly low 57%. The New Zealand species are united with bootstrap support of 96%. From the ITS phylogeny therefore, it is clear that despite morphological differences, *G. carlomunozii* is most closely related to *G. scandens*.

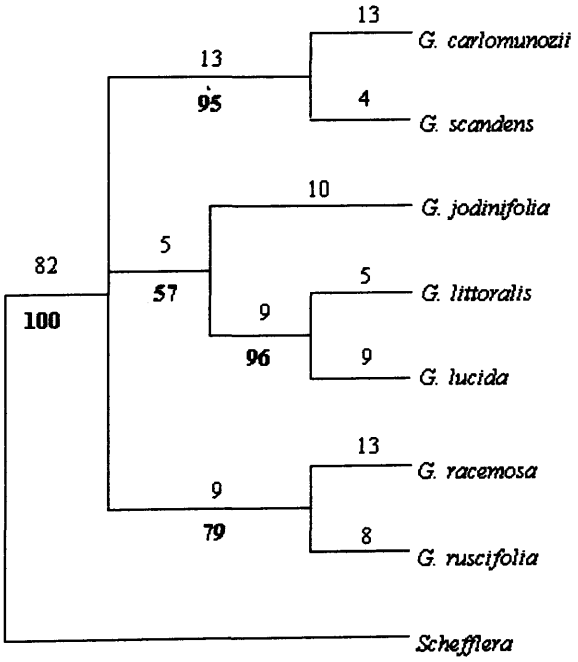

Fig. 7.15. *Griselinia* phylogeny for the ITS data set. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

7.8.9 *ITS variation*

Statistics relating to the properties of the ITS data are given in Table 7.6. In the ITS data set 728 sites (92.86%) are constant, and of the remaining variable sites 20 (2.55%) are autapomorphic and 36 (4.59%) are informative. This relatively large amount of variability (compared to the chloroplast gene regions) is reflected in the average number of steps per character of 0.125. The sequence divergence ranges from 1.52-4.86%.

Each of the seven insertions or deletions are only one base pair in length. The informative indels are at positions 429, 431, 495 and 612. An further observation in this data set is that the majority of the informative base pair mutations are transitions. The percentage transitions is a relatively high 73.21%.

Table 7.6. Statistics relating to ITS sequence data

Length range (bp)	782 – 786
Length mean (bp)	784.14
Number of characters (aligned length + indels)	789
Number of indels (informative)	7 (4)
Size of indels (bp) [uninformative]	All 1
G+C content mean %	60.5
Sequence divergence %	1.52 – 4.86
Number of constant sites (%)	728 (92.86)
Number of autapomorphic sites (%)	20 (2.55)
Number of informative sites (%)	36 (4.59)
Transitions	41
Transversions	15
Percentage transitions	73.21
Average number of steps per characters	0.125

The pattern of variation is illustrated in Fig. 7.16. This shows that only three of the 56 variable bases (7.14%) are found in the 5.8S rRNA gene. The main areas of the ITS region where variation occurs are at the beginning and immediately after the 5.8S gene. This can be seen clearly in the cumulative plot.

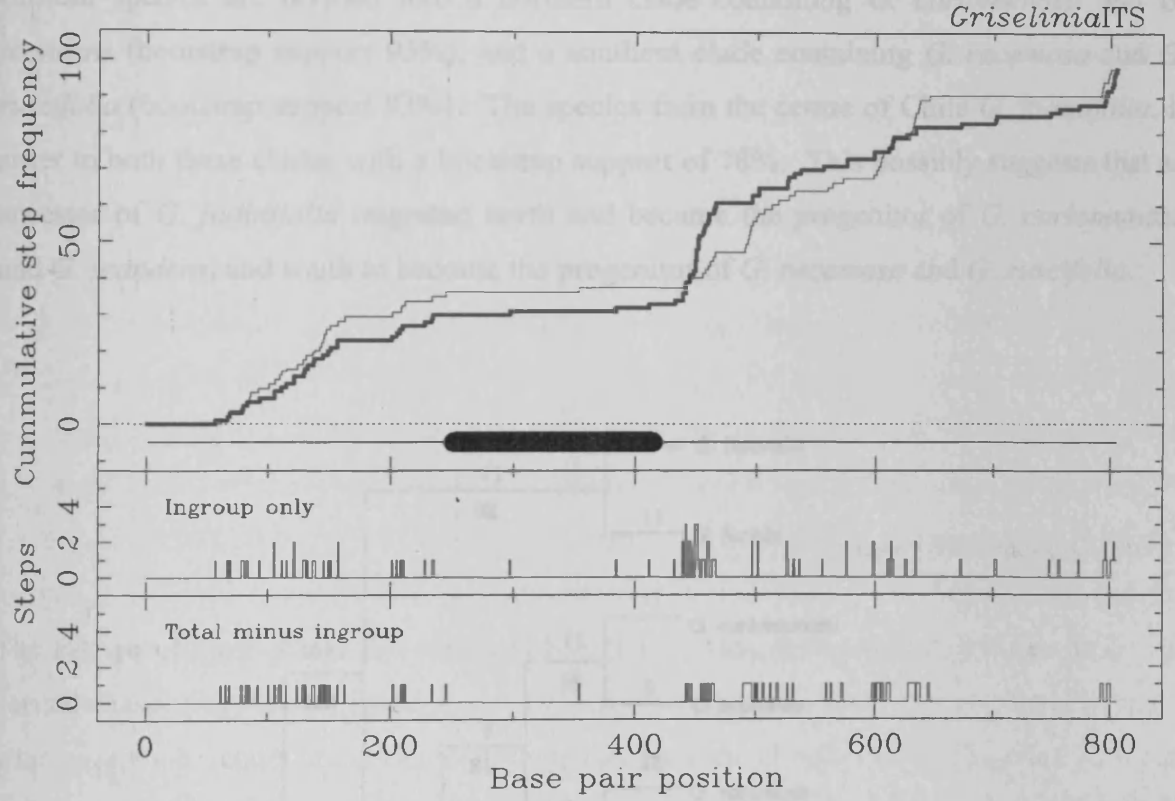

Fig. 7.16. Graph to illustrate the behaviour of ITS in *Griselinia* and its outgroup *Schefflera*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Viburnum*) from the ingroup (*Griselinia*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph. Bases in the bottom panel which are shown in blue are those that are invoked more than once in some of the most parsimonious solutions to the data set.

7.8.10 Combined rpoA, trnL–F, trnH-K and *ITS phylogenies*

The single tree obtained from a combined analysis of all of the data sets (Fig. 7.17) is revealing. Two major clades are resolved, one containing the two New Zealand species (bootstrap support 98%) and one with the Chilean species (bootstrap support 78%). The Chilean species are divided into a northern clade containing *G. carlomunozii* and *G. scandens* (bootstrap support 95%), and a southern clade containing *G. racemosa* and *G. ruscifolia* (bootstrap support 83%). The species from the centre of Chile *G. jodinifolia*, is sister to both these clades with a bootstrap support of 78%. This possibly suggests that an ancestor of *G. jodinifolia* migrated north and became the progenitor of *G. carlomunozii* and *G. scandens*, and south to become the progenitor of *G. racemosa* and *G. ruscifolia*.

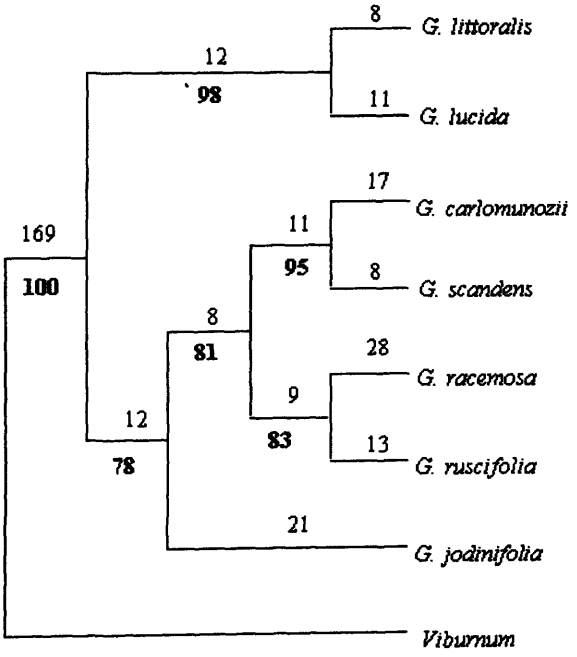

Fig. 7.17. *Griselinia* phylogeny for the combined *rpo*A, *trn*L-F, *trn*H-K and ITS data set. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

7.9 Sequence characteristics of the *Griselinia* data set

Table 7.7 shows the gene sequence characteristics for the *trn*L-F, *trn*H-K, *rpo*A and for ITS 1+2 regions for the *Griselinia* ingroup data. Parameters 1-3 give information about the length of the sequences. Within *Griselinia*, there is very little variation in sequence length. No data set has large insertions or deletions. Again, as there are relatively few insertions or deletions within the *Griselinia* genus data set, this figure is not substantially different from the mean length of sequence.

The next set of parameters shows the number and the percentage of the constant, autapomorphic and informative sites respectively. It is interesting to look at these data in relation to the amount of sequence divergence. We see that the least divergent *rpo*A data set has only one (0.23%) informative site. Although the *trn*H-K data set showed up to 3.67% sequence divergence, this is predominantly accounted for in autapomorphies although 1.41% of the characters are informative. The *trn*L-F data set has only 0.32% informative sites, however it is interesting to note that despite this, the phylogeny produced by these data has more resolution than the *trn*H-K data set. The ITS data set has the highest percentage of informative sites (4.59%) and correspondingly produces the phylogeny with the most structure.

The next set of parameters shows the number of transitions and transversions followed by the percentage of transitions. In the *rpo*A region, the percentage is 50%, but this is not as informative as it is for other gene regions where there is a greater total number of mutations. In the *trn*L-F region this percentage is 46.66% which reflects the fact that most of this region is not coding for sequence. The *trn*H-K has the lowest percentage, 26.92% which reflects the high number of transversions that have occurred. For ITS, the transition to transversion percentage is 73.21% which this time is probably due to the fact that much of the ITS region is coding, or has a use in its secondary structure, in other words has constraints preventing the conversion from a purine to a pyrimidine or vice versa.

The average number of steps per character is roughly similiar for the *trn*L-F and the *trn*H-K regions (0.018 and 0.044) whereas it is a factor of 10 less for the *rpo*A region (0.009) and a factor of ten more for the ITS region (0.125).

In terms of percentage sequence divergence, the least divergent sequence is *rpo*A which is identical in some species and at most only divergent by 0.70%. The sequence divergence range for *trn*H-K is wider than it is for *trn*L-F (0.14 - 2.90% compared to 0.22-0.97%). The ITS sequence is the most divergent with a range of between 1.52% and 4.86%.

The penultimate parameter shows the G+C content, for all the chloroplast regions, this figure stands at less than 40%, somewhere between 34.8% and 37.4%. This is typical for AT rich chloroplast sequence. For the ITS region the G+C content is 60.5%.

The final point concerns the number of most parsimonious trees. These are given both for the phylogenetic analysis of the *Griselina* data alone; the number is given in brackets when the outgroup is included.

Table 7.7. Sequence charateristics of *trn*L-F, *trn*H-K, *rpo*A, ITS 1+2, for the seven species of *Griselinia*.

Parameter	*rpo*A	*trn* L-F	*trn* H-K	ITS 1 + 2
Length range (bp)	429	913-928	702 - 708	782 - 786
Length mean (bp)	429	923.14	703.43	784.14
No. characters (aligned length + indels)	429	931	708	789
No. of indels (informative)	None	2 (2)	2	7 (4)
Size of indels, (bp)	N/A	15 bp 1 bp	1 bp	All 1bp
informative				
[uninformative]				
Number of constant sites (%)	426 (99.08)	908 (98.37)	682 (96.33)	728 (92.86)
No. of autapomorphic sites (%)	3 (0.69)	12 (1.30)	16 (2.26)	20 (2.55)
No. of informative sites (%)	1 (0.23)	3 (0.32)	10 (1.41)	36 (4.59)
Transitions	2	7	7	41
Transversions	2	8	19	15
Percentage transitions	50	46.66	26.92	73.21
Average no. of steps per character	0.009	0.018	0.044	0.125
Seqence divergence (%)	0 – 0.70	0.216 – 0.97	0.14-2.90	1.52 – 4.86
G+C content mean (%)	37.4	34.8	33.6	60.5
Number of trees found	1 (1)	3 (1)	9 (3)	5 (1)

7.10 Molecular clocks: Dating the phylogenetic split

7.10.1 Relative rate tests

Tajima statistics were calculated for each of the sequences studied as described for *Eucryphia*. This statistic has a χ^2 distribution whose critical value at 1 degree of freedom and 99% significance level is 6.63. Tajima statistics are shown for the *rpo*A (Table 7.8), *trn*L-F (Table 7.9), *trn*H-K (Table 7.10) and ITS (Table 7.11). In all tables Carl = *G. carlumonozii*, Jod = *G. jodinifolia*, Lit = *G. littoralis*, Luc = *G. lucida*, Rac = *G. racemosa*, Rusc = *G. ruscifolia*, Scan = *G. scandens*.

Table 7.8. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Griselinia rpoA*.

	Carl	Jod	Lit	Luc	Rac	Rusc	Scan
Carl	X	2,0	2,1	2,1	2,1	2,0	2,0
Jod	2	X	0,1	0,1	0,1	0,0	0,0
Lit	0.333	1	X	0,0	1,1	1,0	1,0
Luc	0.333	1	1	X	1,1	1,0	1,0
Rac	0.333	1	0	0	X	1,0	1,0
Rusc	2	0	1	1	1	X	0,0
Scan	2	0	1	1	1	1	X

All of the values in this table pass the test at the 99% level, on these grounds therefore the molecular clock can be applied.

Table 7.9. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Griselinia trn*L-F.

	Carl	Jod	Lit	Luc	Rac	Rusc	Scan
Carl	X	3,4	4,3	4,3	1,2	1,1	1,1
Jod	0.143	X	6,3	6,3	5,4	5,4	5,3
Lit	0.143	1	X	1,1	3,5	3,4	3,4
Luc	0.143	1	0	X	3,5	3,4	3,4
Rac	0.333	0.111	0.5	0.5	X	2,1	2,1
Rusc	0	0.5	0.143	0.143	0.333	X	1,1
Scan	0	0.5	0.143	0.143	0.333	0	X

All of the values in this table pass the test at the 99% level, on these grounds therefore the molecular clock can be applied.

Table 7.10. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Griselinia trn*H-K.

	Carl	Jod	Lit	Luc	Rac	Rusc	Scan
Carl	X	4,2	3,7	4,6	1,11	1,1	1,0
Jod	0.667	X	4,10	5,9	2,14	2,4	3,4
Lit	1.6	2.571	X	2,0	7,13	6,3	7,3
Luc	0.4	1.423	2	X	6,14	5,4	6,4
Rac	8.3	9.0	1.8	3.2	X	11,1	12,1
Rusc	0	0.667	1	0.111	8.333	X	2,1
Scan	1	0.143	1.6	0.40	9.307	0.333	X

In all cases except relative to the two New Zealand species, *G. littoralis* and *G. lucida*, *G. racemosa* fails the test at the 99% level. All other species pairs pass.

Table 7.11. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for *Griselinia* ITS.

	Carl	Jod	Lit	Luc	Rac	Rusc	Scan
Carl	X	13,8	18,8	17,12	16,9	14,7	8,2
Jod	1.191	X	11,6	10,9	11,8	8,6	9,9
Lit	3.846	1.471	X	4,8	8,11	8,12	10,15
Luc	0.862	0.053	1.333	X	11,10	11,11	13,14
Rac	1.196	0.474	0.474	0.474	X	4,4	10,11
Rusc	2.333	0.286	0.080	0.0	0.0	X	9,10
Scan	3.600	0.0	1.0	0.037	0.0476	0.053	X

All of the values in this table pass the test at the 99% level, on these grounds therefore the molecular clock can be applied.

7.10.2 Calibrating the clock

The clock cannot be calibrated using fossil data as no fossils have been recorded in the literature for *Griselinia*. Extra efforts were also made to find fossil information, by asking paleontologists working in this field, but to no avail.

As the gene regions seemed to evolve at similar rates to *Eucryphia*, and furthermore the two genera are both woody species with similar life histories, the individual gene clocks for *Eucryphia* were employed to estimate the ages of *Griselinia* speciation. The calibrations based on the 60 million year and the 2.6 million year fossils were used. No clock has been applied to the *rpo*A region as it could not be calculated from the *Eucryphia* data; in addition, it is similarly invariable in *Griselinia*. The ITS clock is based on average rate of ITS evolution as shown in work by Richardson *et al.* (2001).

7.10.3 Application of the molecular clock

Table 7.12. Molecular clock calibration table for three gene regions.

Gene region	*trn*H-K	*trn*H-K	*Trn*L-F	*trn*L-F	ITS
Fossil age (MYO)	60	2.6	60	2.6	Ave. clock
Length	708	708	931	931	792
Rate (K) $x10^{-4}$	0.257	0.379	1.53	2.08	2.44
Min K.	0.246	0.158	1.35	0	
Max K.	0.280	0.788	1.71	0.416	

Table 7.12 shows the average, minimum and maximum rate and time of evolution that has occurred over the three different gene regions that passed the rate constancy test in *Griselinia*. The average number of mutations between the South American and the Australasian species is shown in Table 7.13. The divergence times based on the *Eucryphia/* general ITS clock are then given.

Table 7.12. Divergence times for the South American-Australasian disjunction, for the three gene regions.

Gene region	*trn*H-K	*trn*H-K	*trn*L-F	*trn*L-F	ITS
Fossil age (MYO)	60	2.6	60	2.6	Ave. clock
Av. no. of mutations between split	15.2	15.2	7.6	7.6	39.8
Av. divergence time (MYO)	83.54	56.65	53.35	39.25	20.60
Min. divergence time (MYO)	76.68	27.25	47.74	19.62	
Max. divergence time (MYO)	87.27	135.89	60.46	-	

New Zealand became separated from Antarctica between 95 and 82 million years ago and therefore after this time it was no longer associated with South America. In order for the distribution of *Griselinia* to be explained in terms of vicariance, it is necessary to have at least 80 million years worth of divergence. Although these estimated chloroplast clocks based on *Eucryphia* are an approximation of the amount of divergence time, only in one

case, that of the 60 million old fossil and the *trn*H-K sequence, is there enough time to explain the split in terms of vicariance. The average divergence time for the South American-Australasian split for the five clocks is 51 million years. This is approximately half the amount of time needed to explain the split in terms of vicariance. Nevertheless, the estimates of divergence times of the maximum estimates exceed the critical date. It appears therefore that there was a dispersal event between the two continents.

7.11 Conclusions

Having discussed the molecular data in the context of morphological data and having estimated divergence times there are a number of conclusions that can be made. It seems likely that the two species from New Zealand *G. littoralis* and *G. lucida* are sister species. Similarly, the five species from Chile form a well supported monophyletic group. Within this group the two southern hemisphere species *G. racemosa* and *G. ruscifolia* are closely related to each other and thus form a discreet group and similarly the two species from the north *G. carlomunozii* and *G. scandens* form a group. A sister group to these species is the central Chilean species *G. jodinifolia.*

It is not possible from these data to designate the directionality of a dispersal event. However, evidence from outgroup genera being more prevalent in Australasia, suggests that the dispersal event was from Australia to South America. This fits with Dillon's views (pers. comm.) which are based on observations of the morphology. One scenario that would be in accordance with these data is that an ancestor of the species *G. jodinifolia* arrived at the centre of Chile then radiated both north and south where two species of each exist today.

The most likely explanation for the present-day distribution of *Griselinia* is therefore dispersal from Australasia to South America. The method of dispersal could either have been via birds, or, similar to the case in *Eucryphia*, the seeds could have become caught up in debris and rafted from one continent to the other.

CHAPTER 8

THE GENUS CORIARIA

8.1 Introduction

Coriaria is a small genus of about 15 species (Table 8.1). They are small trees, shrubs or sub-shrubs with quadrangular branchlets and simple, opposite or whorled, palmately-veined leaves which lack stipules. Small greenish flowers appear in axillary or terminal racemes. A feature unique to *Coriaria* is that after the ovules have been fertilized, the petals swell and become fleshy to form a dark, purple-black pseudo-fruit that encloses 5-10 achenes

Coriaria species are distributed near the coast around the Mediterranean sea, in eastern Asia, South and Central America and in New Zealand (Good, 1930). In temperate places *Coriaria* is found at low altitudes whereas in the tropics it is generally found at higher elevations. The root nodules of *Coriaria* species contain nitrogen fixing actinomycete from the genus *Frankia* (Daly, 1967). *Coriaria* are successful in establishing themselves in disturbed places, especially on volcanic soils where they are often one of the first plants to be established following eruptions (Wiedman and Weberling, 1993).

The general morphology is illustrated in Fig. 8.1. Work on their vegetative anatomy has been done by Metcalfe and Chalk (1950, 1979); Carlquist (1985); Suzuki and Yoda (1986a,b) and Yoda and Suzuki (1992). In addition, the floral morphology, anatomy and embryology in *C. napalensis* was studied by Sharma (1968) . Further studies on the pericarp anatomy were performed by Tobe *et al.* (1992) The carpel peltation and syncarpy in *C. ruscifolia* was described by Guédes (1971). The pollen has been studied by Praglowski (1970) and, in *C. napalensis* only, by Garg (1980). Cytological work was done by Oginuma *et al.* (1991). A recent molecular survey on two chloroplast genes was carried out by Yokoyama *et al.* (1998).

Fig. 8.1. The shape, and arrangement, of leaf and flower in the southern Hemisphere *Coriaria ruscifolia* . (From Hoffmann, 1997).

A list of the currently accepted names and a synonymy is provided in Table 8.1, data for which were drawn from the work by Cockayne and Allan (1927), Allan (1927), Good (1930), Oliver (1942), Skog (1972) and Thompson (1996).

Table 8.1. List of currently accepted names and synonyms in *Coriaria*.

Coriaria Linnaeus, Sp. Pl., 1037 (1753).

Southern Hemisphere species

New World species

C. microphylla Poiret, Encyc. 6: 87 (1801).
C. ruscifolia ssp. *microphylla* (Poiret) Skog, Rhodora 74: 249. (1972).
C. thymifolia Humboldt & Bonpland *ex* Willdenow, Sp. Pl. 4(2): 819 (1806).
C. phylicifolia Humboldt & Bonpland *ex* Willdenow, Sp. Pl. 4 (2): 819 (1806).
C. atropurpurea Moçino & Sessé *ex* De Candole, Prodr. 1: 740 (1824).
Heterocladus caracasanus Turczaninow, Bull. Soc. Imp. Nat. Moscou 20 (1): 152 (1847).
Heterophylleia caracasanus Bull Soc. Imp. Nat. Moscou 21 (1): 591 (1848).
Vernacular names: Shanshi, piñan, tinta (Ecuador), rebentadera (Columbia), mocotinto, moco de chompipe (Guatemala), tlalocopetate, tlalocopatlatl, helecho *C.* de tierra (Mexico), mio-mio, saca-saca, mio-venenosa, raqui-raqui (Peru), tisis, helecho de playa, helecho noite (Venezuela).

C. ruscifolia Linnaeus, Sp. Pl. 1037 (1753).

New Zealand and Pacific Island species

C. angustissima J.D. Hook., Handb. N.Z. Fl., 47 (1864).

C. arborea Lindsay, Contr. N.Z. Bot. 84 (1868).
C. arborea var. *kermadecensis* W.R.B. Oliver, Rec. Dom. Mus. Wellington 1: 34 (1942).
Vernacular names: Wasalele, mariko ni tambale (Fiji), dedu hique, huiqui, matarrotones (Chile), tutu, tutapike (New Zealand)

C. kingiana Colenso, London J. Bot. 3: 20-21 (1844).
C. thymifolia var. *undulata* Petrie, Trans. N.Z. Inst. 53: 386 (1921).

C. lurida var. *undulata* (Petrie) Allan, Rep. Austr. Assoc. Adv. Sci. 20: 446, fig. 3 (1931).

C. plumosa W.R.B. Oliver, Rec. Dom. Mus. Wellington 1: 27.

C. pottsiana W.R.B. Oliver, Rec. Dom. Mus. Wellington 1: 24, pl. 8, fig. 1 (1942).

C. pteridoides W.R.B. Oliver, Rec. Dom. Mus. Wellington 1: 30, pl. 10 fig. 1 (1942).

C. thymifolia sensu J. D. Hooker, Handb. N.Z. Fl. 47 (1864).

C. lurida var. *acuminata* Cockayne & Allan, Trans. N.Z. Inst. 57, 51 – 52 (1927).

C. sarmentosa G. Forster, Pl. Esc. 46 (August –September 1786); Fl. Ins. Austr., 71, n. 377.

C. Tutu Lindsay, Cont. N.Z. Bot. 84 (1868).

C. papuana Warburg, Bot. Jahrb. 16 (1): 22 (1892).

Vernacular names. Umiaukawe, ikapeh (New Guinea).

Northern Hemisphere species

C. japonica A. Gray, Mem. Amer. Acad. Sci., n.s., 6: 383 (1857).

Vernacular name: Doku-utsugi

C. intermedia Matsum.

C. myrtifolia Linnaeus, Sp. Pl., 1037 (1753).

C. hermaphrodita Turra, Farset. Nov. Gen. 13 (1765).

? *C. procumbens* Whitl. *ex* Hoffmannsegg, Verz. Pfl. Nachtr. 3: 9 (1824), *nomen nudum.*

C. tinctoria Grenier & Godron *ex* Dulac, Fl. Hautes-Pyr. 243 (1867).

Vernacular names: Coroyère, redoul (France), emborrachacabras, roldón (Spain).

C. nepalensis Wallich, Pl. As. Rar. 3: 67, pl. 289 (1832).

C. sinica Maximowicz, Mem. Acad. Petersburg, ser. 7, 29 (3): 9. (1881).

C. kweichovensis Hu, Bull. Fan Mem. Inst. Biol. Peiping Bot. Ser. 7: 213 – 214.

Vernacular names: Ma-sang –(Hubeh), limphu shi, nimbo (Bhutan).

C. terminalis Hemsley, Hook. Icon. Pl. 23: pl. 2220 (1894).

It is important to point out that species delimination in *Coriaria* is highly contentious and estimates have ranged from 5 – 20 species depending on the authors. Authors recognising five species are Skog (1972), Cronquist (1981) and Thorne (1983). Ten species were recognised by Sholz (1964). Shaw (1973) recognised 15 species and Good (1930) recognised 20. I have generally followed the system adopted in the most recent revision of the genus by Thompson (1996), although I have included *C. arborea* and *C. intermedia* as species separate from *C. ruscifolia* and *C. japonica* respectively (Table 8.1).

The distributions of the species are discussed in section 8.4. There are also numerous *Coriaria* hybrids; these are not directly relevant to this study so are not discussed here. Details can be found in Thompson (1996) and references therein.

The genus was first described in 1753 by Linnaeus who described the species *C. myrtifolia* which was found near Montpellier and *C. ruscifolia* which was collected in Chile. Due to the protogyny and the significantly different flowering times for male and hermaphrodite flowers, he placed *Coriaria* in his Dioecia Decandria class. This assumption of dioecy was accepted for many years although Turra (1765) did recognise that *C. myrtifolia* had hermaphrodite flowers.

Poiret (1804) listed four species of *Coriaria* and he gave the name *C. microphylla* to the small-leaved species from Peru. He assigned the name *C. ruscifolia* to those plants from Chile which had much larger leaves than *C. myrtifolia*. He described *C. myrtifolia* in more detail than Linnaeus and also listed *C. sarmentosa* which had been described from New Zealand by Forster in 1786.

The Asiatic species *C. napalensis* was described by Wallich (1832) although he considered the flowers to be hermaphroditic. The monoecious nature of *C. japonica* was first noted by Gray (1857). Helmsley (1894) was the first person to formally describe *C. terminalis*. The first person to review the genus on a worldwide basis was Maximowicz (1881). He gave detailed descriptions of all species mentioned so far and included *C. sinica* from

China plus two species from New Zealand which were unknown to him. He observed the fact that the northern Hemisphere species have both male and hermaphrodite or female flowers and that the male flowers appear before the leaves. He also noted that the flowers from the southern Hemisphere species are always hermaphrodite. Furthermore he recognised the differences in that some species flower from old wood whereas others do so from new wood. Warburg (1892) described a new species from New Guinea which he named *C. papuana*.

Hooker (1864) was the first person to recognise a similarity between species on both sides of the Pacific. He felt that one of the *Coriaria* species from New Zealand was the same as the *Coriaria* found in South America. The variability of *Coriaria* in New Zealand was clarified in a review by Oliver (1942) in which he recognised seven variants as species: *C. angustissima*, *C. arborea*, *C. kingiana*, *C. plumosa*, *C. pottsiana*, *C. pteridoides* and *C. sarmentosa*. Allan (1961) agreed with these species designations.

8.2 Relationships to other families and genera

Coriaria is the only genus in the family Coriariaceae. The relationships between it and other families are unclear. Many different affinities have been proposed. Hooker (1853) considered *Coriaria* to be related to the Rutaceae but felt that it could be nearer either the Chenopodiaceae or the Phytolaccaceae. Maximowicz (1881) believed that the closest relatives were Simaroubaceae and Anacardiaceae, or possibly Phytolaccaceae. The proximity to a sapindalean-rutalean group was also favoured by Scholz (1964), Takhtajan (1980), Dahlgren (1983) and Thorne (1983). This viewpoint was supported by work on pollen structure (Garg, 1981), and on chemical compounds (Bohm and Ornduff, 1981).

Croizat (1952) and Cronquist (1981, 1988) disagreed with the above assessments and they tentatively put *Coriaria* in the Ranunculales. They felt that the floral characters and stem anatomy were more similar to the Ranunculales than they were to any other order. This was supported by work on the seed coat structures of *Coriaria* by Corner (1976). Protein data suggested a different set of relationships, it seems that *Coriaria* is in a clade with Leguminoseae, Rosaceae, Saxifagaceae and Proteaceae (Martin and Dowd, 1993).

The uncertainty regarding close relatives is highlighted by the fact that Coriariaceae has been placed in six different orders: Ranunculales, Rutales, Sapindales, Terebinthales, Coriariales, Rosales and Celastrales (Goldberg, 1986; Bohm and Ornduff, 1981).

Molecular data from the chloroplast *rbc*L gene disagreed completely with previous affinities and assigned *Coriaria* to a Hamamelid clade as a sister group of *Datisca*, *Begonia* and *Cucurbita* (Chase *et al.* 1993; Swensen *et al.* 1994; Swensen 1996). Most recently the genus *Corynocarpus* has been identified as the sister group to *Coriaria* within this Cucurbitalean alliance on the basis of *rbc*L sequences (Wagstaff and Dawson, 2000).

8.3 Reproductive biology

The flowers of *Coriaria* are small and pentamerous, usually in racemes which possess few to many flowers. The sepals are imbricate and persistent, and the petals are keeled on the inside. The 10 stamens have filaments (mostly 1.5 – 5 mm long) which are either free, or those opposed to the petals are adnate. The pollen grains are small and spherical (17-30 μm). The tectum has numerous small projections (Praglowski, 1970). There are 5- 10 carpels, depending on the species, which are free except at their base where they are fused. They contain a single pendulous ovule and terminate in elongated, linear styles.

The species are self-compatible, and exhibit both inter and intra-floral dichogamy (Thompson & Gornall, 1995). The Himalayan *C. terminalis*, as well as most species of the southern Hemisphere, (*C. arborea, C. kingiana, C. microphylla, C. papuana, C. pottsiana, C. pteridoides C. ruscifolia* and *C. sarmentosa*) are evergreen phanerophytes. *C. angustissima* and *C. plumosa* however are high elevation hemicryptophytes. The flowers of these latter species are protogynous and hermaphroditic and are arranged in racemes either terminal or more usually in pairs in the axils at each node. All of these species flower on stems produced in the current season of growth.

In contrast, species from the northern Hemisphere (except *C. terminalis*, see above) are deciduous phanerophytes which flower on old wood. In this case the racemes have few flowers and are aggregated in the leaf axils in clusters of up to 12. Some species are andromonoecious and others are monoecious; both show interfloral protandry and the former display intrafloral protogyny (Thompson & Gornall, 1995).

The evolution of these northern characters, such as the deciduous habit, inter-floral protandry, and the aggregation of racemes, is thought to be related to the selection pressures in a seasonal climate (Thompson and Gornall, 1995).

The flowers are scentless and in field studies no insects were observed visiting the flowers (Thompson and Gornall, 1995). These observations, together with the fact that the stamens and styles are exserted from the flower, dangling free in the air, suggests that the flowers are wind-pollinated.

The fruits are achenes but these are surrounded when ripe, by the enlarged succulent petals. The seeds are compressed with straight embryos and a thin endosperm.
In New Zealand it is known that the fruit pigeon (*Carphphaga novae-zealandiae*) eats the fruits (Thompson *et. al.* 1990), and it is therefore likely that other birds are similiarly attracted to them.

8.4 Distribution and Ecology

The present day distribution of *Coriaria* is shown in Fig. 8.2. The distribution is between 45° N and 45°S, and between 0° and 110° E. Apart from in the Himalayas and in China, the species localities are no further than 20 km from the coast. The species are allopatric except in the Himalayas, where *C. napalensis* and *C. terminalis* are found in the same regions, but possibly not at the same altitudes (Thompson, 1996), and in New Zealand where six species occur, but again with some altitudinal differences. Of the 15 species of *Coriaria* accepted in this work, ten are found in the southern Hemisphere and five in the northern Hemisphere.

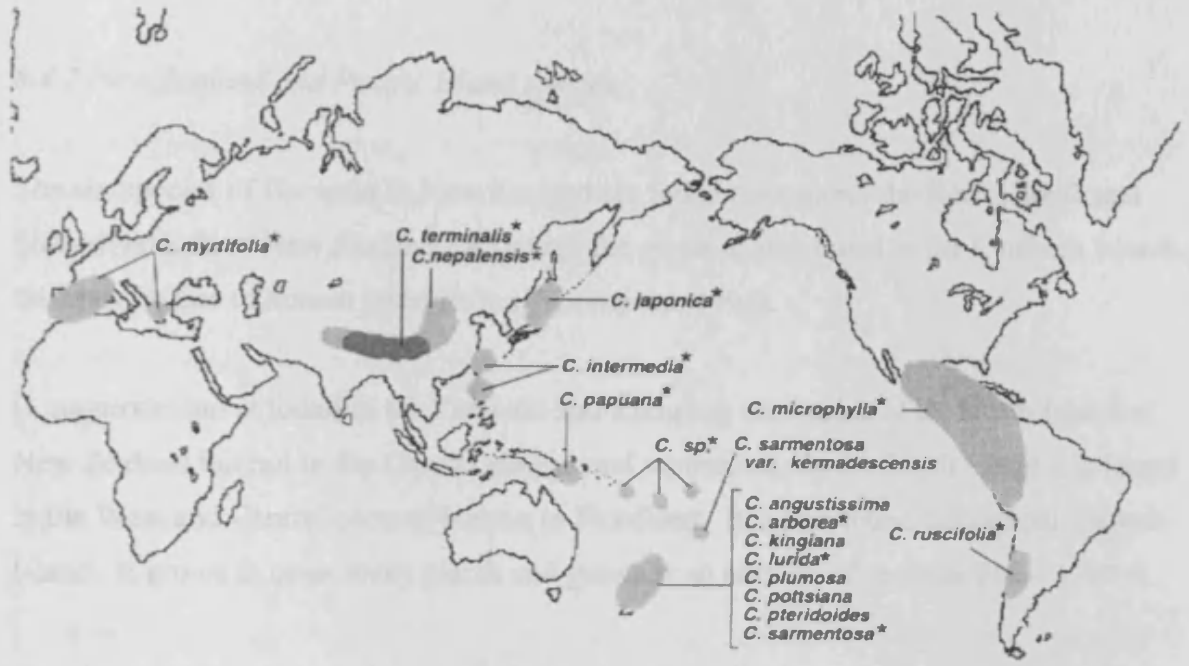

Fig. 8.2. Distribution of *Coriaria*. From Yokoyama *et al.* (1998). The asterisks indicate the speices used in their study. Under the species delimitation described here (section 8.1) *C. lurida* is considered to be a subspecies of *C. sarmentosa*.

8.4.1 South American species

C. microphylla is found north of the Atacama desert in Chile. It is found on the eastern range of the Andes in Peru, and on both the western and eastern ranges in Ecuador and Colombia and is also found in Merida in Venezuela. There is then the Darien Gap before *C. microphylla* again reappears in the volcanic areas of western Panama and in similar habitats of central Costa Rica. It is then found in the Southern and Western Highlands of Guatemala, and then in the southern Mexican state of Chiapas. There is then another gap in the distribution before it reappears in the Central Volcanic Valley, there is then a final gap before it appears at its northerly limit around Durango.

C. ruscifolia is found in Chile where it has a patchy distribution throughout its range which is generally around the Central Valley. It does cross the Andes at one point where it spreads into the province of Neuquén in Argentina.

In terms of habitat, It is generally found by the sides of streams or by damp gorges, on loose roadside banks and on new lava flows. The altitudinal range in temperate areas is from sea level to 1200 m and in the tropics it is found at the height of 2100 m.

8.4.2 New Zealand and Pacific Island species

The six species of *Coriaria* in New Zealand are found throughout the North, South and Stewart Islands of New Zealand. Although the genus is also found in the Chatham Islands, this may be due to human intervention (Thompson, 1996).

C. angustissima is found in the Taranaki and Kikurang mountains of the North Island of New Zealand but not in the Central plateau and mountains. In the South Island it is found in the West and Central parts of Nelson to Fiordland. It is also found throughout Stewart Island. It grows in open rocky places and grows at an altitude of between 500 – 1700 m.

C. arborea is found throughout both the north and south islands from coastal regions to montane forests mainly in the forest margins, and also in shrublands on alluvial ground.

C. kingiana is found in the Central plateau and inland from Poverty Bay to Hawke Bay in the North Island of New Zealand and in the South Island it is found in Nelson and Otago. It grows by the roadside banks and in open scrub and between 50 - 700m altitude.

C. plumosa grows on the North Island of New Zealand from Mount Egmont and Mount Hikurang southwards. It is also found throughout the South Island where it grows on lowland to lower subalpine ground, grassland, stream-beds and on lower mountain slopes.

C. pottsiana is found in the screes on Mount Hikurangi and at high altitude environments around an altitude of 1000m.

C. pteridoides is found in the North Island of New Zealand around Mt. Taranaki and the Central volcanic plateau and mountains. In the South Island it is found near Nelson, Canterbury, Otago and Westland. It is found from 500-1300 m and grows on banks and river banks.

C. sarmentosa is found in the North Island of New Zealand along the coast of Cook Strait. It is also found throughout the South Island and Stewart Island. It grows in open places and in low scrub land, from sea level to 1000 m.

C. papuana is found in Papua New Guinea where it is confined to New Guinea on the mainland in the Finisterre and Bismarck mountain ranges. It is also found in New Ireland. It grows on mountain slopes and on areas where there have been landslides and it also grows in open areas of forest. Its altitudinal range is between 600-3000 m.

8.4.3 Northern Hemisphere species

C. intermedia is found throughout the mountains of Taiwan and the Philippines where it is particularly found in Benguet Province.

C. japonica is found in the central area of Japan, around Honshu and around Hokkaido. It grows on the sides of mountains and in stony areas, particularly by water courses and low scrub. Its altitudinal range is from 500 to 3400 m.

C. myrtifolia is found along most of the Mediterranean coast of Spain. Its range extends disjunctively into France, Italy, Greece, Morocco and Algeria. It grows on roadside banks and on loose mountains sides, it also favours the cooler, damper faces of riverbanks. It is found at an altitude of between sea level and 1500 m.

C. napalensis is found in India were it grows in the Himalayas eastwards from the Kashmir border. It is also found in Nepal, Sikkim, Bhutan, the Burma-Tibet Triangle and China in the provinces of Yunan, Kwangsi, Kweichow, Sichwan, Shaanxi and Hubeh. It is found at an altitudinal range of between 1400-3200 m where it grows in open woodland and scrub and beside gorges and streams.

C. terminalis is found in the east of Nepal, Sikkim, Bhutan, in the Burma-Tibet Triangle, China (in the north-west of Yunan and in the West of Sichwan. It grows on open hillsides and in forest clearings and scrub at altitude from 2000-4000 m.

8.5 Fossil record of *Coriaria*

Although many fossils have been reported for *Coriaria*, the majority of them are difficult to relate to extant species because they are either from detached leaves (Ozaki, 1991), fossil pollen (Praglowski, 1970; Muller, 1981), or seeds (Gregor, 1980). An exception to this is a fossil shoot containing inflorescences and leaves named *C. longaeba* (Saporta, 1865). This fossil was found in the Armissan beds of the Oligocene age, about 35 million years ago. The inflorescence has a long terminal raceme which is similar to that of *C. terminalis*.

The only other fossils found and definitely attributed to *Coriaria* thus far are those of pollen. Definitive *Coriaria* pollen has been found from the Miocene (between 5 and 23 MYO) deposits of Spain (Van Campo, 1976). It has also been found in 11 million year old deposits of New Zealand (Couper, 1953, 1960). Fossils have also been found from deposits of the Pliocene of New Zealand (between 5 and 1.6 MYO; Mildenhall, 1980). The most recent fossils of *Coriaria* pollen were found albeit sparsely in the New Zealand deposits of the Pleistocene (1.6 MYO; Horrocks and Ogden, 1994; Lees, 1986).

Fossils therefore which will be used to date the nodes on the phylogram are the 35 million year old *C. longaeba* and the 11 million year old pollen from New Zealand.

8.6 Relationships within *Coriaria*

8.6.1. Morphological evidence

In his study of *Coriaria*, Good (1930) recognised three groups within the genus. These were based on geographic distribution, flower type and sexuality, and features of the winter bud scales. His first group, Group A, contained only *C. terminalis* which unusually has terminal inflorescences as opposed to the lateral and axillary inflorescences which all the other species of *Coriaria* possess. His second group, Group B, consisted of all other Eurasian species apart from *C. terminalis*, i.e. *C. intermedia*, *C. japonica*, *C. myrtifolia* and *C. napalensis*. He included all of the other species from the southern Hemisphere in group C.

8.6.2. Flavonoid evidence

As described in Chapter 6, flavonoids can sometimes provide useful evidence in understanding plant relationships. The flavonoid content of 12 species of *Coriaria* was determined by Bohm and Ornduff (1981). They included all of the northern Hemisphere species and most from the southern Hemisphere (excluding *C. papuana* and *C. ruscifolia*). The major compounds they identified were monoglycosides of quercetin and kaempferol, comprising 3-0-arabinosides, 3-0-glucosides, 3-0-galactosides, 3-0- xylosides and 3-0-rhamnosides. A small amount of diglycosides were found but no triglycosides. The distribution of flavonoids amongst the species was very similar and the only major difference was that *C. angustissima* contained no kaempferol monoglycosides. All these compounds are commonly found in a wide range of terrestrial dicots and hence provide little information regarding relationships.

8.6.3 Molecular evidence

A recent molecular study of *Coriaria* has been carried out by Yokoyama *et al.* (1998). They used the chloroplast gene regions *rbc*L and *mat*K. They studied 12 species, which included representatives from all geographical regions (Fig. 8.3). For outgroup species they used *Begonia evansiana* Andr. and *Datisca cannabina* L. In the 1224bp *rbc*L data set they found no insertions or deletions and 97 (7.92%) variable bases of which 34 (2.77%) were potentially informative and in the 1186 pb *mat*K data set they found no insertions or deletions within *Coriaria* and 184 (15.51%) variable sites of which 110 (9.27%) were potentially informative. Although these figures seem fairly high, they include variation between *Coriaria* and the two outgroups.

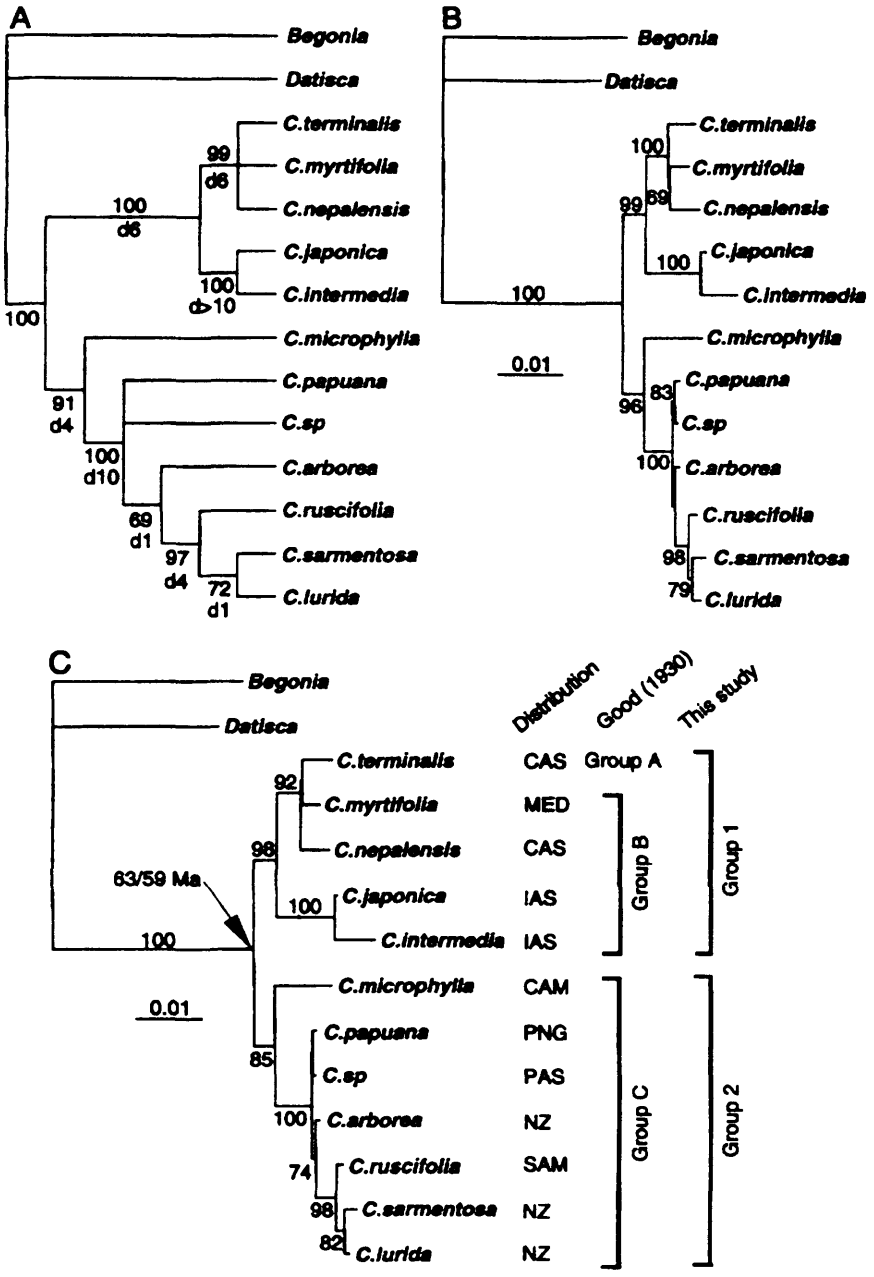

Fig. 8.3. Phylogenetic trees of *Coriaria* species based on the combined *rbc*L and *mat*K sequences. A) Strict consensus of 6 trees using parsimony, B) neighbor- joining tree, C) maximum likelihood tree. (from Yokoyama *et al.*, 1998.) Numbers above the clades are bootstrap values. Abbreviations used are: CAS, continental Asia around the Himalayas; CAM, central America; IAS, insular Asia; MED, the Mediterranean; NZ, New Zealand; PAS, the Pacific islands; PNG, Papua New Guinea; SAM, South America.

This study by Yokoyama *et al* (1998) was interesting but has various limitations which are described as follows. My extended study has many advantages. Firstly, I was able to obtain material from many more of the New Zealand species, their study contained only *C. arborea*, *C. sarmentosa* and what they call *C. lurida* (which under the scheme that I am

following is a variant of *C. sarmentosa*). In addition to *C. arborea* and *C. sarmentosa*, I studied material from *C. angustissima*, *C. kingiana*, *C. pottsiana* and *C. pteridoides*. Furthermore the Yokoyama study only looked at one accession from each species whereas this study looked at multiple accessions where possible. Finally, as well as looking at the chloroplast genome, I have produced a nuclear phylogeny based on the ITS 1 and 2 spacer region and have preliminary data on the *Adh* gene.

8.7 Molecular analyses: materials and methods

8.7.1 Plant material

Table 8.2. Collection details of accessions of *Coriaria* used in the present study. Voucher specimens are lodged in the Leicester University Herbarium.

C. angustissima

Mt. Cook New Zealand. *Thompson*. (*ex* hort. Leicester University Botanic Garden accession T422).

Mt. Cook New Zealand. *Thompson*. (*ex* hort. Leicester University Botanic Garden accession T423).

C. arborea

L. Wakatipu, New Zealand. *Thompson*. (*ex*. hort. Leicester University Botanic Garden, accession T431).

L. Wakatipu, New Zealand. *Thompson*. (*ex*. hort. Leicester University Botanic Garden, accession T433).

L. Wakatipu, New Zealand. *Thompson*. (*ex*. hort. Leicester University Botanic Garden, accession T438).

C. intermedia

ex hort. National *Coriaria* collection Wales (Japan)

ex hort. National *Coriaria* collection Wales (Philippines)

C. japonica

ex hort. Leicester University Botanic Garden, accession Hillier

ex hort. Leicester University Botanic Garden, accession Tubingen

C. kingiana

nr. Gisborne, New Zealand. *Thompson.* ex hort. Leicester University Botanic Garden, accession (T460).

C. microphylla

Ex Leicester University Botanic Garden, accession Kew

C. myrtifolia

Genoa, Italy. Thompson. *ex* hort. Leicester University Botanic garden.

Guadalest, Spain. *Thompson.* *ex* hort. Leicester University Botanic Garden, accession T400.

C. napalensis

ex hort. Leicester University Botanic Garden, accession Glasnevin

C. pottsiana

ex hort. Leicester University Botanic Garden, accession T470

C. pteridoides

ex hort. Leicester University Botanic Garden, accession T450

ex hort. Leicester University Botanic Garden, accession T451

C. ruscifolia

Lincan Ray. Chile. *Thompson.* *ex* hort. Leicester University Botanic garden, accession T605

ex hort. Leicester University Botanic Garden, accession T606

Chile, Región VII [Prov. Cauqueres], at beach at Pelluhue, (35° 49′S, 72° 35′W) 0 m, 2/9/1999. *M.R.J. Clokie & Patricio Peñoilillo*

C. sarmentosa

L. Wakatipu. *Thompson. ex* hort. Leicester University Botanic Garden, accession T401

L. Wakatipu. *Thompson. ex* hort. Leicester University Botanic Garden, accession T402

L. Wakatipu. *Thompson. ex* hort. Leicester University Botanic Garden, accession T403

C. terminalis

Ex hort. Leicester University Botanic garden, accession Treasusres

Begonia

ex Leicester University coffee room

Datisca cannabina

ex hort. Leicester University Botanic garden

8.7.2 Methods

DNA sequences from three gene regions were obtained following the methods described in Chapter 4. Two of the sequences were from the chloroplast (*rpo*A and *trn*L-F) and one was from the nuclear genome ITS 1 and 2. A partial data set was also obtained for the nuclear gene *Adh*.

8.7.3 Analyses

Phylogenies were derived from the *rpo*A, *trn*L-F and ITS data sets. The outgroups were selected on the basis of results from analyses of molecular data (Chase *et al.*, 1993; Swensen *et al.*, 1994; Swensen, 1996).

A point to make about the ITS data set is that originally data for *C. angustissima* was included. On analysis however, this taxon formed a tight group with the Central American *C. microphylla*. Material was derived from seeds that had been harvested from a *C. angustissima* which had been growing in close proximity to *C. microphylla* in the glass

house. The plant seedlings were therefore suspected to be hybrid and thus not included in further ITS analysis. *C. angustissima* was included in the chloroplast phylogenies as cpDNA is inherited maternally without recombination.

8.7.4 Phylogenies

The individual phylogenies are given and discussed for all genes starting with the chloroplast sequences *rpo*A, *trn*L–F, and followed by the nuclear sequence ITS. A phylogeny containing the partial data set for *Adh* is then given. In addition, a tree is given for the combined chloroplast genes, and finally a tree derived from a combination of both the chloroplast genes and the ITS region together.

8.8 Molecular phylogenies

8.8.1 rpoA phylogeny

The *rpo*A sequences produced seven most parsimonious trees of length 42 (Fig. 8.4). There was homoplasy (CI = 0.796, RI = 0.947, RC = 0.925, HI = 0.024). This tree shows that *Coriaria* is clearly monophyletic with respect to its close outgroup *Datisca* (it remained monophyletic when further outgroups were included). Furthermore, all of the New Zealand species apart from *C. angustissima*, together with the Chilean *C. ruscifolia*, form a group which share a single point mutation but which is supported by a bootstrap of 59%. More convincingly all of the southern Hemisphere species form a group which is supported by three point mutations and has a bootstrap value of 89%. The only other structure in this phylogenetic tree is the grouping of the two Asian species *C. intermedia* and *C. japonica* which share two point mutations supported by a bootstrap value of 86%.

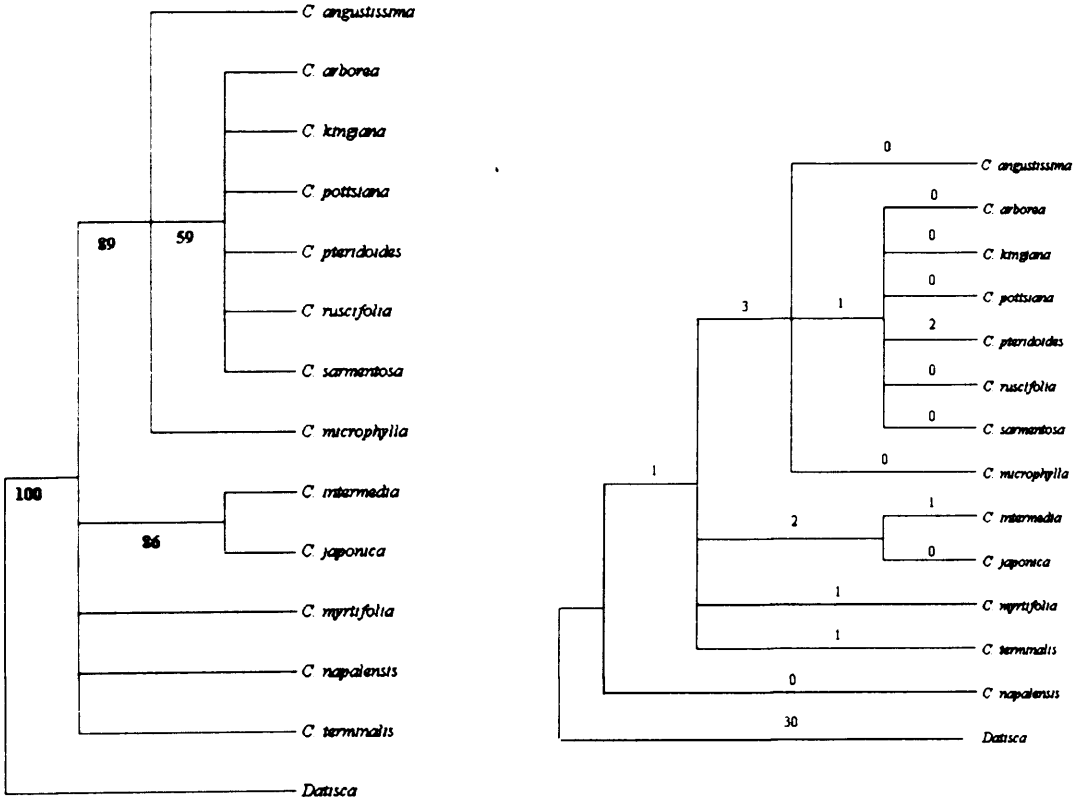

Fig. 8.4 a. Consensus tree for *rpo*A . **b.** Example of shortest tree. For the consensus tree bootstrap values where they occur are below in bold. For the example of the shortest tree the numbers of mutations are given above the clades.

8.8.2 rpoA variation

Statistics relating to the properties of *rpo*A data are given in Table 8.3. From this information the following points can be observed; there are no insertions or deletions. Of the 12 (0.27%) variable bases, half (0.14%) are informative. These are enough to separate the southern Hemisphere species from the northern Hemisphere species. The average number of steps per character (0.027) is the highest of all data sets for this gene. There are more transversions than there are transitions. There is no sequence divergence between some species of *Coriaria* for this gene (Fig. 8.4b), it can be observed that four of the species from New Zealand: *C. arborea*, *C. kingiana*, *C. pottsiana* and *C. sarmentosa* are identical to *C. ruscifolia* from Chile). Maximum sequence divergence of 1.37% is again higher than it is for other DNA regions. The GC content (36.1%) is typical for a chloroplast gene (Gillham, 1994).

Table 8.3 Statistics relating to *rpo*A sequence data in *Coriaria*

Length range (bp)	437
Length mean (bp)	437
Number of characters (aligned length + indels)	437
Number of indels	none
Size of indels (bp)	N/A
G+C content mean (%)	36.1
Sequence divergence (%)	0 – 1.37
Number of constant sites (%)	431 (98.63)
Number of autapomorphic sites (%)	6 (1.37%)
Number of informative sites (%)	6 (1.37%)
Transitions	5
Tranversions	7
Percentage transitions	41.67
Average number of steps per character	0.027

The pattern of variation is illustrated in Fig. 8.5, where it can be seen that all but one of the variable bases in the ingroup-only data set change once and the remaining point mutation occurs twice. The changes in the ingroup data set occur towards either side of the gene region and there is no variation between positions 175 and 315. There is a fairly good correlation between both of the data sets, with the majority of the changes again occurring either side of the gene region. It can be seen that there are two blue peaks in the graph which represent base pair mutations which are invoked more than once only for some of the seven most parsimonious trees.

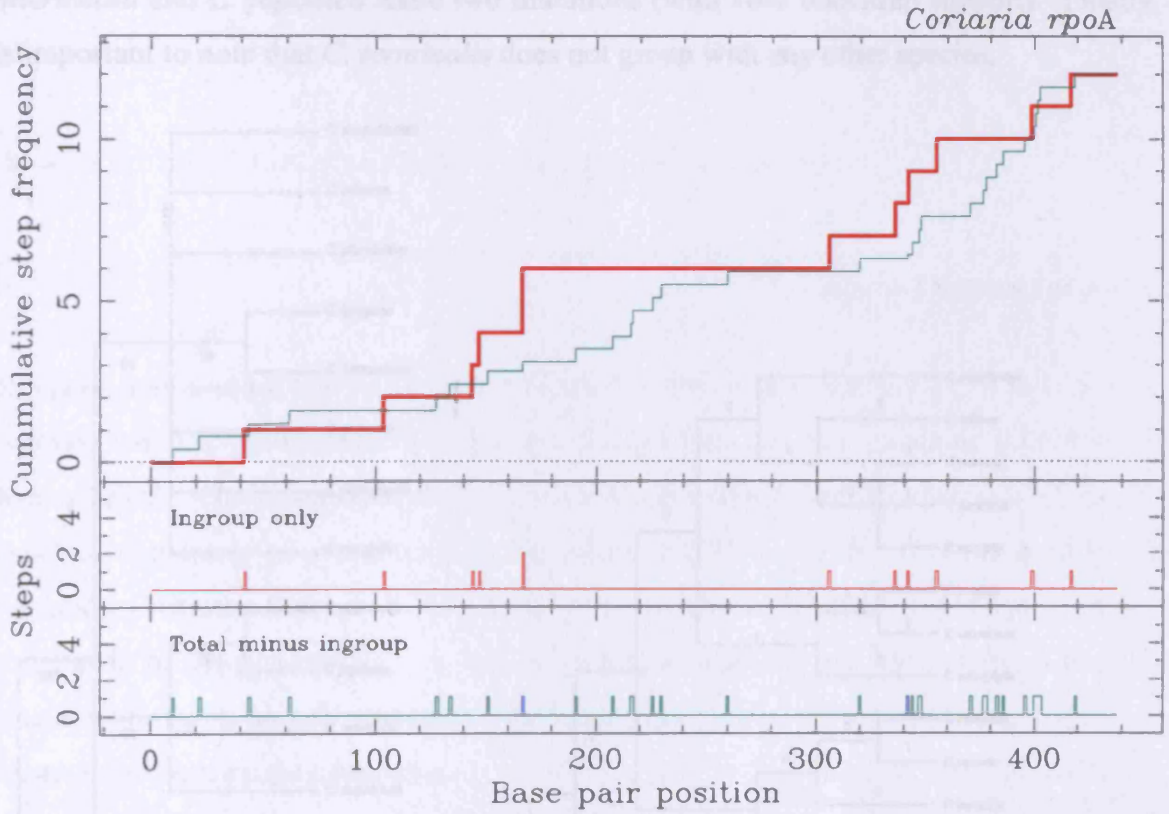

Fig. 8.5. Graph to illustrate the behaviour of *rpo*A in *Coriaria* and its outgroup *Datisca*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Datisca*) from the ingroup (*Coriaria*), plotted in green. These data produce 7 equally most parsimonious trees and include two positions which are invoked more than once in more than one of these trees, such bases are plotted in blue. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

8.8.3 trnL-F phylogeny

The *trn*L-F sequences produced 27 most parsimonious trees of length 96 (Fig. 8.6). There was some homoplasy in this data set (CI = 0.917, RI = 0.830, RC = 0.761, HI = 0.083). *Coriaria* is clearly monophyletic. Within *Coriaria* there are two main well supported clades. The first consists of all of the species from the southern Hemisphere; five characters hold the clade together and it is supported by a bootstrap value of 81%. The only further structure within this group is the group which contains *C. kingiana* and *C. sarmentosa* this group is supported by a 98% bootstrap value. The other main clade contains the species from Eurasia (73% bootstrap support). The two Asian species *C. intermedia* and *C. japonica* share two mutations (with 70% bootstrap support). Finally, it is important to note that *C. terminalis* does not group with any other species.

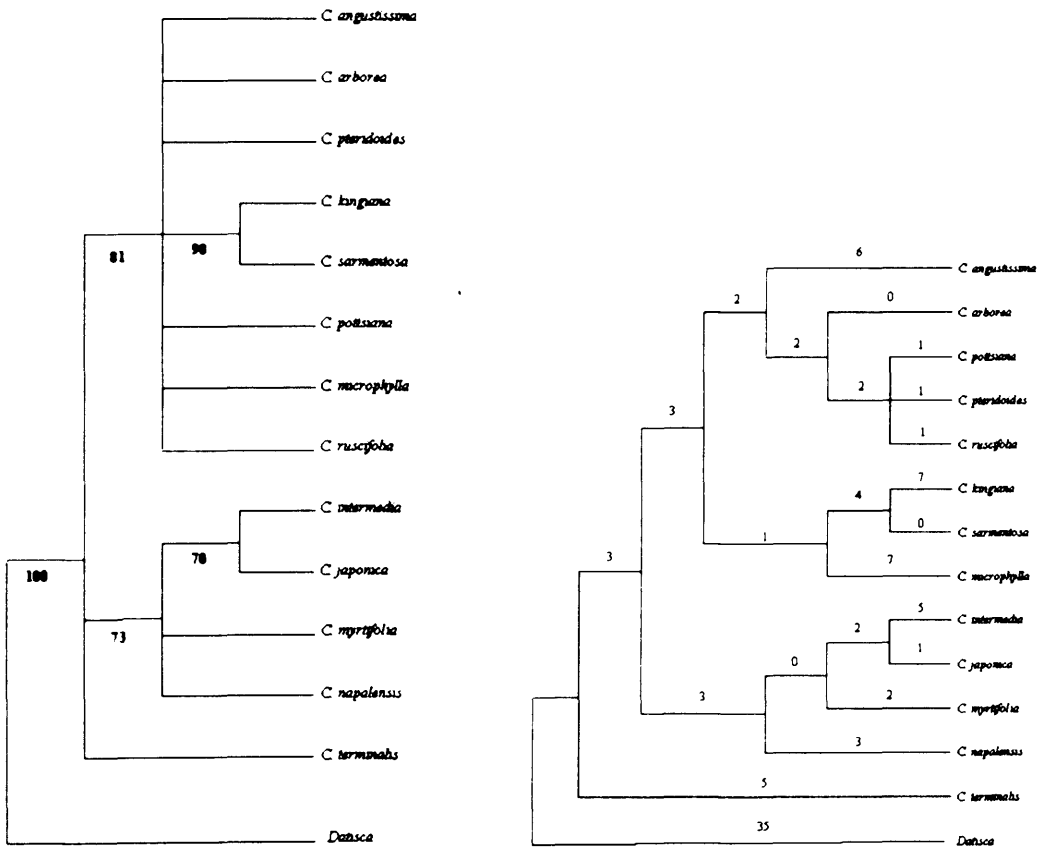

Fig. 8.6. a. Consensus tree for *trn*L-F . **b.** Example of shortest tree. For the consensus tree bootstrap values where they occur are below. For the example of one of the shortest trees, the numbers of mutations are given above the clades.

8.8.4 trnL-F variation

Statistics relating to the properties of *trn*L-F data are given in Table 8.4. In this data set 4.31% of the bases are variable, but only 0.41% are informative. However this is still enough to give some structure to the phylogenetic tree. The average number of steps per character is a relatively high (0.262), compared to other chloroplast regions. There are five indels of which three are informative; these are 5 bp, 7 bp and a variable 1-4 bp in length. The two autapomorphic indels are 2 and 1 base pairs respectively. Interestingly, as was the case for the *rpo*A gene region, there are more transversions than transitions (46.51% transitions). The sequence divergence ranges from 0.27 to a ten-fold difference of 2.732. The GC content is 59.7% which is surprisingly high for the chloroplast genome (Gillham, 1994).

Table 8.4. Statistics relating to *trn*L-F sequence data in *Coriaria*

Length range (bp)	739 - 748
Length mean (bp)	742.61
Number of characters (aligned length + indels)	754
Number of indels (informative)	5 (3)
Size of indels (bp) [uninformative]	5,7, 1-4 [2,1]
G+C content mean %	59.7
Sequence divergence %	0.27 – 2.732
Number of constant sites (%)	710 (95.69)
Number of autapomorphic sites (%)	29 (3.90)
Number of informative sites (%)	14 (0.41)
Transitions	20
Transversions	23
Percentage transitions	46.51
Average number of steps per character	0.262

The pattern of variation is illustrated in Fig. 8.7. When the *Coriaria* alone is considered, there are only two mutations in the first 100 bp, there are then many mutations for the next two hundred base pairs. The gene region is then relatively conserved between positions 300 and 600 bp (with the exception of one base which changes three times). The final stretch of the region sequenced is rich in mutations. In the data set that includes the outgroup *Datisca*, initially there is some correlation with the *Coriaria*-only data set. There are only four mutations until about position 175 bp when there are many mutations for one hundred base pairs. The distribution of variation after this point is fairly even until a position of 550 bp. The reason for there being no mutations after base pair position 580 is because sequence information stops at this point (due to technical difficulties in sequencing the full sequence for *Datisca*).

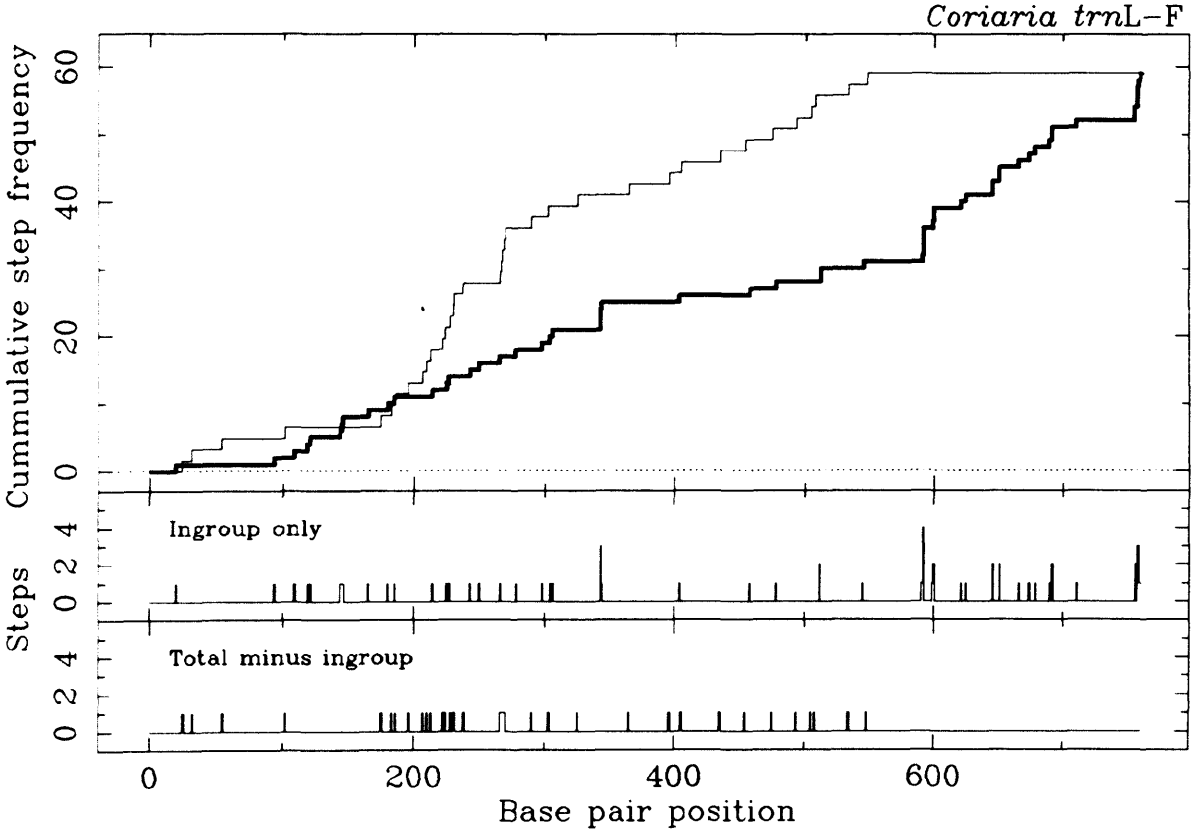

Fig. 8.7. Graph to illustrate the behaviour of *trn*L-F in *Coriaria* and its outgroup *Datisca*. The bottom panel shows the positions and number of mutations that distinguish the outgroup (*Datisca*) from the ingroup (*Coriaria*), plotted in green. The central panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these two data sets, with the plot relating to the outgroup (green) normalised to that of the ingroup (red) so it can fit on the same graph.

8.8.5 Combined rpoA and trnL-F phylogenies

The combined chloroplast data set produces 9 equally most parsimonious trees of length 139 (Fig. 8.8; CI = 0.953, RI = 0.895, RC = 0.812, HI = 0.047). These data result in two groups, one of which contains all of the Eurasian species, and the other consists of all those of the southern Hemisphere. The bootstrap value that includes *C. terminalis* within the other Eurasian species is a low 50%. The two Asian species *C. intermedia* and *C. japonica* form a group which is held together by four characters (90% bootstrap support). The European *C. myrtifolia* and the Asian *C. napalensis* are both sister species to this group, supported by three characters (60% bootstrap support). Within the southern Hemisphere group, the only strongly supported structure is from the two New Zealand species *C. kingiana* and *C. sarmentosa* which form a group supported by five characters (92% bootstrap support). The New Zealand species *C. angustissima* and the Central American species *C. microphylla* both have unresolved positions. The remaining group of the New Zealand species *C. arborea*, *C. pottsiana*, *C. pteridoides* and the Chilean *C. ruscifolia* is only held together by 51% bootstrap support.

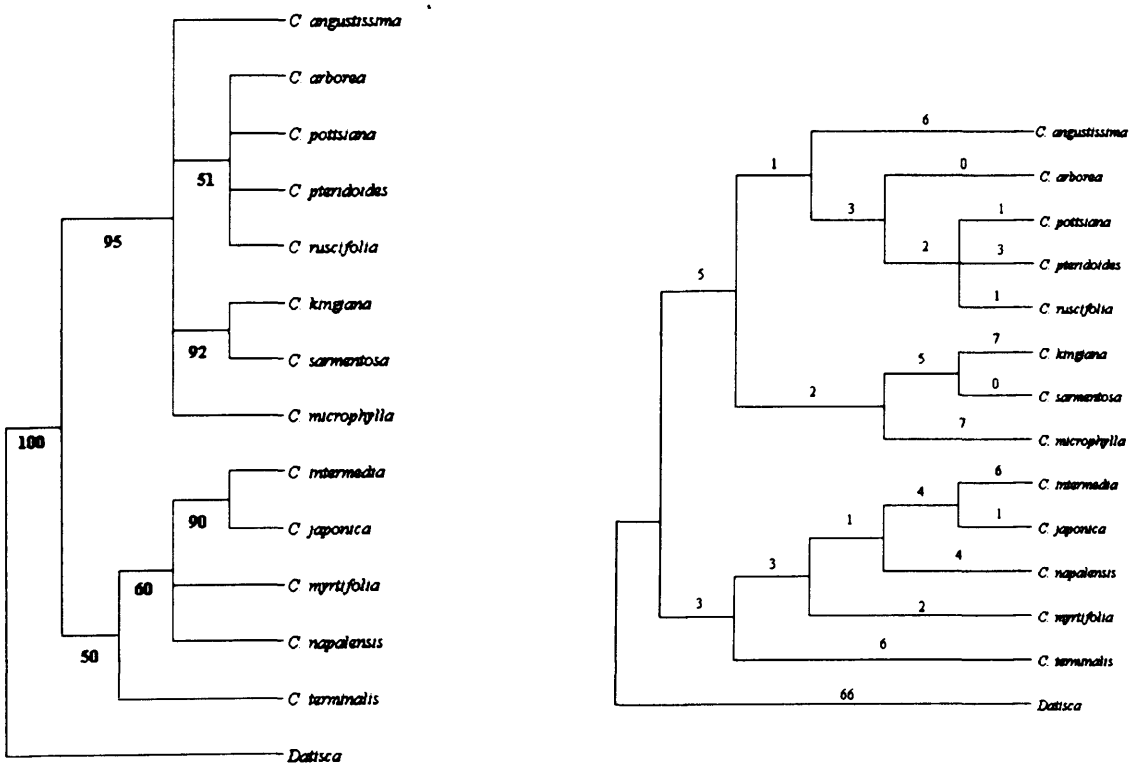

Fig. 8.8. a. Consensus tree for combined chloroplast data **b.** Example of shortest tree.
For the consensus tree bootstrap values where they occur are below in bold; the numbers of mutations are given above the clades.

8.8.6 ITS phylogeny

The ITS sequences produced one most parsimonious tree of length 153 (Fig. 8.9). There was homoplasy (CI = 0.915, RI = 0.856, RC = 0.783, HI = 0.085). Due to the high amount of sequence divergence, it was not possible to align *Coriaria* sequences with the outgroup species *Datisca* (despite the 5.8S coding regions. The tree has been rooted with the Himalayan *C. terminalis*.

The tree shows a highly branched structure. The two Asian species *C. intermedia* and *C. japonica* have 29 characters which group them together, supported by a bootstrap value of 99%. Sister to these species is the European *C. myrtifolia*, five characters hold this clade together with 73% bootstrap support. The Himalayan *C. napalensis* is then sister to the other Eurasian species with 14 characters and 95% bootstrap support. *C. terminalis* does not group with the other species. There is strong support for *C. ruscifolia* being held together with the rest of this group with only seven characters and 90% bootstrap support. The New Zealand species *C. arborea* and *C. pottsiana* are then included within a clade that contains the other New Zealand species and the Central American *C. microphylla*.

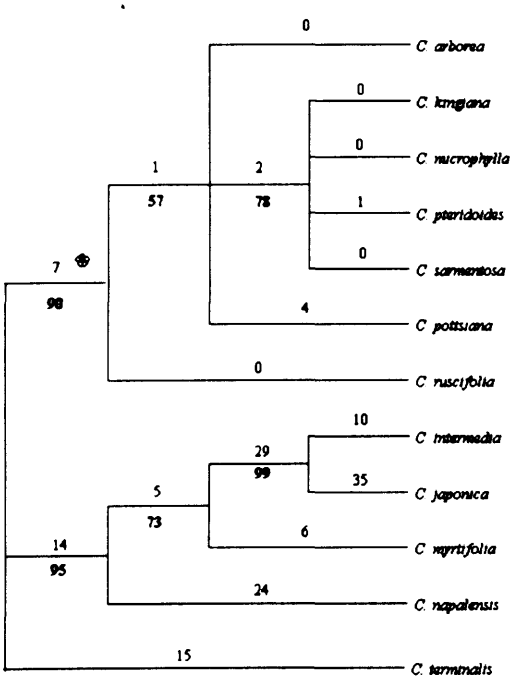

Fig. 8.9. Phylogeny for ITS. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold.

8.8.7 ITS variation

Statistics relating to the properties of ITS data are given in Table 8.5. In the ITS data set, 501 sites (83.36%) are constant, 63 (10.48%) are autapomorphic and 37 (6.16%) are informative. This relatively high amount of variability compared to the chloroplast sequences is reflected in the average number of steps per character being 0.25. The percentage of transitions is 73%. There is a great range in sequence divergence, with none between the New Zealand species *C. kingiana* and *C. sarmentosa* but 14.81% between the New Zealand species and *C. japonica*. In fact, between most other species (except *C. intermedia*) there is approximately 10% sequence divergence, similarly between *C. intermedia* and all the other species (except *C. japonica*) there is approximately 7.5% sequence divergence. Of the 10 indels, only one 1 bp mutation is synapomorphic and all the others are autapomorphies. The GC content is a fairly high 59.7%.

Table 8.5 Statistics relating to ITS sequence data

Length range (bp)	600 - 602
Length mean (bp)	601.42
Number of characters (aligned length + indels)	618
Number of indels (informative)	10
Size of indels (bp) [uninformative]	1, [3,2,2, rest 1]
G+C content mean %	59.7
Sequence divergence %	0 – 14.81
Number of constant sites (%)	501 (83.36)
Number of autapomorphic sites (%)	63 (10.48)
Number of informative sites (%)	37 (6.16)
Transitions	73
Tranversions	27
Percentage transitions	73
Average number of steps per character	0.25

The pattern of variation is illustrated in Fig. 8.10. This shows that the gradient of the cumulative plot is at its most shallow in the 5.8S coding region. Apart from there being less variation in this 5.8S region, the variation along the rest of the sequence is fairly even with no great areas of clustering.

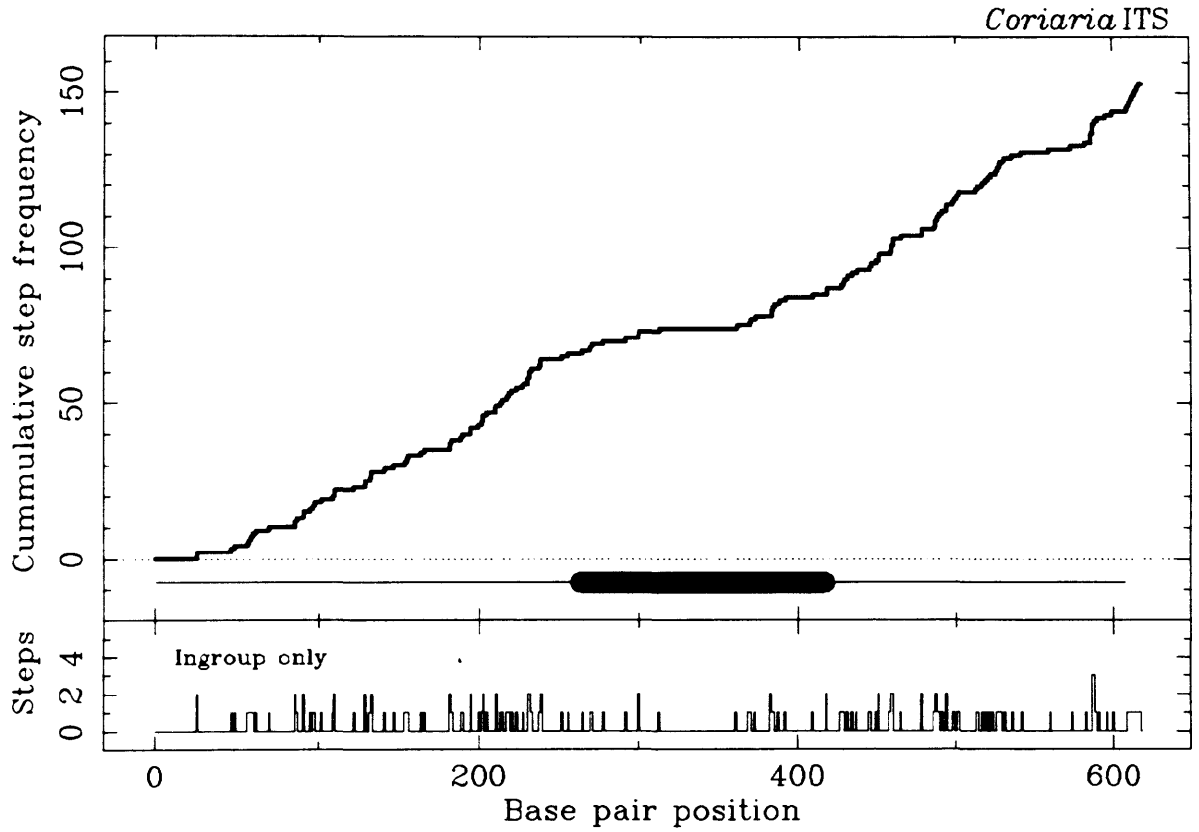

Fig. 8.9. Graph to illustrate the behaviour of ITS in *Coriaria*. ITS spacer 1 (fine black line) is from 1 - 263bp. The 5.8S rRNA gene (indicated by a thick black block) is from 264 - 416. The spacer 2 is shown again as a thin line from 417 - 607 bp. The bottom panel shows the positions and number of mutations that occur in the ingroup sequences only, plotted in red. The top panel shows a cumulative plot of these data.

8.8.8 Adh data set

Unfortunately data were obtained only for nine of the 15 species and no outgroup species. However data were obtained from representative species from all geographical areas apart from the South American *C. ruscifolia*. 350 bp of good sequence data were obtained for all nine species. Two equally most parsimonious trees were produced of length 44 (Fig. 8.10). There was a fairly large level of homoplasy present (CI = 0.773, RI = 0.615, RC = 0.476 and HI = 0.227). Six characters with 72% bootstrap support hold together a clade which consists of the 3 New Zealand species sequenced (*C. arborea*, *C. pottsiana*, and *C. pteridoides*) and the Central American *C. microphylla*. The two Asian species *C. japonica* and *C. intermedia* are supported by characters with 98% bootstrap support. The other two Eurasian species are unresolved. Due to the limitations of the data set, no extensive analysis of the sequence data has been carried out.

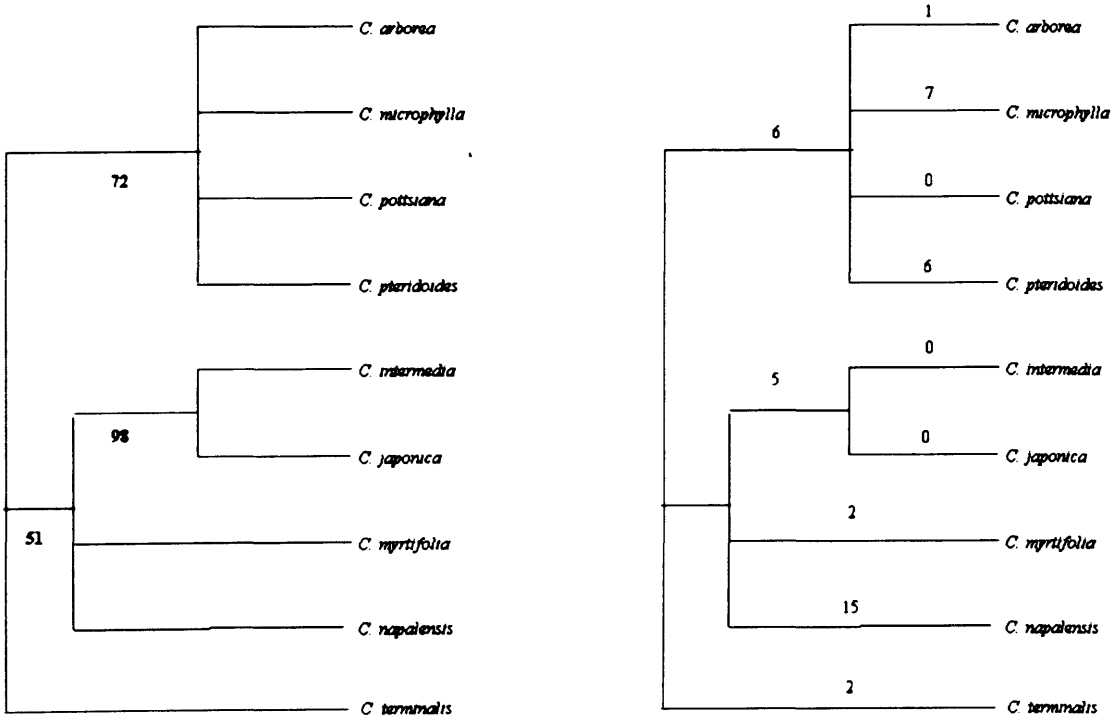

Fig. 8.10. Consensus tree for *Adh* . **b.** Example of shortest tree. For the consensus tree bootstrap values where they occur are below in bold and the numbers of mutations are given above the clades. These trees were rooted with the Himalayan *C. terminalis*.

8.8.9 Combined rpoA, trnL-F and ITS phylogenies

The combined *rpo*A, *trn*L-F and ITS data set produced 13 equally most parsimonious trees of length 294 (CI = 0.868, RI = 0.767, RC = 0.853, HI = 0.047). There are three main groups the first of which consists of the Himalayan species *C. terminalis*. The second contains the Eurasian species. Within this group the two E. Asian species *C. intermedia* and *C. japonica*, are held together with 100% bootstrap support. Sister to this group is the European *C. myrtifolia* with a 75% bootstrap value, and then *C. napalensis*, with bootstrap support of 69%. The third main group contains all of the Southern Hemisphere species which are now held together with 63% bootstrap support. Within this group the only structure is of the group which contains the two New Zealand species *C. kingiana* and *C. sarmentosa* which are supported by a bootstrap value of 97%.

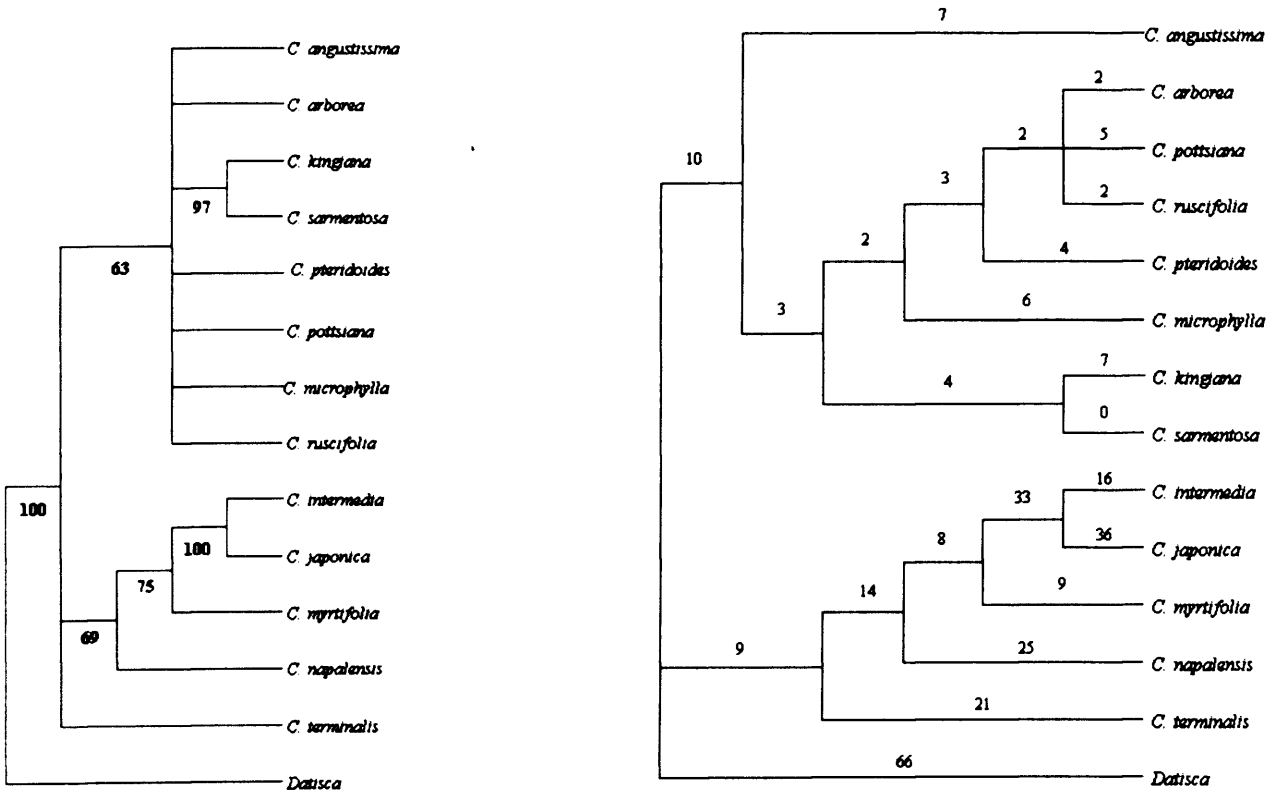

Fig. 8.12. a. Consensus phylogeny for *rpo*A, *trn*L-F and ITS. **b.** example of consensus tree. Numbers of mutations are given above the clades, and bootstrap values where they occur are below in bold. Missing data has been coded as question marks.

8.9 Problems with dating nodes on *Coriaria* phylogenetic trees

Unlike the situation in *Eucryphia*, where the fossil record is comprehensive and well understood, for *Coriaria* only a small amount of fossil data are available. It is much more difficult to accurately place these fossils on a phylogenetic tree and consequently to use them for the calibration of rates of evolution.

As discussed in Section 8.5, only one fossil inflorescence in *Coriaria* has been found. This has been dated at 35 million years old and named *C. longaeba*. It is thought to be the ancestor of *C. terminalis*, but in the absence of a well-defined fossil record it is not clear whether it should be placed in the region of x or y in the schematic tree shown in Fig. 8.13. Even if it were possible to associate it with one of these branches, we would not know whether it belongs before or after the bulk of the mutations on that branch. In the case of *Eucryphia* because far more work has been done on the fossil record, we can be more confident that the fossil occurs before the bulk of the mutations that separate it from the outgroup.

As on the basis of morphological evidence the fossil is clearly an ancestor of *C. terminalis*, it might seem most appropriate to place the fossil at position y on Fig. 8.13. However this would limit the number of mutations which have occurred in 35 million years to very few: one in the case of the *rpo*A gene, six for the *trn*L-F gene and 15 for the ITS. Based on average rates by Wolfe et al. (1987), these would translate to rates of 3.2×10^{-5}, 1.15×10^{-4} and 3.6×10^{-4} substitutions per site per million years for *rpo*A, *trn*L-F and ITS respectively, which are much slower than the average rate of chloroplast gene region evolution, typically $1.1 - 2.9 \times 10^{-3}$ substitutions per site per million years (and note that *rpo*A and *trn*L-F were chosen because they acquire mutations faster than average).

The rate for ITS is also slower than the average, 2.44×10^{-4}, though only by a small factor. To summarise therefore, placing the fossil at point y in the tree only makes sense if the chloroplast genes are evolving extremely slowly.

In contrast, placing the fossil at the root of all *Coriaria* gives mutation rates that are consistent with the average rates: 1.1×10^{-3} and 0.9×10^{-3} respectively for the *rpo*A and

*trn*L-F regions. It is not possible to calibrate a rate for the ITS data if the fossil is placed here as there is no outgroup species.

In conclusion, it seems more likely that this fossil is at the root of all of the *Coriaria* species rather than just the direct ancestor of *C. terminalis*, but without further knowledge of the fossil record it is not possible to say precisely where it should be placed. Consequently it is not possible to accurately calibrate the rates evolution in *Coriaria*.

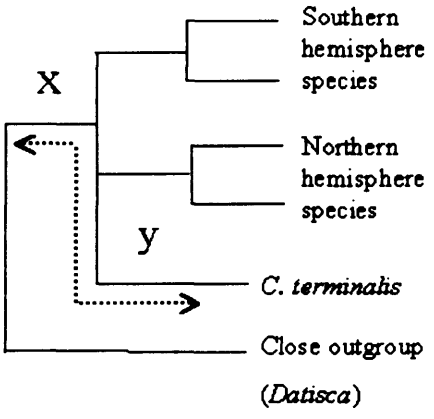

Fig. 8.13. The alternative positions in which to place the 35 MYO *Coriaria* fossil are x and y but it probably belongs somewhere in between. This tree is generalised but fits well with the morphological work of Good (1930), and with the *trn*L-F and the ITS data. x is ancestral to the genus, and y is ancestral only to *C. terminalis*, the alternative placements have quite different effects on dating the rates of evolution. See text for full details.

Fortunately this is not the case with the 11 million year old fossil pollen, with which correct placement on the phylogenetic trees is possible. It is known that this pollen is from a southern Hemisphere species of genus *Coriaria* and that it was found in New Zealand. It was therefore placed at the node of the southern Hemisphere *Coriaria* and used to date the phylogenetic tree.

8.10 Molecular clocks: dating the phylogenetic split

8.10.1 Relative rate tests

Relative rate tests were performed within the southern Hemisphere clade. Tests were not performed for the rest of the species as a) they were outwith the disjunction of interest (they concerned the northern Hemisphere-southern Hemisphere disjunction), and b) no fossil data could be placed on their nodes, and c) it was apparent that they would not pass the tests. Tests were not carried out for the *rpo*A region as there are an insufficient number of mutations for the species to be distinguished.

Table 8.6. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for the southern Hemisphere *Coriaria trnL-F*.

	Ang	Arb	Kin	Mic	Pot	Pter	Rusc	Sar
Ang	X	2,5	14,6	8,5	2,5	2,6	2,6	9,6
Arb	1.29	X	15,3	7,1	1,1	1,1	2,1	9,3
Kin	3.20	8.00	X	9,14	4,15	3,15	4,15	1,6
Mic	0.64	4.5	1.09	X	1,7	0,7	1,7	8,8
Pot	1.29	0	6.37	4.50	X	0,1	1,1	9,3
Pter	2.00	0	8.00	7.00	1.0	X	1,0	9,2
Rusc	2.00	0.33	6.37	4.50	0	1.00	X	9,3
Sar	0.60	3.00	3.57	0	3.00	4.46	3.00	X

Table (8.6) shows that even at the 95% level, there is no rate consistency in the *trn*L-F data set. *C. arborea* fails both with *C. kingiana* and *C. microphylla*. *C. kingiana* and *C. microphylla* both fail with respect to *C. pottisana*, *C. ruscifolia* and *C. sarmentosa*. *C. pteridoides* also fails with *C. sarmentosa*. It is therefore not acceptable to apply a molecular clock to this data set. With respect to the discussion in Section 8.9, even if we could had fixed a fossil to a node, no further analysis could be carried out.

Table 8.7. m_A and m_B values are given above the diagonal and the $(m_A - m_B)^2 / (m_A + m_B)$ ratio below the diagonal for the southern Hemisphere *Coriaria* ITS.

	Arb	Kin	Mic	Pot	Pter	Rusc	Sar
Arb	X	2,0	2,0	4,0	2,0	0,1	2,0
Kin	2.00	X	0,0	4,2	1,0	0,3	0,0
Mic	2.00	0	X	4,2	1,0	0,3	0,0
Pot	4.00	0.67	0.67	X	3,4	0,5	4,2
Pter	2.00	1.00	1.00	0.143	X	0,4	1,0
Rusc	1.00	3.00	3.00	5.00	4.00	X	0,3
Sar	2.00	0	0	0.67	1.00	3.00	X

Some species pairs do not pass at the test at the 95% level ($x^2 = 3.84$), these are *C. arborea* compared to *C. pottsiana*, and *C. ruscifolia* compared both to *C. pottsiana* and *C. pteridoides*. Widening the test to the 99% ($x^2 = 6.63$) level these combinations do pass, so it is acceptable to carry on with the application of a molecular clock.

In conclusion, in all three molecular data sets, there are very few mutations distinguishing the South American *C. ruscifolia* from the New Zealand species. The *rpo*A data set has between zero and two mutations separating the species thus no further analysis was performed. The *trn*L-F data set has some mutations which separate the species but the data fail the relative rate test. Finally, in the ITS data set the relative rate test is passed at the 5% level thus molecular clock was calculated.

8.10.2 Calibrating the clock

The only clock to be calculated was that for the ITS gene. The fossil used in the dating process was the 11 million year old fossil pollen. Rates of evolution were calculated as described in Section 6.10.2. Data are presented in table 8.8.

Table 8.8. Molecular clock calibration table for the ITS region.

Gene region	ITS
Fossil age (MYO)	11
Length	596
Average no. of mutations	23.71
Minimum no. of mutations	21
Maximum no. of mutations	26
Rate (K) $\times 10^{-4}$	1.809
Minimum K $\times 10^{-4}$	1.602
Maximum	1.983

From these data the average age of the divergence was calculated as described in section 6.10.2

Table 8.9. Divergence times for the South American-Australasian disjunction, for the ITS region.

Gene region	ITS
Average no. of mutations between split	3.2
Average divergence time (MYO)	2.97
Minimum divegence time (MYO)	2.71
Maximum divegence time (MYO)	3.35

In conclusion, the average divergence time, from a calibration of the ITS clock is 2.97 million years. It is extremely unlikely that these species could remain so similar if they had been breeding in isolation since the separation of New Zealand from South America (95 – 82 million years ago).

8.11 Discussion and conclusions

The most important phylogenetic split for comparative purposes with the other two genera in this thesis is the split between the South American *C. ruscifolia* and the many species from New Zealand (*C. arborea*, *C. kingiana*, *C. sarmentosa*, *C. pteridoides* and *C. pottsiana*). New Zealand became separated from Antarctica some time between 95 and 82 million years ago, after which it has not been associated with South America. The remarkable similarity between *C. ruscifolia* and the New Zealand species and the molecular clock data (average divergence time approximately 3 million years) show that they have been in contact far more recently than this. Therefore the distribution of *Coriaria* must be explained by dispersal and not vicariance. It may have been dispersed by birds, or by ocean currents, but the split appears to be so recent that human involvement could be an explanation. It is known that the seeds from *Coriaria* contain highly toxic compounds which cause paralysis (Thompson, 1996). Therefore it is possible that man used these seeds to stun prey or in war, for example and so planted *Coriaria* as they inhabited new territories.

In addition to the South American-Australasian disjunction (shared with *Eucryphia* and *Griselinia*) *Coriaria* has a second disjunction between the northern and southern Hemisphere. A particularly interesting species is *C. terminalis* which is only found in the North but has all the morphological features of a southern Hemisphere species. Thompson (1996) studying just the *trn*L intron concluded that *C. terminalis* is basal to all the other *Coriaria* species. In contrast Yokoyama *et al.* (1998) combined data from *rbc*L and *mat*K

genes and concluded that *C. terminalis* forms a sister group with the European *C. myrtifolia*. My work with both the ITS and entire *trn*L-F region demonstrates that *C. terminalis* actually forms a third separate clade. This confirms the conclusions of Good (1930) based on a careful morphological analysis.

CHAPTER 9

DISCUSSION AND CONCLUSIONS

9.1 Molecular phylogenies

The first aim of the project was to establish well-supported molecular phylogenies for three genera of flowering plants (*Eucryphia* (Eucryphiaceae), *Griselinia* (Griseliniaceae) and *Coriaria* (Coriariaceae)), using sequences from the chloroplast and the nuclear genomes. These have been discussed in detail in the relevant chapters, referring to the different genera (Chapters 6-8) and are shown here for general comparative purposes. Three phylogenetic trees are given for each genus (Fig. 9.1 – 9.3): the tree from the combined chloroplast data, the combined nuclear data, and all data combined. For added clarity, in all cases the South American species are in red font, Australasian species in blue font and in *Coriaria* the northern Hemisphere species are in green font. For all trees the numbers below the branches are bootstrap values.

For *Eucryphia* (Fig. 9.1), all three data sets show the Chilean species (in red) form a well supported monophyletic clade sister to the Australian species (in blue), thus indicating a single trans-Antarctic disjunction. The relationships suggested for the Australian species are slightly different in the two genomes. In the chloroplast data there is strong bootstrap support for the Tasmanian clade which contains both subspecies of *E. milliganii* and *E. lucida*. However, in the nuclear data set the Tasmanian species are retained in a clade which also contains the southern Australian *E. moorei*. The combined data agree with the nuclear tree, and makes good biogeographical sense. The Tasmanian species are included with the southern Australian *E. moorei* and the other two species *E. jinksii* and *E. wilkiei* are both unresolved within the monophyletic *Eucryphia*.

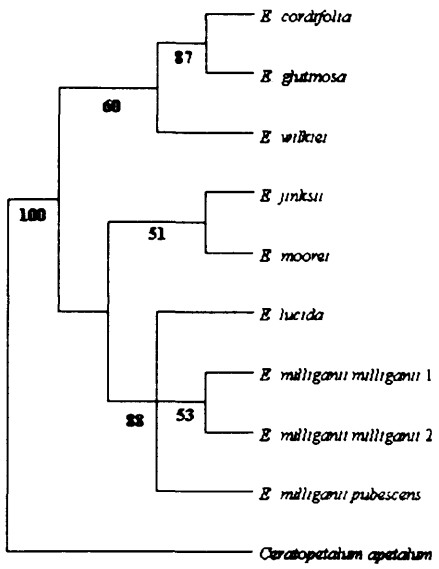

Fig. 9.1a.
Eucryphia: combined chloroplast data (based on *rpo*A, *trn*L-F and *trn*H-K)
CI , 0.96; RI, 0.85; RC, 0.82; HO, 0.036
No. trees, 1; Length, 83.
Numbers below the branches are bootstrap values.

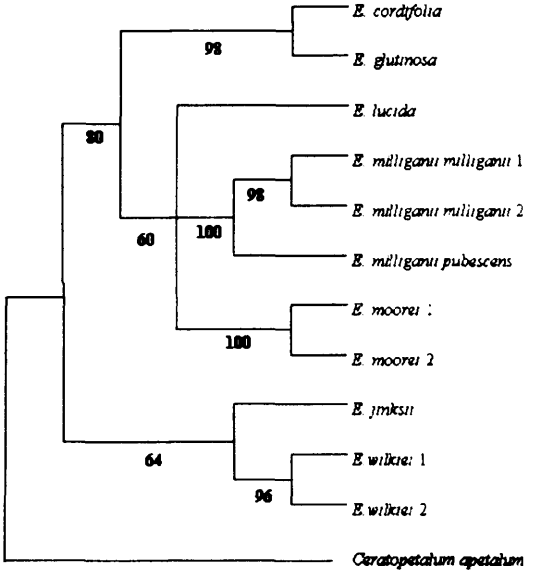

Fig. 9.1b.
Eucryphia: strict consensus combined nuclear data (based on ITS, *G3pdh*-1 and *G3pdh*-2)
CI , 0.96; RI, 0.91; RC, 0.87; HO, 0.045
No. trees, 2; Length,331.
Numbers below the branches are bootstrap values.

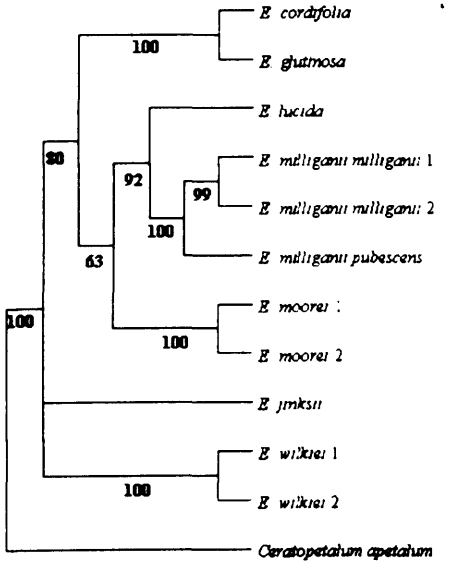

Fig. 9.1c.

Eucryphia: strict consensus for combined chloroplast and nuclear data
(based on *rpo*A, *trn*L-F and *trn*H-K, ITS, *G3pdh*-1 and *G3pdh*-2)
CI , 0.95; RI, 0.89; RC, 0.85; Ho, 0.05
No trees, 2; Length, 417.
Numbers below the branches are bootstrap values.

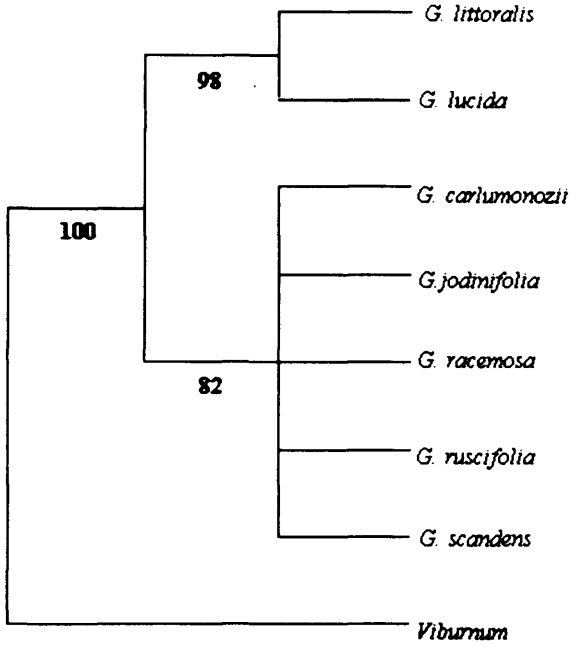

Fig. 9.2a.
Griselinia: strict consensus, combined chloroplast data (based on *rpo*A, *trn*L-F and *trn*H-K).
CI , 0.84; RI, 0.57; RC, 0.48; HO, 0.16
No. trees, 14; Length, 226.
Numbers below the branches are bootstrap values.

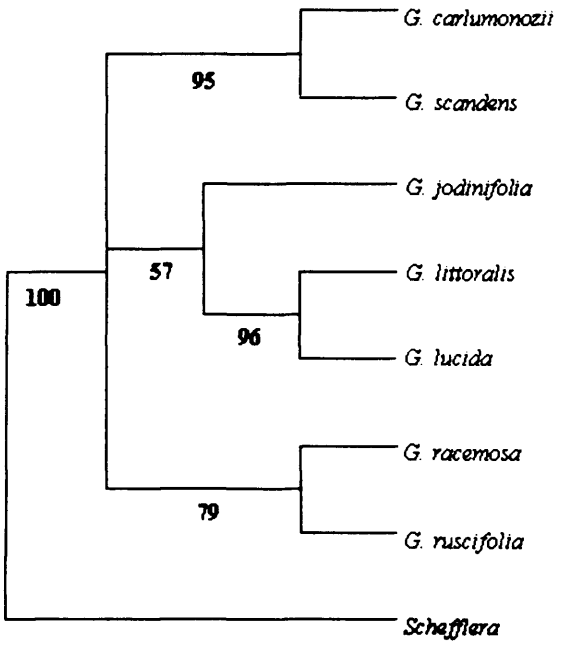

Fig. 9.2b.
Griselinia: strict consensus, nuclear data (based on ITS)
CI , 0.83; RI, 0.65; RC, 0.54; HO, 0.017
No. trees, 1; Length, 100.
Numbers below the branches are bootstrap values.

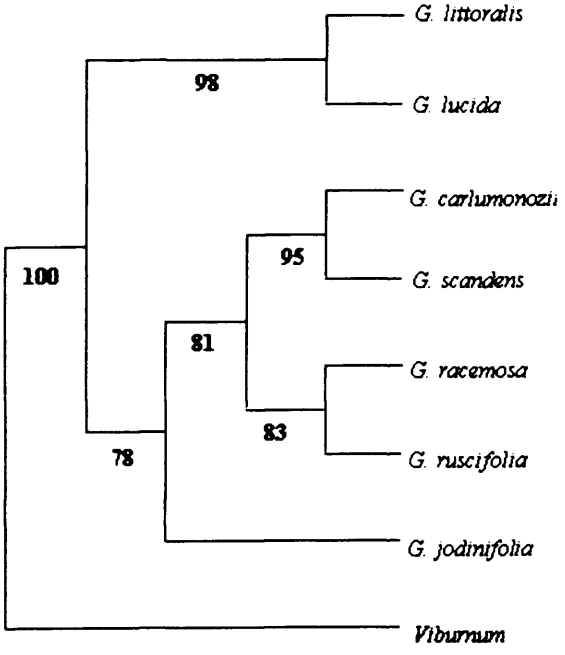

Fig. 9.2c.

Griselinia: strict consensus for combined chloroplast and nuclear data (based on *rpo*A, *trn*L-F and *trn*H-K and ITS)
CI , 83; RI, 62; RC, 51; HO, 0.020
No. trees, 1; Length 334.
Numbers below the branches are bootstrap values.

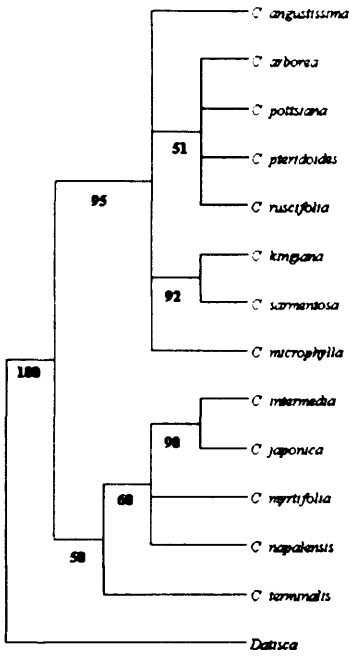

Fig. 9.3a.
Coriaria: strict consensus combined chloroplast data (based on *rpo*A and *trn*L-F).
CI , 0.953|; RI, 0.895; RC, 0.812; HO, 047
No. trees, 9; Length, 139.
Numbers below the branches are bootstrap values.

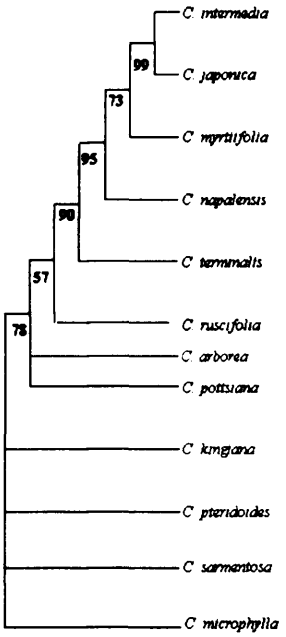

Fig. 9.3b.
Coriaria: nuclear data (based on ITS)
CI , 0.92; RI, 0.86; RC, 0.78; HO, 0.085
No. trees, 1; Length, 153.
Numbers below the branches are bootstrap values.

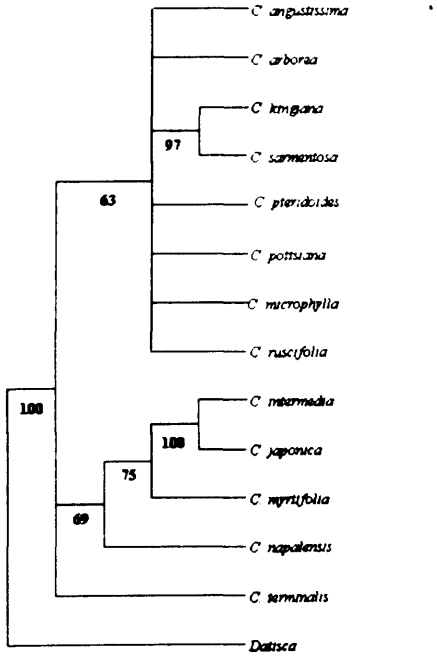

Fig. 9.3a.

Coriaria: strict consensus for combined chloroplast and nuclear data
(based on *rpo*A, *trn*L-F and and ITS)
CI , 0.860; RI, 0.767; RC, 853; Ho, 0.047
No trees, 13; Length, 294.
Numbers below the branches are bootstrap values.

For *Griselinia* (Fig. 9.2a-c), in all cases, it is the New Zealand species (in blue) which form a well supported monophyletic clade, again indicating a single trans-Antarctic disjunction.. In the chloroplast genome the Chilean species (in red) form a monophyletic but unresolved group. The nuclear data support the further grouping of the northern Chilean species *G. carlomunozii* and *G. scandens* and the southern Hemisphere species *G. racemosa* and *G. ruscifolia.* The combined data set clearly shows good bootstrap support for the monophyletic grouping of the Chilean species and the further aforementioned subdivisions. The central Chilean *G. jodinifolia* is basal to the other Chilean species.

In the case of *Coriaria* (Figs. 9.3a-c), for both genomes, the grouping of the northern Hemisphere species (in green) is very clear (although not necessarily well supported in the cpDNA data). There is strong bootstrap support in the chloroplast genome for the grouping of all species from the southern Hemisphere. There is also strong support for a tight grouping of the New Zealand species *C. kingiana* and *C. sarmentosa.* In the combined tree the southern Hemisphere species form a largely unresolved monophyletic group. It has been claimed that the South American species *C. ruscifolia* and the central American *C. microphylla* (in red) form a sister group, but the chloroplast and nuclear trees suggest that this is not the case. At least one trans-Antarctic disjunction is indicated, involving the Chilean *C.ruscifolia* and the New Zealand species. Within the northern Hemisphere species the two Asian species *C. intermedia* and *C. japonica* form a well supported group which is sister to the European *C. myrtifolia.* This in turn is sister to the Himalayan *C. napalensis.* The final Himalayan species *C. terminalis* is unresolved within the monophyletic *Coriaria* assemblage.

9.2 Evolutionary rates

The background to DNA sequence evolution was given in Chapter 3 where the merits and the problems of utilising different genomes were discussed. The specific way in which they were used in this thesis was discussed in Chapter 4. It is important to reiterate that evolution is not constant over the individual genomes. Figures are often quoted for average rates of evolution, for example the 'average' chloroplast mutation rate is often cited as 1.1 - 2.9 x 10^{-9} substitutions per synonymous site per year (Wolfe *et. al.*, 1987, 1989). In comparison, mitochondrial genes have much lower rates of evolution at 0.2 - 1.1

x 10^{-9} substitutions per synonymous site per year (Wolfe *et. al.*, 1987, 1989). In and the nuclear genes have much higher rates of evolution of up to 31.5 x 10^{-9} substitutions per synonymous site per year. These figures are based on an analysis of 31 genes, representing all three plant genomes (Wolfe *et. al.*, 1987, 1989). Gaut (1998) carried out a larger study, calculating the rates of evolution of sequence data from 94 rice and maize genes and arrived at very similar results.

I was interested in the rate of mutation in specific gene regions. Particularly, I wanted to find gene regions that evolved at a suitable rate to produce robust phylogenies, preferably for the three different genera. As was pointed out in chapter 3, certain chloroplast regions are more commonly used for studies at a species level than others. However there has been no systematic comparison of the average rates of these genes. Based on the divergence between monocot and dicots, Wolfe *et al.* (1987, 1989) calculated that in the chloroplast genome, genes in the inverted-repeat region evolved considerably slower (approximately by a factor of five) than the single copy genes. All of the chloroplast genes in the present study are in the single copy region and thus might be expected to evolve fast enough to elucidate species relationships.

The reasons for selecting different gene regions studied in this project are given in chapter 5. In many cases, the *trn*L-F region has been shown to evolve suitably quickly to produce a robust phylogeny, for example in Acanthaceae it was shown to evolve faster than either the *rbc*L or the *ndh*F region (McDade, 1999). Similarly to the chloroplast genome, there has not been a comprehensive review of the rates of evolution of nuclear genes. Studies of the commonly used ITS spacer regions 1 and 2 have been reviewed recently by Richardson *et al.* (2001) and they concluded that the average rate of evolution in angiosperms is 2.44 x 10^{-9} mutations per site per year. The only study thus far in which the *G3pdh* gene region has been used taxonomically is a survey of cassava (*Manihot esculenta*) (Olsen and Schaal, 1999). Interestingly, in *Manihot,* the variation was sufficient to resolve relationships at the subspecies level. It was therefore of great interest to see what level of taxonomic resolution that this gene region would provide.

In my studies, the *rpo*A region was not particularly variable at the desired species level. It was, however, consistently amplified in every species for which it was tried (including outgroups). In the case of *Eucryphia*, it did not produce any meaningful phylogeny, for

Griselinia however it did differentiate the New Zealand species from those from Chile. In *Coriaria*, there was enough variability in this region to differentiate the southern Hemisphere species from the northern Hemisphere ones. In summary therefore this region is particularly useful in establishing relationships at a genus level and provides a small but consistent amount of resolution of the level of species.

There is good reason why the *trn*L-F region of DNA has commonly been the region of choice for molecular phylogeneticists. It was amplified consistently for all three genera and for most of the outgroup species. It also provided a reasonable level of resolution between species in all three genera. In all three cases, the number of variable bases was fairly consistent throughout the intron and the IGS region although the initial part of the IGS contained fewer variable positions. This region was the most informative chloroplast sequence used in the study.

The *trn*H-K region was able to distinguish the Chilean species from those from Australasia for *Eucryphia*. It was particularly useful in the case of *Griselinia*, providing a similar level of resolution to the *trn*L-F region. The rate of change in *Eucryphia* is a little faster than *trn*L-F, at $2.57 \times 10^{-4} - 3.79 \times 10^{-4}$ changes per million years. Unfortunately it could not be sequenced for *Coriaria* due to its high AT content which in many species caused slippage at the 5' end of the region. In conclusion therefore this chloroplast region of DNA is potentially useful but must be assayed before establishing whether or not it is suitable.

It is worth mentioning that the average rates of evolution are similar for all three chloroplast regions, but it is the distribution of mutations which differentiates useful regions from those which are less informative. Most of the mutations present in the *rpo*A region and to a certain extent the *trn*H-K region were shared by all *Eucryphia* species and by all southen Hemisphere *Coriaria* species. In contrast the *trn*L-F mutations occurred further down the branches of the phylogenetic tree. All chloroplast regions calculated here are slower than the supposed average of $1.1 - 2.9 \times 10^{-9}$ substitutions per synonymous site per year. This shows the importance of not using general clocks but specifically calculating ones which apply to the specific regions of interest.

The ITS gene region was consistently informative. One of the main problems with this region is the dissimilarity between the ingroup and the outgroup species, thus making

alignment difficult, even impossible. No outgroups were used for this region in either *Coriaria* or *Eucryphia,* and in *Griselinia* an outgroup had to be used which was different from that which was used for the other gene regions. One of the other main concerns regarding this region is its multi-copy nature and associated problems (discussed fully in Chapter 3). The average clock for ITS evolution was used to calculate the evolutionary time that had elapsed for *Griselinia.* Despite a lack of rate constancy, rates of evolution were calculated for *Eucryphia* and *Coriaria* and are 2.13 and 1.81×10^{-9} changes per site per year respectively. This compares well with the average rates of mutation for ITS of 2.44×10^{-9} changes per site per year (Richardson *et al.*, 2001).

Low copy nuclear genes are increasingly being utilised to assess relationships between plant species. In this study the *Adh* region looked potentially useful. Unfortunately technical difficulties meant that only a partial phylogeny was produced for *Coriaria* and nothing for the other two genera. The two loci that were amplified for the *G3pdh* region proved to be very useful and informative for *Eucryphia.* The actual rates of change were around a factor of ten higher in *Eucryphia* than the chloroplast regions were. Although similar size products were amplified for *Coriaria,* they could not be sequenced directly. It was not possible to amplify this region in *Griselinia.*

It is worth mentioning that even in the short duration of this project there have been enormous improvements in the polymerase enzyme which is used to amplify single or low copy genes. Biotechnology companies have realised the importance of low copy amplification and have modified both *Taq* and other polymerase enzymes with this aim. The synthesis of primers has also become much cheaper thus making it easier to optimise them, as primers which do not show potential at an early stage can be re-designed.

Single or low copy genes are increasingly easier to work with than they have been and provide an interesting data source. *G3pdh* in particular shows great potential as an informative marker at the species level, being able to detect differences between sub-species. In *Eucryphia* it evolved approximately ten times faster than the *trn*L-F region. Although it evolved around half the rate of the *G3pdh* loci, they are much longer so therefore there is more data. Thus these *G3pdh* regions were able to distinguish relationships which neither chloroplast nor ITS regions of DNA detected.

Table 9.1. Rates of evolution among the different DNA sequences studied. Values given are average number of mutations per site per year x 10^{-9}.

Calibration/Sequence	*rpo*A	*trn*H-K	*trn*L-F	ITS	*G3pdh*-1	*G3pdh*-2
Eucryphia 60	0.13	0.26	0.18	-	0.98	0.61
Eucryphia 2.6	0.43	0.35	0.21	2.13	0.65	0.70
Coriaria 11	0.42	-	0.67	1.81	-	-
Average for these two genera	**0.33**	**0.31**	**0.35**	**1.97**	**0.82**	**0.66**

9.3 Evolutionary rate-constancy

In this thesis molecular clocks were only applied where it could be shown that mutation rates were constant between the different lineages. This was not the case in either the Wolfe (1987, 1989) or Gaut (1998) studies.

Constancy of mutation rate was tested using the rate heterogeneity test described by Tajima (1993). This is described in Chapter 4, and in Chapters 6-8. Where at least one mutation was present to distinguish each species, sequences for all three genera were tested. For *Eucryphia* rate constancy was established for *rpo*A, *trn*H-K, *trn*L-F, *G3pdh-1* and *G3pdh-2*. Of the regions tested only the ITS region failed, although this failure was fairly comprehensive, with most species comparisons failing the test. For *Griselinia*, *rpo*A, *trn*L-F and ITS passed the rate constancy test. The chloroplast region *trn*H-K did not pass due to the comparison of the southern Chilean *G. racemosa* with the New Zealand species. In *Coriaria,* rate constancy was established only in the ITS sequences of the southern Hemisphere species.

Where rate constancy was established, the rates are given in Table 9.1.

9.4 Critique of molecular clocks

An introduction to molecular clocks was given in Chapter 3. One question which must be answered is: are molecular clocks useful? It is argued by some researchers that the error involved in molecular clocks is so large as to render them useless. In a recent review, however, Sanderson (1998) concluded that molecular rate constancy is a reasonable model across many parts of the tree of life, despite a few clear exceptions.

One major problem with molecular clocks is the difficulty in calibrating them. Obviously factors such as the completeness of the fossil record can enormously alter the time over which the mutations occur. For example we know that the fossil record for *Eucryphia* is fairly well characterised and understood. Therefore we can feel confident by placing the 60 million years old fossil at the root of the *Eucryphia* tree. However, if the record were not so strong, it would be possible that a newly-discovered 100 million year old fossil could be placed at the root of the tree: displacing the original fossil and considerably slowing down the measured rate of evolution across the branches of the tree. It is particularly lucky for the purposes of this project that the *Eucryphia* fossils have been so well studied. We can thus be confident in the calibration of the trees. The fossil record is much less complete in *Coriaria,* and is absent in *Griselinia.*

I think that the concept of a molecular clock is an extremely valuable tool, although extreme care must be taken in their calibration and interpretation. As more sequence data become available, particularly for taxa with well understood fossil records, it will be possible to create increasingly accurate general clocks. This will be useful for calculating divergences in taxa that do not have well documented fossil histories.

9.5 Divergence times between the South American-Australasian disjunction

The splitting of continents was reviewed fully in Chapter 2. The significant dates for *Griselinia* and *Coriaria* are 95 to 82 million years ago when New Zealand became separated from Antarctica and therefore was no longer associated with South America. For *Eucryphia*, this date is later as Australia did not become separated from Antarctica

until around 40 million years ago. This means that for the South American – Australasian disjunction to be explained in terms of vicariance there must be 95-80 millions years of molecular divergence between species for *Griselinia* and *Coriaria*, and 40 million years worth of divergence for *Eucryphia*.

Tables containing divergence times based on the rates above, for the South American-Australasian disjunction are given below for *Eucryphia* (Table 9.2), *Griselinia* (Table 9.3) and *Coriaria* (Table 9.4). Average divergence times are shown in bold type.

Table 9.2. Divergence times for the four gene regions for *Eucryphia*.

Gene region	***trn*H-K**	***trn*H-K**	***trn*L-F**	***trn*L-F**	***G3pdh*-1**	***G3pdh*-1**	***G3pdh*-2**	***G3pdh*-2**
Fossil age (MYO)	60	2.6	60	2.6	60	2.6	60	2.6
Av. divergence time (MYO)	**11.49**	**7.78**	**47.39**	**34.86**	**15.88**	**26.42**	**19.04**	**22.98**
Min. divergence time (MYO)	10.54	3.74	42.40	17.43	15.35	13.20	18.77	22.98
Max. divergence time (MYO)	12.00	18.67	53.70	-	16.47	79.38	19.64	22.98

Table 9.3. Divergence times for the three gene regions for *Griselinia*.

Gene region	***trn*H-K**	***trn*H-K**	***trn*L-F**	***trn*L-F**	**ITS**
Fossil age (MYO)	60	2.6	60	2.6	Ave. clock
Av. divergence time (MYO)	**83.54**	**56.65**	**53.35**	**39.25**	**20.60**
Min. divergence time (MYO)	76.68	27.25	47.74	19.62	
Max. divergence time (MYO)	87.27	135.89	60.46	-	

Table 9.4. Divergence times for the ITS gene region for *Coriaria*.

Gene region	ITS
Fossil age (MYO)	11
Average divergence time (MYO)	**2.97**
Minimum divegence time (MYO)	2.71
Maximum divegence time (MYO)	3.35

For *Eucryphia*, only one of the average divergence times is consistent with the 40 million years required by vicariance (that for the *trn*L-F clock calibrated with the 60 million year old fossil). Including the maximum divergence times adds only one more clock (that for the *G3pdh*-1 region calibrated with the 2.6 million year old fossil). Thus I conclude that the disjunctive distribution of *Eucryphia* most likely arose via dispersal.

Although the divergence times are longer for *Griselinia* than they are for *Eucryphia*, the 95 - 80 million year age of separation of New Zealand from South America means that two of the five clocks show enough variation to account for the distribution in terms of vicariance. The other three suggest that the present day distribution of *Griselinia* arose by dispersal.

Coriaria has by far the lowest divergence times for any of the genera and thus its present day distribution must also be attributable to dispersal. Applying ITS rates suggests the disjunction arose within the last three million years.

9.6 *Eucryphia*, *Griselina* and *Coriaria* in the context of other taxonomic groups

In Chapter 2, I discussed all literature that described molecular work on plant and animal taxa that share the disjuntive distribution between South America and Australasia. This is the first review of these groups together, as previously each genus or family had been looked at in isolation.

It became clear that there are different reasons for the disjunctions of taxa that share this distribution, and that there is not one mechanism to explain all of the groups. Above the

level of the genus, in groups such as the marsupials, ratite birds, or the Araucariaceae plant family, it seems that vicariance is consistent with the large amount of molecular variation observed.

In contrast, the disjunctions of groups at the level of the genus and below (for example *Abrotanella* (Asteraceae), *Sophora* (Fabaceae) and the fish *Galaxias maculatus*) have to be explained by invoking dispersal as a mechanism, rather than vicariance. In other cases, such as *Nothofagus*, it seems likely that the distribution is due in part to vicariance and in part to dispersal. Humphries (1999) envisaged that all such cases of disjunction were part of a generalised track caused by vicariant events. But when the groups are studied more closely, different underlying factors seem to have caused their present day distribution. Clearly, there is no such thing as one simple southern Hemisphere disjunctive distribution.

The results from this thesis take the findings from the literature one step further. The small amount of molecular variation observed in three unrelated plant genera *Eucryphia*, *Griselinia* and *Coriaria*, indicates that in all cases their trans-Antarctic distribution probably arose through dispersal, although the case of *Griselinia* is open to more doubt than the others. This is the case despite different dispersal mechanisms: dry winged seeds (*Eucryphia*) or fleshy diaspores (*Griselinia* and *Coriaria*).

Although all the distributions may possibly be attributed to dispersal, I have also shown that these dispersal events have happened at different times. In *Eucryphia* I calculated that the disjunction arose about 23 million years ago whereas South America and Australasia became geographically separated 50 – 40 million years ago. The average divergence time for *Griselinia* species on either side of the Pacific is 51 million years whereas the continental split of New Zealand from South America occurred approximately 95 – 80 million years ago. The *Coriaria* split occurred far more recently, at about three million years ago.

References

Adachi, J., Kosuge, K., Denda, T., and Watanabe, K. (1995). Phylogenetic relationship of the Berberidaceae based on partial sequence of the *gapA* gene. Plant Systematics and Evolution *9 (Suppl.)*, 351-353.

Allan, H. H. (1961). *Coriaria*. In Flora of New Zealand (Wellington: Government Printer), pp. 300 - 305.

Allan, H. H. (1927). Illustrations of wild hybrids in the New Zealand flora. Genetica *9*, 145 - 146.

Archer, M., Godthelp, H., and Hand, S. J. (1993). Early Eocene marsupial from Australia. Kaupia *3*, 193 - 200.

Arroyo, M. T. K., Riveros, M., Penaloza, A., Cavieres, L., and Faggi, A. M. (1996). Phytogeographic relations and regional richness patterns of cool temperate rainforest flora of southern South America. In high latitude rainforests and associated ecosystems of the west coast of the Americas
climate, hydrology, ecology and conservation, R. G. Lawford, P. Alaback, G and E. Fuentes, eds.: Springer, pp. 134-172.

Ashton, D. H. (1976). Foreword to 'Seed dispersal'. In Seed Dispersal, D. Murray, ed. (Sydney: Academic Press), pp. xi - xii.

The Australian National University website. 1998. Leatherwood honey from the forest. Last update: 29 10 1999.

Avise, J. C. (1986). Mitochondrial DNA and the evolutionary genetics of higher animals. Philosophical Transactions of the Royal Society of London B *312*, 325 - 342.

Avise, J. C. (1994). Molecular markers. natural history and evolution (London: Chapman & Hall).

Baillon. Histoire naturelle des plantes, pp. 402.

Bailey, C. D., and Doyle, J. J. (1999). Potential phylogenetic utility of the low-copy nuclear gene pistillata in dicotyledonous plants: comparison to nrDNA ITS and trnL intron in Shaerocardamum and other Brassicaceae. Molecular phylogenetics and evolution *13*, 20-30.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., and Donoghue, M. J. (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden *82*, 247 - 277.

Barnes, R. W., Hill, R., and Bradford, J. C. (2001). The history of Cunoniaceae in Australia from macrofossil evidence. Australian Journal of Botany *49*, 301 - 320.

Barnes, R. W., and Jordan, G. I. (2000). *Eucryphia* (Cunoniaceae) reproductive and leaf macrofossils from Australian Cainozoic sediments. Australian Systematic Botany *13*, 373 - 394.

Barnes, R. W., Jordan, G. J., Hill, R. S., and McCoull, C. J. (2000). A common boundary between distinct northern and southern morphotypes in two unrelated Tasmanian rainforest species. Australian Jounal of Botany *48*, 481 - 491.

Barnsley, B. (1983). Eucryphiaceae. In Flowering plants in Austalia, H. R. Toelken, ed. (Adelaide: Rigby Publishers), pp. 137 - 138.

Bate-Smith, E. C., Davenport, S. M., and Harbourne, J. B. (1967). Comparative biochemistry of flavonoids - III. A correlation between chemistry and plant geography in the genus *Eucryphia*. Phytochemistry *6*, 1407 - 1413.

Bausch, J. (1983). A revision of the Eucryphiaceae. Kew Bulletin *8*, 317 - 349.

Beadle, N. C. W. (1981). Origins of the Australian flora. In Ecological Biogeography of Australia, A. Keast, ed. (The Hague: Junk), pp. 407 - 426.

Beck, C. B. (1976). Origin and Early Evolution of Angiosperms (New York: Columbia University Press).

Becker, A., Winter, K. U., and Meyer, B. (2000). MADS-box gene diversity in seed plants 300 million years ago. Molecular Biology and Evolution *17*, 1425-1434.

Berra, T. M., Crowley, L. E. L. M., Ivantsoff, W., and Fuerst, P. (1997). *Galaxias maculatus*: an explanation of its biogeography. Marine and Freshwater Research *48*, 80 - 89.

Beuzenberg, E. J., and Hair, J. B. (1959). Contributions to a chromosome atlas of the New Zealand flora. 5. Miscellaneous families. New Zealand Journal of Botany *1*, 53-67.

Billardière, L. (1799/1800). Relation du voyage à la recherche de la Pérouse, pp. 16.

Bohn, B. A., and Ornduff, R. (1981). Leaf flavonoids and ordinal affinities of Coriariaceae. Systematic Botany *6*, 15 - 26.

Bonaparte, J. F. (1990). New Late Cretaceous mammals from the Los alamitos Formation, northern Patagonia. National Geographic Research *6*, 63 - 93.

Bousquet, J. e. a. (1992). Extensive variation in the evolutionary rate of *rbc*L gene sequence among seed plants. Proceedings of the National Academy of Sciences, USA *89*, 7844 - 7848.

Brain. (1984). The terminal Miocene event: a critical environmental and evolutionary episode? In Late Cainozoic palaeoclimates of the Southern Hemisphere, J. C. Vogel, ed. (Rotterdam: Balkema), pp. 5 - 20.

Bremer, K. (1992). Ancestral areas - a cladistic reinterpretation of the center of origin concept. Systematic Biology *41*, 436 - 445.

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

Brundin, L. (1966). Transantarctic relationships and their significance, as evidenced by chironomid midges, Volume 4.

Burbidge, N. T. (1960). The phytogeography of the Australian region. Australian Journal of Botany *8*, 75 - 209.

Burrows, C. J. (1995). Germination behaviour of seeds of the New Zealand species *Fuchsia excorticata*, *Griselinia littoralis*, *Macropiper excelsum*, and *Melicytus ramiflorus*. New Zealand Journal of Botany *33*, 131 - 140.

Cain, M. L., Milligan, B. G., and Strand, A. E. (2000). Long-distance seed dispersal in plant populations. American Journal of Botany *87*, 1217 - 1227.

Cain, S. A. (1944). Foundations of Plant Geography (New York: Harper & Row).

Cain, M. L., Milligan, B. G., and Strand, A. E. (2000). Long-distance seed dispersal in plant populations. American Journal of Botany *87*, 1217 - 1227.

Carlquist, S. (1966). The biota of long-distance dispersal. I. Principles of dispersal and evolution. Quarterly Review of Biology *41*, 247 - 270.

Carlquist, S. (1967). The biota of long-distance dispersal. V. Plant dispersal to Pacific islands. Bulletin of the Torrey Botanical Club *94*, 129 - 162.

Cavanilles, J. (1797). . In Icones, pp. 49.

Charlesworth, D., Liu, F. L., and Zhang, L. (1998). The evolution of the alcohol dehydrogenase family by loss of introns in plants of the genus *Leavenworthia* (Brassicaceae). Molecular Biology and Evolution *15*, 552 - 559.

Chase, M. W., and Albert, V. A. (1998). A perspective on the contribution of plastid rbcL DNA on sequences to angiosperm phylogenetics. In Molecular Systematics of Plants II: DNA sequencing, D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. (London: Kluwer academic publishers), pp. 488 - 507.

Chase, M. W., *et al.* (1993). Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*l. Annals of the Missouri Botanical Garden *80*, 528 - 580.

Cifelli, R. L. (1990a). Cretaceous mammals of southern Utah. I. Marsupials from the Kaiparowits Formation Judithian. Journal of Vertebrate Paleaontology *10*, 295 - 319.

Cifelli, R. L. (1990b). Cretaceous mammals of southern Utah. II. Marsupials and marsupial-like mammals from the Wahweap Formation early Campanian. Journal of vertebrate Paleontology *10*, 295 - 319.

Cifelli, R. L. (1993). Theria of metatherian-eutherian grade and the origin of marsupial orders. In Mammal Phylogeny, F. Szalay, M. J. Novacek and M. C. McKenna, eds. (New York: Springer-Verlag).

Clegg, M. T., and Zurawski, G. (1992). Chloroplast DNA and the study of phylogeny: Status and future prospects. In Molecular Systematics of Plants, P. S. Soltis, D. E. Soltis and J. J. Doyle, eds. (New York: Chapman and Hall), pp. 1 - 13.

Cockayne, L., and Allan, H. H. (1927). Notes on New Zealand floristic botany, including descriptions of new species (No. 5). Transactions of the New Zealand Institute *57*, 48 - 72.

Corner, E. J. H. (1976). The seeds of dicotyledons (Cambridge: Cambridge University Press).

Cox, C. B., and Moore, P. (2000). Biogeography, 6th Edition (London: Blackwell Science).

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

Cranston, P. S., and Naumann, I. D. (1991). Biogeography. In The insects of Australia, I. D. Nauman, ed. (Melbourne: Melbourne University Press), pp. 180 - 197.

Cranwell, L. M. (1963). *Nothofagus*: living and fossil. In Pacific basin biogeography, J. L. Gressitt, ed. (Honolulu: Bishop Museum Press), pp. 387 - 400.

Crayn, D. M., Kron, K. A., and Gadeck, P. A. (1996). Delimitation of Epacridaceae: Preliminary molecular evidence. Annals of Botany *77*, 317 - 321.

Crisci, J. V., Cigliano, M. M., Morrone, J. J., and Roig-Junent, S. (1991). Historical biogeography of southern South America. Systematic Zoology *40*.

Crisp, M. D., Linder, H. P., and Weston, P. H. (1995). Cladistic biogeography of plants in Australia and New Guinea: congruent pattern reveals two endemic tropical tracks. Systematic Biology *44*, 457 - 473.

Crisp, M. D., West, J. G., and Linder, H. P. (1999). Biogeography of the terrestrial flora. In Flora of Australia, A. Orchard, ed. (Melbourne: ABRS/SIRD), pp. 332 - 367.

Croizat, L. (1952). Manual of phytogeography: or an account of plant-dispersal throughout the world, W. Junk, ed. (The Hague.

Croizat, L. (1958). Panbiogeography: or, an introductory synthesis of zoogeography, phytogeography, and geology: with notes on evolution, systematics, ecology, anthropology etc., L. Croziat, ed. (Caracas.

Cronquist, A. (1973). The evolution and classification of flowering plants (Boston: Houghton Mifflin).

Cronquist, A. (1981). An integrated system of classification of flowering plants (New York: Columbia University Press).

Cronquist. (1988). The evolution and classification of flowering plants, 2 Edition (New York: New York Botanical Garden).

Curtis, W. M. (1956). Eucryphiaceae, Tasmania, ed. (Hobart: Government Printer), pp. 180.

Dahlgren, R. M. T. (1980). A revised system of classification of the angiosperms. Botanical Journal of the Linnean Society *80*, 91 - 124.

Daly, G. T. (1967). Non-leguminous nitrogen-fixers in New Zealand. Canterbury Chambers Comm. Agr. Bull. No. 449.

Darlington, P. J. (1965). Biogeography of the southern end of the world (Massachusetts: Harvard University Press).

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

Demesure, B., Sodzin, N., and Petit, R. J. (1995). A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology *4*, 129-131.

de Miera L., and de la Vega, M. P. (1998). A comparative study of vicilin genes in *Lens*: Negative evidence of concerted evolution. Molecular Biology and Evolution *15*, 303 - 311.

Desmond, A., and Moore, J. (1992). Darwin: The life of a tormented evolutionist (New York: Warner Books, Inc.).

Dettmann, M. E. D., Pocknall, D. T., Romero, E. J., and Zamaloa, M. C. (1990). *Nothofagidites* Erdtman ex Potonie, 1960: a catalogue of species with notes on the paleogeographic distribution of *Nothofagus* Bl. (Southern Beech) (New Zealand.

Dickison, W. C. (1978). Comparative anatomy of Eucryphiaceae. American Journal of Botany *65*, 722 - 735.

Dickison, W. C., and Rutishauser, R. (1990). Developmental morphology of stipules and systematics of the Cunoniaceae and presumed allies. II. Taxa without interpetiolar stipules and conclusions. Botanica Helvetica *100*, 75 - 95.

Dillon, M. O., and Muñoz-Schick, M. (1993). A revision of the dioecious genus *Griselinia* (Griseliniaceae), including a new species from the coastal Atacama Desert of northern Chile. Brittonia *45*, 261 - 274.

Downie, S. R. and J. D. Palmer (1992). Use of chloroplast rearrangements in reconstructing plant phylogeny. In Molecular Systematics of Plants, P. S. Soltis, D. E. Soltis and J. F. Doyle, eds. (New York: Chapman and Hall), pp. 14 - 35.

Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. Focus *12*, 13 - 15.

Doyle, J. J., Kanazin, V., Shoemaker, R. C., and Williams, D. M. (1994). Phylotenetic utility of histone H3 sequencesin the perrenial relatives of soybean (Glycine: Leguminosae). Moleuclar Phylogetics and Evolution *6*, 438-447.

Dress, W. J. (1956). A review of the genus *Eucryphia*. Baileya *4*, 116 - 122.

Eaton, J. G. (1993). Marsupial dispersal. National Geographic Research and Exploration *9*, 436 - 443.

Edwards, D. (1990). Constraints on Silurian and early Devonian phytogeogra;hic analysis based on megafossils. In Palaezoic Palaeogeography and Biogeography, W. S. McKerrow and C. R. Scotese, eds.: Geological Society Memoir), pp. 233 - 242.

Emshwiller, E., and Doyle, J. J. (1999). Chloroplast-expressed glutamine synthetase (ncpGS): Potential utility for phylogenetic studies with an example from Oxalis (Oxalidaceae). Molecular Phylogenetics and Evolution *12*, 310-319.

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

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

Ferguson, I. K., and Hideux, M. J. (1978). Some aspects of the pollen morphology and its taxonomic significance in Cornaceae *sens. lat.* In Proceedings of the IV international palynological Conference, Lucknow, ed., pp. 240 - 249.

Ferris, C., King, R. A., and Gray, A. J. (1997). Molecular evidence for the maternal parentage in the hybrid origin of *Spartina anglica*. Molecular Ecology *6*, 185 - 187.

Ferris, C., Oliver, R. P., and Davy, A. J. (1993). Native oak chloroplasts reveal an ancient divide accross Europe. Molecular Ecology *2*, 337 - 344.

Flook, P. K. S., Klee, S., and Rowell, C. H. F. (1999). Combined molecular phylogenetic analysis of the Orthoptera (Arthropoda, Insecta) and implications for their higher systematics. Systematic Biology *48*, 233 - 253.

Forster, P. I., and Hyland, B. P. M. (1997). Two new species of *Eucryphia* Cav. (Cunoniaceae) from Queensland. Austrobaileya *4*, 589 - 596.

Futuyma, D., J. (1986). Evolutionary Biology, Second Edition (Sunderland, Massachusetts: Sinauer Associates Inc.).

Gandolfo, M. A., and Romero, E. J. (1992). Leaf morphology and a key to species of *Nothofagus*. Bulletin of the Torrey Botanical Club, 152 - 166.

Gardes, M., and Bruns, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. Molecular Ecology *2*, 113 - 118.

Gardiner, B. G. (1997). Picture Quiz [J.D. Hooker]. Journal of the Linnean Society *13*, 16-22.

Garg, M. (1980). Pollen morphology and systematic position of *Coriaria*. Phytomorphology *30*, 5 - 10.

Gaut, B. S. (1998). Molecular clocks and nucleotide substitution rates in higher plants. Evolutionary Biology *30*, 93 - 120.

Gaut, B. S., Muse, S. V., and Clegg, M. T. (1992). Relative rates of nucleotide substitution at the *rbc*L locus of monocotyledonous plants. Journal of molecular evolution *35*, 292 - 303.

Gayet, M., Marshall, L. G., and Sempere, T. (1991). The Mesozoic and Paleocene vertebrates of Bolivia and their stratigraphic context: a review. Revista Tecnica de YPFB *12*, 339 - 433.

Gielly, L., and Taberlet, P. (1994). The use of chloroplast DNA to resolve plant phylogenies: non-coding versus *rbc*L sequences. Molecular Biology and Evolution *11*, 769 - 777.

Gielly, L., and Taberlet, P. (1996). A phylogeny of the European gentians inferred from chloroplast *trn*L (UAA) intron sequences. Botanical Journal of the Linnean Society *120*, 57 - 75.

Gilg, E. (1925). Eucryphiaceae. In Engler and Prantl, Natürl. Pflanzenfam.

Gillham, N. W. (1994). Organelle Genes and Genomes (New York: Oxford University Press).

Godley. (1975). Biogeography and Ecology of New Zealand. In Flora and vegetation (Amsterdam: Junk), pp. 177- 229.

Godthelp, H., Archer, M., and Cifellir, R. (1992). Earliest known Australian tertiary mammal fauna. Nature *356*, 514 - 516.

Goldberg, A. (1986). Classification, Evolution and Phylogeny of the Families of Dicotyledons. In Smithsonian Contribution to Botany (Smithsonian Institution Press: Washington).

Goldblatt, P. (1976). New or noteworthy chromosome records in the Angiosperms. Annals of the Missouri Botanical Gardens *63*, 889 - 895.

Good, R. D. O. (1974). The geography of the flowering plants (London: Longman).

Good, R. D. O. (1930). The geography of the genus *Coriaria*. New Phytologist *29*, 170 - 198.

Gottlieb, L. D., and Ford, V. S. (1996). Phylogenetic relationships among the sections of *Clarkia* (Onagraceae) inferred from the nucleotide sequences of *PgiC*. Systematic Botany *21*, 45-62.

Grant-Mackie, J. A. (1979). Cretaceous - Recent plate tectonic history and paleoceanographic development of the Southern Hemisphere. In Proceedings of the International Symposium on marine biogeography and evolution in the southern hemisphere (Aukland, New Zealand: N. Z. DSIR Inform. Ser.), pp. 27 - 42.

Gray, A. (1857). On the botany of Japan. Mem. Amer. Acad. Sc. New Ser. 6, 383.

Green, P. T. (1999). Greta's Garbo: stranded seeds and fruits from Greta Beach, Christmas Island, Indian Ocean. Journal of Biogeography *26*, 937 - 946.

Gregor, H.-J. (1980). Seeds of the genus *Coriaria* Linne (Coriariaceae) in the European Neogene. Tertiary Research *3*, 61 - 69.

Guédes, M. (1971). Carpel peltation and syncarpy in *Coriaria ruscifolia* L. New Phytologist, 213 - 227.

Gunn, C. R., and Dennis, D. J. (1976). World guide to tropical drift seeds and fruits (New York: Quadrangle/The New York Times Book Company).

Guppy, H. B. (1906). Plant dispersal, Volume 2 (London: Macmillan).

Guppy, H. B. (1917). Plants, seeds and currents in the West Indies and Azores (London: Williams and Norgate).

Hamby, K. R., and Zimmer, E. A. (1992). Ribosomal RNA as a phylogenetic tool in plant systematics. In Molecular Systematics of Plants, P. Soltis, D. Soltis and J. Doyle, eds. (New York: Chapman & Hall), pp. 50 - 91.

Hamel, J. L. (1956). Contribution a l'étude caryo-taxinomique des eucryplhiacées. Bulletin du Muséum, 2^{e} série *XXXI*, 526 - 535.

Hanks, S. L., and Fairbrothers, D. E. (1976). Palynotaxonomic investigation of *Fagus* L. and *Nothofagus*, Bl.: light microscopy, scanning electron microscopy, and computer analysis. In Botanical systematics, V. H. Heywood, ed. (London: Academic Press), pp. 1 - 142.

Harden, G. J. (1991). Eucryphiaceae. In Flora of New South Wales, G. J. Harden, ed. (Kensington: New South Wales University Press), pp. 523 - 524.

Harris, S. A., and Ingram, R. (1991). Chloroplast DNA and biosystematics - the effects of intraspecific diversity and plastid transmission. Taxon *40*, 393 - 412.

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

Helfman, G. S., Collette, B. B., and Facey, D. E. (1997). The Diversity of fishes: Blackwell Science).

Hemsley, W. B. (1894). C. terminalis, pl. 2220. In W. J. Hooker's Icones Plantarm, D. Oliver, ed. (London: Longmans).

Herbert, D. A. (1953). The climatic sifting of Australian vegetation. ANZAAS Rep. *22*, 349 -370.

Herbert, D. A. (1967). Ecological segregation and Australian phytogeographic elements. Proceedings of the royal society Queensland *78*, 101 - 111.

Herbert, D. A. (1950). Present day distribution and the geological past. Victorian Naturalist *66*, 227 -232.

Herendeen, P. S. (1992). The fossil history of the Leguminosae from the Eocene of Southeastern North America. In Advances in Legume Systematics 4. The fossil record.

Hershkovitz, M. A., and Zimmer, E. A. (2000). Ribosomal DNA evidence and disjunctions of western American Portulacaceae. Molecular Phylogenetics and Evolution *15*, 419 - 439.

Hideux, M. J., and Ferguson, I. K. (1976). The stereostructure of the exine and its evolutionary significance in Saxifragaceae *sensu lato*. In The evolutionary significance of the exine: Linnean Society Symposium Series 1).

Hill, R. S. (1991a). Leaves of *Eucryphia* (Eucryphiaceae) from Tertiary sediments in southeastern Australia. Australian Systematic Botany *4*, 481 - 497.

Hill, R. S. (1991b). Tertiary *Nothofagus* (Fagaceae) macrofossils from Tasmania and Antarctica and their bearing on the evoltuion of the genus. Botanical Journal of the Linnean Society *105*, 73 - 112.

Hill, R. S. (1994). The history of selected Australian taxa. In History of the Australian Vegetation: Cretaceous to recent (Cambridge: Cambridge University Press).

Hill, R. S., and Carpenter, R. J. (1991). Evolution of Acmopyle and Dacrycarpus (Podocarpaceae) foliage as inferred from macrofossils in south-eastern Australia. Australian Systematic Botany *4*, 449 - 479.

Hill, R. S., and Read, J. (1991). A revised infrageneric classification of Nothofagus (Fagaceae). Botanical Journal of the Linnean Society *105*, 37 - 72.

Hill, R. S., Read, J., and Busby, J. R. (1988). The temperature-dependence of photosynthesis of some australian temperate rainforest trees and its biogeographical significance. Journal of Biogeography *15*, 431 - 49.

Hillis, D., and Dixon, M. T. (1991). Ribosomal DNA: molecular and phylogenetic inference. Quarterly Review of Biology *66*, 410 - 453.

Hillis, D. M., Moritz, C., and Mable, B. (1996). Molecular systematics (Sunderland, Massachusettes: Sinauer Associates).

Hodgeson, C. J., and Henderson, R. C. (2000). Fauna of New Zealand
Coccidae (Insecta: Hemiptera: Coccoidea). *Fauna of New Zealand 41*, Volume 41, T. Crosby, ed. (Auckland: Manaaki Whenua Press).

Hoffmann, J. (1997). Flora Silvestre de Chile; zona araucana; arboles, arbustos y enredaderas leñosas (Santiago, Chile: Imprenta Salesianos A.A. General Bulnes 19).

Holdgate, M. W. (1961). Vegetation and soil in the south Chilean islands. Journal of Ecology *49*, 559 - 580.

Hooker, J. D. (1853). Introductory Essay to the Flora of New Zealand. In Botany of the Antarctic Voyage of H.M. Discovery Shiips Erebus and Terror in the years 1839 -1843, III. Flora Tasmaniae (London: Lovell Reeve), pp. i - cxxviii.

Hooker, J. D. (1860). Botany of the Antarctic voyage of H. M. discovery ships Erebus and Terror in the years 1839 - 1843, Volume III (London: Lovell Reeve).

Hooker, J. D. (1864). Coriariaceae. In Handbook of the New Zealand Flora (London: Reeve & Co.).

Hoot, S. B., and Douglas, A. W. (1998). Phylogeny of the Proteaceae based on *atp*B and *atp*B-*rbc*L intergenic spacer region sequences. Australian Systematic Botany *11*, 301 - 320.

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

Howe, H. F., and Smallwood, J. (1982). Ecology of seed dispersal. Annual Review of Ecology and Systematics *13*, 201 - 228.

Hufford, L., and Dickison, W. (1992). A phylogenetic analysis of Cunoniaceae. Systematic Botany *17*, 181 - 200.

Humphries, C. J. (1981). Biogeographical methods and the southern beeches (Fagaceae: *Nothofagus*). In Advances in cladistics: proceedings of the first meeting of the Hennigian Society, V. A. Funk and D. R. Brooks, eds. (New York: New York Botanical Garden), pp. 177 - 207.

Humphries, C. J., and Parenti, L. (1999). Cladistic biogeography, Second Edition (Oxford: Oxford University Press).

Humphries, C. J., and Parenti, L. R. (1986). Cladistic biogeography (Oxford: Clarendon Press).

Hurr, K. A., Lockhart, P. J., Heenan, P. B., and Penny, D. (1999). Evidence for the recent dispersal of *Sophora* (Leguminosae) around the Southern Oceans: molecular data. Journal of Biogeography *26*, 565 - 577.

Hutchinson, J. (1973). The families of floweing plants (Oxford: Oxford University Press).

Isoda, K., Shiraishi, S., Watanabe, S., and Kitamura, K. (2000). Molecular evidence of natural hybridization between *Abies veitchii* and *Ahomlepis* (Pinaceae) revealed by chloroplast, mitochondrial and nuclear DNA markers. Molecular Ecology *9*, 1965 - 1974.

Jay, M. (1968). Distribution des flavonoides chez les Cunoniacées. Taxon *17*, 489 - 495.

Jessup, L. W. (1994). Eucryphiaceae. In Queensland Vascular Plants, Names and Distribution, R. J. F. Henderson, ed. (Brisbane: Queensland Department of Environment & Heritage).

Johnson, B. D., Powell, C. M., and Veevers, J. J. (1980). Early spreading of the Indian Ocean between India and Australia. Earth Planet. Sci. Lett. *47*, 113 - 143.

Johnson, L. A. S. (1998). Proteaceae - where are we? Australian Systematic Botany *11*, 251 - 255.

Johnson, L. S., and Soltis, D. E. (1995). Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *mat*K sequences. Annals of the Missouri Botanical Garden *82*, 149 - 175.

Johnson, S. D., and Steiner, K. E. (2000). Generalization versus specialization in plant pollination systems. Trends in Ecology and Evolution *15*, 140 - 143.

Johnston, I. M. (1929). Papers on the flora of northern Chile. Contr. Gray Herb. *4*, 1 - 172.

Jordan, G. J., Carpenter, R. J., and Hill, R. S. (1991). The macrofossil record of Proteaceae in Tasmania: a review with new species. Australian Systematic Botany *11*, 465 - 501.

Kalkman, C. (1988). The phylogeny of the Rosaceae. Botanical Journal of the Linean Society *98*, 37 - 59.

Kennett, J. P. (1977). Cenozoic evolution of tAntactic glaciation, the circum-Antarctic Ocean, and their impact of global paleoceanography. Journal of Geophysical Research *82*, 3843 - 3860.

Kirsch, J. A. W., Lapointe, F.-J., and Springer, M. S. (1997). DNA-hybridisation studies of marsupials and their implications for metatherian classification. Aurstalian Jounal of Zoololgy *45*, 211 - 280.

Kosuge, K., Swada, K., Denda, T., Adachi, J., and Watanabe, K. (1995). Phylogenetic relationships of some genera in the Ranunculaceae based on alcohol dehydrogenase genes. Plant Systematics and Evolution *Suppl. 9*, 263-271.

Kron, K. A. (1996). Phylogenetic relationships of Empetraceae, Epacridaceae, Ericaceae, Monotropaceae, and Pyrolaceae: Evidence from nuclear ribosomal 18s sequence data. Annals of Botany *77*, 293 - 303.

Kruglyak, S., Durrett, R., Schug, M. D., and Aquardo, C. F. (2000). Distribution and abundance of microsatellites in the yeast genome can be explained by a balance between slippage events and point mutations. Molecular Biology and Evolution *17*, 1210 - 1219.

Kubitzki, K. (1993). Fagaceae. In The families and genera of vascular plants, J. G. Rohwer and V. Bittrich, eds. (Berlin: Springer - Verlag), pp. 301 - 309.

Kurochkin, E. N. (1995). Morphological differentiation of palaeognathous and neognathous birds. Cour. Forschunginst. Senckenb. *181*, 79 - 88.

LaBrecque, J. L., and Barker, P. (1981). The age of the Weddell basin. Nature *290*, 489 - 492.

Laroche, J., Peng, L., Laurent, M., and Bousquet, J. (1997). Molecular evolution of angiosperm mitochondrial introns and exons. Proceedings of the National Acadamy of Sciences, USA *89*, 5722 - 5727.

Lavin, M., Eshbaugh, E., Hu, J. M., Mathews, S., and Sharock, R. A. (1998). Monophyletic subgroups of the tribe Millettieae (Leguminosae) as revealed by phytochrome nucleotide sequence data. American Journal of Botany *85*, 412-433.

Li, W.-H., and Tanimura, M. (1987). The molecular clock runs more slowly in man than in apes and monkeys. Nature *326*, 93 - 96.

Liao, D. Q. (2000). Gene conversion drives within genic sequences: concerted evolution of ribosomal RNA genes in bacteria and archaea. Journal of Molecular Evolution *51*, 305 - 317.

Linder, H. P., and Crisp, M. D. (1995). *Nothofagus* and Pacific biogeography. Cladistics *11*, 5-32.

Lingen, G. J., Swanson, K. M. (1994). Evolution of the Tasman Sea basin (Rotterdam: Balkema).

Lu, M.-Z., Szmidt, A. E., and Wang, X.-R. (1998). RNA editing in gymnosperms and its impact on the evolution of the mitochondrial coxI gene. Plant Molecular Biology *37*, 5722 - 5727.

Lüning, K. (1990). Temperate and Polar Seaweed Vegetation of the Southern Hemisphere. In Seaweeds their environment, biogeography and ecophysiology, C. Yarish and H. Kirkman, eds. (New York: John Wiley & Sons, Inc.), pp. 235 - 233.

Macphail, M. K. (1997). Comment on M. Pole (1994): 'New Zealand flora - entirely long distance dispersal?'. Journal of Biogeography *24*, 113-117.

Manos, P., S (1997). Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): Taxonomic congruence with morphology and plastid sequences. American Journal of Botany *84*, 1137 - 1155.

Manos, P. S., and Steele, K. P. (1997). A phylogenetic analysis of 'higher Hamamelididae based on morphology and sequences from plastid and nuclear DNA. American Journal of Botany *10*, 1407 - 1419.

Martin, P. G., and Dowd, J. M. (1993). Using sequences of *rbc*L to study phylogeny and biogeography of *Nothofagus* species. Australian Systematic Botany *6*, 441-447.

Mathews, S., Lavin, M., and Sharrock, R. A. (1995). Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. Annals of the Missouri Botanical Garden *82*, 296 - 321.

Mathews, S., and Sharrock, R. A. (1997). Phytochrome gene diversity. Plant Cell and Environment *20*, 666 - 671.

Maximowicz, C. J. (1881). Coriaria ilice et monochasmate. Mem. Acad. Imp. Sciences de St. Petersbourg, VIIth. Ser. *29*, 113.

McDade, L. A., and Moody, M. L. (1999). Phylogenetic relationships among Acanthaceae: Evidence from noncoding *trn*L-*trn*F chloroplast DNA sequences. American Journal of Botany *86*, 70 - 80.

McDowall, R. M. (1978). Generalized tracks and dispersal in biogeography. Systematic Zoology *27*, 88 - 104.

Melville, R. (1973). Continental drift and plant distribution. In Implications of continental drift to the earth sciences, D. H. Tarling and S. K. Runcorn, eds. (London: Academic Press), pp. 439-446.

Mercer, J. H. (1983). Cenozoic glaciation in the Southern Hemisphere. Annual Review of Earth Planet Sciences *11*, 99 - 132.

Metcalfe, C. R., and Chalk, L. (1950). Anatomy of the Dicotyledons (Oxford: Clarendon Press).

Metcalfe, C. R., and Chalk, L. (1979). Anatomy of Dicotyledons. 2nd Edition (Oxford: Clarendon Press).

Mildenhall, D. C. (1980). New Zealand Late Cretaceous and Cenozoic plant biogeography: a contribution. Palaeogeography, Palaeclimatology, Palaeoecology *31*, 197 - 233.

Miller, C. N. (1977). Mesozoic conifers. Botanical Review *43*, 217-283.

Miller, C. N. (1988). The origin of modern conifer families. In Origin and evolution of gymnosperms, C. B. Beck, ed. (New York: Columbia University Press), pp. 448-486.

Monge-Najera, J. (1995). Phylogeny, biogeography and reproductive trends in the *Onychophora*. Zoological Journal of the Linnean Society *114*, 21 - 60.

Moore, W. S., and DeFilippis, V. R. (1997). The window of taxonomic resolution for phylogenies based on mitochondrial cytochrome b. In Avian molecular evolution and systematics, D. P. Mindell, ed. (San Diego: Academic Press), pp. 83 - 120.

Morgan, D. R., and Soltis, D. E. (1993). Phylogenetic relationships among members of saxifragaceae *sensu* lato based on *rbc*L sequence data. Annals of the Missouri Botanical Garden *80*, 631 - 660.

Mourer - Chauvire, C. B., Senut, B., Pickford, M., and Mein, P. (1996). Le plus ancien representant du genre Struthio (Aves, Struthionidae), Struthio coppensi n. sp., du Miocene inferieure de Namibie. C. R. Acad. Sci. Paris *322*, 325 - 332.

Muray, D. R. (1986). Seed dispersal by water. In Seed Dispersal (Australia: Academic Press).

Muse, S. V. (2000). Examining rates and pattens of nucleotide substitution in plants. Plant Molecular Biology *42*, 25-43.

Muse, S. V., and Weir, B. S. (1992). Testing for equality of evolutionary rates. Genetics *1*, 269 - 276.

Nelson, G., and Rosen, D. E. (1981). Vicariance Biogeography: a critique (New York: Columbia University Press).

Nickrent, D. L., and Soltis, D. E. (1995). A comparison of angiosperm phylogenies from nuclear 18S rDNA and *rbc*L sequences. Annals of the Missouri Botanical Garden *82*, 208 - 234.

Nixon, K. C. (1989). The origin of Fagaceae: evolution, systematics, and fossil history of the Hamamelidae. In Systematic Association Special Volume, P. R. Crane and S. Blackmore, eds. (London: Clarendon Press), pp. 23 - 43.

Norberg, R. A. (1973). Autorotation, self-stability, and structure of single winged fruits and seeds (samaras) with comparative remarks on animal flight. Biological Reviews *48*, 561 - 596.

Oginuma, K., Nakath, M., Suzuki, M., and Tobe, H. (1991). Karyomorphology of *Coriaria* (Coriariaceae) Taxonomic implications. Botanical Magazine *103*, 297 - 308.

Ohsawa, T., Nishida, H., and Nishida, M. (1995). Yezonia, a new section of *Araucaria* (Araucariaceae) based on permineralized vegetative and reproductive organs of *A. vulgaris comb. nov.* from the upper Cretaceous of Hokkaido, Japan. Journal of Plant Research *108*, 25-39.

Oliver, W. R. B. (1942). The genus *Coriaria*. New Zealand Recordings of the Dom.. Mus, i, Wellington, 27.

Olmstead, R. G., and Palmer, J. D. (1994). Chloroplast DNA systematics: a review of methods and data analysis. American Journal of Botany *81*, 1205 - 1244.

Olmstead, R. G., and Reeves, P. A. (1995). Evidence for the polyphyly of the Scrophulariaceae based on chloroplast *rbc*L and *ndh*F sequences. Annals of the Missouri Botanical Garden *82*, 176 - 193.

Olsen, K. M., and Schaal, B. A. (1999). Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proceedings of the National Academy of Sciences USA *96*, 5586 - 5591.

Owen. (1983). Atlas of continental displacement, 200 million years ago to the present (Cambridge: University Press).

Ozaki, K. (1991). Late Miocene and Pliocene floras in Central Honshu (Yokohama: Kanagawa Pref. Mus.).

Page, R., and Holmes, E. (1998). Molecular evolution a phylogenetic approach (Oxford: Blackwell Science Ltd.).

Palmer, J. D. (1992). Mitochondrial DNA in plant systematics: applications and limitations. In Molecular Systematics of Plants, P. S. Soltis, D. E. Soltis and J. J. Doyle, eds. (New York: Chapman & Hall), pp. 36 - 49.

Parkinson, C. G. (1999). Extraordinary changes in mutation rates in *Pelargonium* E. In XVI International Botanical Conference St. Louis, USA.
Penny, D., Hendy, M. D., and Henderson, I. M. (1987). Reliability of evolutionary trees. Cold Spring Harbour Symposium of Quaternary Biology *52*, 857 - 862.

Philipson, W. R. (1967). *Griselinia* Forst. fil. - Anomaly or link. New Zealand Journal of Botany *5*, 134 - 165.

Philipson, W. R. (1977). Ovular morphology and the classification of the dicotyledons. Plant Systematics and Evolution *Supplement 1*, 123 - 140.

Philipson, W. R., and Philipson, M. N. (1988). A classification of the genus *Nothofagus* (Fagaceae). Botanical Journal of the Linnean Society *98*, 27-36.

Platnick, N. I., and Nelson, G. (1978). A method of analysis for historical biogeography. Systematic Zoology *27*, 1 - 16.

Plunkett, G. M., Soltis, D. E., and Soltis, P. S. (1996). Higher level relationships involving Apiales (Apiaceae and Araliaceae) based on phylogenetic analysis of *rbc*L sequences. American Journal of Botany *83*, 499 - 515.

Poeppig, and Endlicher. (1839). Nova Genera ac Species Plantarum, Volume 2.

Poiret, J. L. M. (1804). *C. microphylla*. In Encyclopedia methodique Botanique, 6. (Paris: Chez Panouke).

Pole, M. S. (1994). The New Zealand flora - entirely long distance dispersal. Journal of Biogeography *21*, 625-635.

Powell, J. M., Crayn, D. M., Gadek, P. A., Quinn, C. J., Morrison, D. A., and Chapman, A. R. (1996). A re-assessment of relationships within Epacridaceae. Annals of Botany *77*, 305 - 315.

Praglowski, J. (1970). Coriariaceae. In World pollen flora, Erdtman, ed. (Copenhagen, pp. 15 - 22.

Premoli, A. C. (1996). Leaf architecture of South American Nothofagus (Nothofagaceae) using traditional and new methods in morphometrics. Botanical Journal of the Linnean Society *121*, 25-40.

Quilty, P. G. (1994). The background: 144 million years of Australian palaeoclimate and palaeogeography. In History of the Australian Vegetation: Cretaceous to Recent, R. S. Hill, ed. (Cambridge: Cambridge University Press), pp. 14 - 43.

Radford, A. E., Dickison, W. C., Massey, J. R., and Bell, C. R. (1974). Vascular Plant Systematics (London: Parper & Row Publishers).

Raven, P. H., and Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. Annals of the Mosouri Botanical Garden *61*, 539 - 673.

Raven, P. H., and Polhill, R. M. (1981). Biogeography of the Leguminosae, Volume 2: Royal Botanic Gardens, Kew).

Read, J., and Busby, J. R. (1990). Comparative responses to temperature of the major canopy species of Tasmanian cool temperate rainforest and their ecological significance II. Net photosynthesis and climate analysis. Australian Journal of Botany *38*, 185 - 205.

Renner, S. S. (2000). Timing transantarctic disjunctions in the Atherospermataceae (Laurales): Evidence from coding and noncoding chloroplast sequences. Systematic Biology *49*, 579 - 591.

Richardson, J. E., Weitz, F. M., Fay, M. F., Cronk, Q. C. B., Linder, H. P., Reeves, G., and Chase, M. (2001). Phylogenetic analysis of *Phylica* L. (Rhamnaceae) with an emphasis on island species: evidence from plastid *trn*L-F DNA and nuclear internal transcribed spacer (ribosomal DNA) sequences. Taxon *In press*.

Ridley. (1930). The dispersal of plants throughout the world (London, UK: Reeve).

Rieseberg, L. H., and Brunsfeld, S. J. (1992). Molecular evidence and plant introgression. In Molecular systematics of plants, P. S. Soltis, D. E. Soltis and J. J. Doyle, eds. (New York: Chapman & Hall), pp. 151 - 176.

Rodriguez, R. R., Matthei, O. S., and Quezada, M. M. (1983). Flora Arbórea de Chile, (Universidad de Concepción: Chile.).

Rosen, D. E. (1978). Vicariant patterns and historical explanation in biogeography. Systematic Zoology *27*, 159 - 188.

Rozefelds, A. C., and Drinnan, A. N. (1994). Ontogeny and floral diversity of staminate flowers in the southern beeches (*Nothofagus*). Australian Systematic Botany, 28-29.

Rundel, P. W., Dillon, M. O., Palma, B., Mooney, H. A., Gulmon, S. L., and Ehleringer, J. R. (1991). The phytogeography and ecology of the coastal Atacama and Peruvian deserts. Aliso *13*, 1 - 50.

Sanderson, M. J. (1998). Estimating rate and time in molecular phylogenies: beyond the molecular clock. In Molecular Systematics of Plants II: DNA sequencing, D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. (Boston: Kluwer Academic Publishers), pp. 242- 264.

Sanderson, M. J., and Doyle, J. J. (1992). Reconstruction of organismal gene phylogenies from data on multigene families: concerted evolution, homoplasy and confidence. Systematic Biology *41*, 4-17.

Savard, L. *et al.* (1994). Chloroplast and nuclear gene sequences indicate late Pennsylvanian time for the last common ancestor of extant seed plants. Proceedings of the National Acadamy of Sciences *89*, 5163 - 5176.

Schlessman, M. A., Lloyd, D. G., and Lowry, P. P. (1990). Evolution of sexual systems in New Caledonia Araliaceae. Memoirs of the New York Botanical Garden *55*, 105 - 117.

Schodde, R. (1969). A monograph of the family Atherospermataceae R. Br. (Adelaide: University of Adelaide). Australia.

Scholz, H. (1964). Sapindales. In A. Engler's Syllabus der Pflanzenfamilien, H. Melchior, ed. (Berlin: Borntraeger), pp. 277 - 288.

Setoguchi, H., *et al.* (1998). Phylogenetic relationships within Aracariaceae based on *rbc*L gene sequences. American Journal of Botany *85*, 1507 - 1516.

Sharma, V. K. (1968). Floral morphology, anatomy and embryology of *C. nepalensis* Wall., with a discussion on the inter-relatonships of the family Coriariaceae. Phytomorphology *18*, 143 - 153.

Shih, M. C., Heinrich, P., and Goodman, H. M. (1991). Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase from Arabidopsis thaliana. Gene *104*, 133-138.

Simpson, G. G. (1944). Tempo and mode in evolution (New York: Columbia University Press).

Skog, L. E. (1972). The genus *Coriaria* (Coriariaceae) in the W. Hemisphere. Rhodora *74*, 242 - 253.

Soltis, D. E., and Soltis, P. S. (1998). Choosing an approach for an appropriate gene for phylogenetic analysis. In Molecular Systematics of Plants II: DNA sequencing, D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. (Boston: Kluwer Academic Publishers), pp. 1 - 42.

Soltis, D. E., and Soltis, P. S. (1997). Phylogenetic relationships in Saxifragaceae *sensu lato*: A comparison of topologies based on 18S rDNA and *rbc*L sequences. American Journal of Botany *84*, 504 - 522.

Soltis, D. E., Soltis, P. S., Morgan, D. R., Swenson, S. M., Mullin, B. C., Dowd, J. M., and Martin, P. G. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for systematic nitrogen fixation in angiosperms. Proceedings of the National Academy of Sciences, USA *92*, 2647 - 2651.

Soltis, P. S., Soltis, D. E., and Doyle, J. (1992). Molecular Systematics of Plants (New York: Chapman and Hall).

Sorensen, A. E. (1986). Seed dispersal by adhesion. Annual Review of Ecology and Systematics *17*, 443 - 463.

Spach. (1836). Histoire naturelle des végétaux. Phanérogames, pp. 345 - 346.

Springer, M., Kirsch, J. A. W., and Chase, J. A. (1997). The chronicle of marsupial evolution. In Molecular Evolution and Adaptive Radiation, T. J. Givnish and K. J. Sytsma, eds. (Cambridge: Cambridge University Press), pp. 129 - 163.

Springer, M. S., Westerman, M., Kavanagh, J., R., Burk, A., Woodburne, M. O., Kao, D. J., and Krajewski, C. (1998). The origin of the Australian marsupial fauna and the phylogenetic affinities of the enigmatic monito del monte and marsupial mole. Proceedings of the Royal Society of London Series B-Biological Sciences *265*, 2381- 2386.

Stace, C. A. (1989). Plant taxonomy and biosystematics, 2nd Edition (Cambridge: Cambridge University Press).

Stebbins, G. L. (1974). Flowering plants: evolution above the species level (London: Edward Arnold).

Stockey, R. A. (1982). The Araucariaceae: an evolutionary perspective. Review of Palaeobotany and Palynology *37*, 133-154.

Strand, J., Leebens-Mack, J., and Milligan, B. G. (1997). Nuclear DNA based markers for plant evolutionary biology. Molecular Ecology *6*, 113 - 118.

Suzuki, M., and Yoda, K. (1986a). Comparative wood anatomy of *Coriaria* of E. Asia (1). Journal of Japanese Botany *61*, 289 - 296.

Suzuki, M., and Yoda, K. (1986b). Comparative wood anatomy of *Coriaria* of E. Asia (2). Journal of Japanese Botany *61*, 333 - 342.

Swensen, S. M. (1996). The evolution of actinorhizal symbioses: evidence for multiple origins of the symbiotic association. American Journal of Botany *83*, 1503 - 1512.

Swensen, S. M., Mullin, B. C., and Chase, M. W. (1994). Phylogenetic affinities of Datiscaceae based on an analysis of nucleotide sequences from the plastid *rbc*L gene. Systematic Botany *19*, 157 - 168.

Swenson, U., and Bremer, K. (1997). Patterns of floral evolution of four Asteraceae genera (Senecioneae, Blennospermatinae) and the origin of white flowers in New Zealand. Systematic Biology *46*, 407 - 425.

Swofford, D. L. (1998). Phylogenetic analysis using parsimony version 4.0. .

Sykes, W. R., and Godley, E. (1968). Transoceanic dispersal in *Sophora* and other genera. Nature *218*, 495-496.

Szalay, F. S. (1994). Evolutionary history of the marsupials and an analysis of osteological characters (New York: NY: Cambridge University Press).

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

Takhtajan, A. (1996). Diversity and classification of flowering plants: Columbia University Press).

Takhtajan, A. (1969). Flowering plants. Origin and Dispersal (Washington: Smithsonian Institution Press).

Takhtajan, A. (1980). Outline of the classification of flowering plants (Magnoliophyta). Botanical Reviews *46*, 225 - 359.

Tanai, T. (1986). Phytogeographic and phylogenetic history of the genus *Nothofagus* Bl. (Fagaceae) in the southern Hemisphere. Journal of the Faculty of Science, Hokkaido University, series 4 *21*, 505-582.

Taubert, P. (1892). Pevision der Gattung *Griselinia*. Bot. Jahrb. Syst. *16*, 386 - 392.

Taylor, F., and Hill, R. S. (1996). A phylogenetic analysis of the Eucryphiaceae. Australian Systematic Botany, 533 - 558.

Thompson, P. N. (1996). The biogeography and systematics of *Coriaria* (Coriariaceae). PhD Thesis. (Leicester: University of Leicester), pp. 244.

Thompson, P. N., and Gornall, R. J. (1995). Breeding systems in *Coriaria* (Coriariaceae). Botanical Journal of the Linnean Society *117*, 293 - 304.

Thorne, R. F. (1968). Synopsis of a putatively phylogenetic classification of the flowering plants. Aliso *6*, 57 - 66.

Thorne, R. F. (1972). Major disjunctions in the geographic ranges of seed plants. Quaterly Review of Biology *47*, 365 - 411.

Thorne, R. F. (1973). A phylogenetic classification of the Angiospermae. Evolutionary Biology *9*, 35 - 106.

Thorne, R. F. (1983). Proposed new realignments in the angiosperms. Nordic Journal of Botany *3*, 85 - 117.

Thorne, R. F. (1992). An updated phylogenetic clasification of the flowering plants. Aliso, 365 - 389.

Tobe, H., Suzuki, M., and Fukuhara, T. (1992). Pericarp anatomy and evolution in *Coriaria* (Coriaricaeae). Botanical Magazine of Tokyo *105*, 289 - 302.

Trofimov, B. A., and Szalay, F. J. (1994). New Cretaceous marsupial from Mongolia and the early radiation of Metatheria. Proceedings of the National Academy of Sciences USA *91*, 12569 - 12573.

Truswell, E. M., Dershaw, A. P., and Sluiter, I. R. (1987). The Australian-south-east Asian connection: evidence from the paleobotanical record. In Biogeographical evolution of the Malay Archipelago, T. C. Whitmore, ed. (Oxford: Clarendon Press), pp. 32 - 49.

Tseng, C. C. (1980). The systematic position of *Aralidium* Miq. - A multidisciplinary study. 3. Pollen Morphology. Taxon *29*, 407 - 409.

Turra, A. (1765). *C. hermaphrodita.* Farset. Nov. Gen., Venetiis, 13.

van der Pijl, L. (1972). Principles of dispersal in higher plants, 2nd Edition (Berlin: Springer).

Van Steenis, C. G. G. J. (1962). The land-bridge theory. Blumea *11*, 235 - 542.

Van Steenis, C. G. G. J. (1971). *Nothofagus*, key genus of plant geography, in time and space, living and fosssil, ecology and phylogeny. Blumea *11*, 235 - 372.

van Tuinen, M., Sibley, C. G., and Hedges, S. B. (1998). Phylogeny and biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. Molecular Biology and Evolution *15*, 370 - 376.

Veblen, T. T., Schlegel, F. M., and Oltemari, J. V. (1983). Temperate broad-leaved evergreen forests of South America. In Temperate Broadleaved Evergreen Forests, O. J. D., ed. (Amsterdam: Elsevier), pp. 5 - 31.

Wagstaff, S. J., and Dawson, M. I. (2000). Classification, origin and patterns of divergence of Corynocarpus (Corynocarpaceae) inferred from DNA sequences. Systematic Botany *25*, 134 - 149.

Wagstaff, S. J., Heenan, P. B., and Sanderson, M. J. (1999). Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). American Journal of Botany *86*, 1346 - 1356.

Wallich, N. (1832). C. napalensis. In Plantae Asiaticae Rariores (London: Würtz & Richter), pp. 67.

Wanscher, J. H. (1933). Studies on the chromosome numbers of the Umbelliferae. III. Bot. Tidsskr. *42*, 49-50.

Warburg, O. (1892). C. papuana. Bot. Jahrb. *16*, 22.

Wardle, P., and Lee, W. G. (1990). Environmental and floristic gradients in westland, New-Zealand, and the discontinuous distribution of *Nothofagus* - comment. New Zealand Journal of Botany *4*, 479 - 481.

Waters, E. R. (1995). An evaluation of the usefulness of the small heat shock genes for phylogenetic analysis in plants. Annals of the Missouri Botanicl Garden *82*, 278-295.

Waters, J. M., and Burridge, C. P. (1999). Extreme intraspecific mitochondrial DNA sequence divergence in *Galaxias maculatus* (Osteichths : Galaxiidae). One of the world's most widespread freshwater fish. Molecular Phylogenetics and Evolution *11*, 1 - 12.

Wegener, A. (1915). Die Entstehung der Kontinente und Ozeane (Brunswick: Sammlung Vieweg).

Wendel, J. F., and Doyle, J. J. (1998). Phylogenetic Incongruence: window into genome history and molecular evolution. In Molecular systematics of plants II, DNA sequencing,

D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. (Boston: Kluwer Academic Publishers), pp. 265 - 297.

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In PCR protocols: a guide to methods and applications, M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White, eds. (New York: Academic press), pp. 315 - 322.

Wiedman, S., and Weberling, F. (1993). Studies on growth forms in the Paramo of Costa Rica. Flora *188*, 321 - 342.

Wilde, M. H., and Eames, A. J. (1952). The ovule of *Araucaria bidwilli* with discussion of the taxonomy of the genus II. Taxonomy. Annals of Botany New Series *16*, 27-47.

Wilson, M. A., Gaut, B., and Clegg, M. T. (1990). Chloroplast DNA evolves slowly in the palm family (Arecaceae). Molecular Biology and Evolution *7*, 303 - 314.

Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M., and Li, W.-H. (1989). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. Proceedings of the National Acadamy of Sciences USA *86*, 6201 - 6205.

Wolfe, K. H., Li, W.-H., and Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. Proceedings of the national acadamy of Sciences USA *84*, 9054 - 9058.

Wollenweber, E., Dörr, M., Rozefelds, D., Minchin, P., and Forster, P. I. (2000). Variation in flavonoid exudates in *Eucryphia* species from Australia and South America. Biochemical Systematics and Ecology *28*, 111-118.

Wray, G. A., Levinton, J. S., and Shapiro, L. H. (1996). Molecular evidence for deep Precambrian divergences among metazoan phyla. Science *274*, 568 - 573.

Xiang, Q.-Y., Soltis, D. E., and Soltis, P. S. (1998). Phylogenetic relationships of Cornaceae and close relatives inferred from *mat*K and *rbc*L sequences. American Journal of Botany *85*, 285 - 297.

Yang, Z. H. (1994). Maximum-likelihood phylogenetic estimation from DNA-sequences with variable rates over sites - approximate methods. Journal of molecular evolution *39*, 306 - 314.

Yoda, K., and Suzuki, M. (1992). Comparative wood anatomy of *Coriaria*. Botanical Magazine Tokyo *105*, 235 - 245.

Yokoyama, J., Suzuki, M., Iwatsuki, K., and Hasebes, M. (1998). Molecular phylogeny of *Coriaria*, with special emphasis on the disjuct distribution. Molecular Phylogenetics and Evolution *14*, 11 - 19.

Young, G. C. (1990). Devonian vertebrate distribution patterns and cladistic analysis of palaeogeographic hypotheses. In Palaezoic Palaeogeography and Biogeography, W. S. McKerrow and C. R. Scotese, eds.: Geological Society Memoir), pp. 243 - 255.

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

Zuckerkandl, E., and Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In In Evolving genes and proteins, V. Vryson and H. J. Vogel, eds. (New York: Academic Press), pp. 97 - 166.