Origins and Evolution of the New Zealand Forest Flora - a Molecular Phylogenetic Approach

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

Doctor of Philosophy
in
Plant Biology

at Massey University, Palmerston North,
New Zealand.

Karen Stöckler
2001

ABSTRACT

The origins and evolution of the New Zealand flora have puzzled the imagination of botanists world-wide. Competing hypotheses have sought to explain the floristic relationships between New Zealand and other Southern Hemisphere landmasses. Scientific approaches have involved geology, plant morphology, palynology and palaeobotany in investigations of the distribution patterns of these floras. Analyses presented in the current thesis use molecular data to investigate phylogenetic relationships of plant lineages native to the New Zealand forest flora.

In the present thesis, molecular work included amplification and sequencing of standard DNA markers such as nuclear ribosomal DNA, *ndh*F and *rbc*L gene sequence. These data were obtained for New Zealand and overseas species of Myrsinaceae, Nothofagaceae and genus *Agathis* (Araucariaceae). Analyses of these data have been presented alongside results and re-analyses of genetic data for Podocarpaceae, Proteaceae, Winteraceae and genus *Metrosideros* (Myrtaceae). These analyses aimed to synthesise recent work and provide a framework for further molecular investigations into the origins of the New Zealand woody forest flora.

Amplified fragment length polymorphism (AFLP) was used to locate polymorphic genome regions that were converted into sequence specific DNA markers. Information from AFLP and AFLP derived markers was used to elucidate evolutionary processes as well as interspecific and intraspecific relationships between closely related taxa of *Myrsine* and *Nothofagus*.

DNA analyses showed that the New Zealand forest hosts plants with very different origins and evolutionary histories. Results presented in the current thesis support hypotheses of vicariance and long-distance dispersal across Southern Hemisphere lands.

Molecular data are consistent with a continuous presence of *Agathis* (Araucariaceae), *Dacrydium* (Podocarpaceae) and *Pseudowintera* (Winteraceae) on the New Zealand archipelago since the break-up of the Gondwanan supercontinent. It is proposed that extant species of these lineages have evolved from ancestors that arrived on the New Zealand landmass during the Cretaceous. In contrast, divergence time estimates on *Nothofagus* suggest that relationships between extant *Fuscaspora* and *Lophozonia* beeches date back to the Mid Tertiary and are not explained by vicariance and continental drift.

Phylogenetic analyses substantiate fossil evidence of a Tertiary arrival of *Metrosideros* (Myrtaceae), *Myrsine* (Myrsinaceae), *Knightia* and *Toronia* (both Proteaceae). Similarly, dispersal from New Zealand to other southern lands has been inferred for *Metrosideros* and *Myrsine*. These findings and those reported earlier for alpine plant groups suggest that trans-oceanic dispersal is likely to be an important explanation of floristic relationships between New Zealand and other distant landmasses.

Molecular studies on New Zealand *Myrsine* suggest recent speciation events, geographic differentiation and ongoing hybridisation between some morphologically and ecologically distinct species. Intraspecific analyses on *Myrsine divaricata* and *Nothofagus menziesii* show that extant distribution patterns within New Zealand are relatively recent and may have developed during the Quaternary. Although both lineages have an ancient history on the New Zealand archipelago, it is concluded that species and their distributions – including that of the monotypic genera *Elingamita* - are of recent origin.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Pete Lockhart, who gave me the chance to work in this exciting field of research: thank you for your scientific experience, for countless discussions and for your support. Thank you for your motivation and your friendship!

I would like to thank my co-supervisor Prof. Dr. Bruce Christie for his efforts in providing ideas and plant material for this study.

I would like to thank my co-supervisor Prof. Dr. Uwe Maier for his encouragement, for providing a lab space in Marburg and for never shutting his door on me.

I would like to thank Prof. Dr. David Penny and Dr. Mark Large for discussions and for their friendly encouragement. It was very much appreciated!

Special thanks go to Trish McLenachan for her help in countless occasions: thank you for your advise in the lab, for listening and talking, for proof-reading my thesis, and..., and...., and.....

Thanks to all the people in the DEB Lab and the Boffin Lounge: to Richard, Carmel, Leon, Dan, Matt, Ant, Abbey, Nicole....

I really appreciated the support of the Institute of molecular BioSciences that gave me the opportunity to attend an international conference and provided possibilities for scientific exchange.

I gratefully acknowledge the Deutsche Akademische Austauschdienst, the New Zealand Marsden Fund and Massey University, without whose financial support this thesis work would not have been possible.

Last but not least, I would like to thank my husband Dr. Volker Sieber: thank you for standing by me after my decision to go to New Zealand, thank you for listening and supporting my efforts. Thank you for your time, patience, and encouragement. Thank you for giving helpful advise while I was writing up - without you I wouldn't have completed this work.

ABBREVIATIONS

A	Adenine
AFLP	Amplified Fragment Length Polymorphism
bp	base pairs
C	Cytosine
DNA	Deoxyribonucleic acid
G	Guanine
Indel	Insertion or deletion
ITS	Internal Transcribed Spacer
ISSR	Inter Simple Sequence Repeats
ML	maximum likelihood
MP	maximum parsimony
m.y.	million years
m.y.a.	million years ago
PAA	polyacrylamide
PCR	Polymerase Chain Reaction
QP	quartet puzzling
rDNA	ribosomal DNA
T	Thymine
RAPD	Random Amplified Polymorphic DNA markers
RFLP	Restriction Fragment Length Polymorphism
SSR	Simple Sequence Repeats

TABLE OF CONTENTS

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

The study of island biotas has been a key to elucidating ways and mechanisms of the evolution of species. Islands have also stimulated discussions on biogeographic relationships between distant landmasses and the dispersal of plants and animals (e.g. Darwin, 1859; Grant, 1998). The New Zealand archipelago offers one opportunity to investigate the evolution of a flora and fauna which has been isolated by more than 1 500 km from its nearest continental neighbour. Being an ancient continental landmass the biota combines aspects of both recent dispersal and ancient floristic and faunistic relationships (e.g. Fleming, 1979, Mildenhall, 1980). Geological, biological and palaeological studies have produced numerous hypotheses trying to explain biogeographic patterns in New Zealand with regard to other landmasses (Cockayne, 1919; Van Steenis, 1962; Raven and Axelrod, 1972). Many of these hypotheses try to incorporate the unstable geological history of the country and its possible influence on the biota (e.g. Pole, 1994). Advances in molecular systematics offer new techniques to analyse evolution and speciation of island biotas and allow a re-investigation of existing hypotheses (Wagstaff and Garnock Jones, 1998; Wagstaff *et al.*, 1999; Lockhart *et al.*, 2001).

1.1 PLATE TECTONICS AND THE GEOLOGICAL HISTORY OF THE SOUTHERN CONTINENTS

1.1.1 The Plate Tectonic Theory

The 'plate tectonic theory' is a comprehensive model of the geologic history of the earth, which was first proposed by the German geologist Alfred Wegener in 1912. It has subsequently been developed to explain major global architectural features of the earth such as mountain chains, ocean basins, and continents. The theory holds that the outer part of the solid earth, the lithosphere, consists of a relatively cool, dense shell, which is riding on a hotter, less dense zone, called the astenosphere. The lithosphere is divided into a number of segments (plates) which are in motion with respect to each other (Moores, 1990). This is the reason for continental drift. The plate movements lead to perpetual epirogenesis, where landscapes are formed and reformed through vertical and horizontal alterations of geological layers.

Most of the present dynamic activity of the earth is concentrated at the plate boundaries. Three types of boundaries are distinguished:

1. At divergent boundaries, plates spread apart. The resulting gap is filled with upwelling material from the earth's interior, forming ocean ridges. The process is called 'sea floor spreading'.

2. Regions, where one plate dives beneath another are called 'subduction zones'. Subduction zones are often subject to earthquakes and the outbreak of magma. Generally, the landmasses in these regions show frequent up and down movements during an unstable geological history.

3. 'Conservative faults' are boundaries where plates slide past each other; the lithosphere is neither created nor destroyed.

1.1.2 The Geological History of Southern Hemisphere Lands

The evidence suggests that, until the middle Cretaceous (see Appendix I for geological time scale), southern hemisphere continents formed a continuous landmass. According to this hypothesis, Australia, New Zealand, Antarctica, and South America were united with Africa, India, and Madagascar forming the super-continent 'Gondwana' (Sproll and Dietz, 1969; Tarling, 1971). An ancestral ocean called 'Tethys' separated Gondwana from today's northern hemisphere lands which also formed one continent called 'Laurasia'.

During Palaeozoic times the position of these two continents was different from what it is today. While Gondwana was located in the Northern Hemisphere, Laurasia was in the south. Later, global movements progressively rotated the continents and moved Gondwana into the Southern Hemisphere. The shapes of the continents changed considerably during the processes of continental drift (Frisch and Loeschke, 1993).

The present regional geography of southern hemisphere lands resulted from the break-up and spreading out of geosynclinal belts of Mesozoic to Palaeozoic age. In the Southern Hemisphere, sea floor spreading had commenced 120 million years ago between Africa and South America. By 90 million years ago South Africa, Malagasy, India and Australia were separated. About 80 million years ago, the Tasman Sea opened between Australia and New Zealand began to drift away from the Gondwanan landmass. New Zealand is completely separated by 65 million years before present.

1.2 BIOGEOGRAPHIC RELATIONSHIPS BETWEEN SOUTHERN HEMISPHERE LANDS

Biogeographic relationships between southern hemisphere lands have been noticed and discussed by a large number of biogeographers. Many widely different taxa including plants and animals are common to the temperate regions of South America, Australia, Tasmania and New Zealand, but do not occur in other parts of the world (e.g. Darwin, 1859; Hooker 1853; Darlington, 1965).

Many forest plant families like Podocarpaceae, Araucariaceae, Araliaceae, Myrsinaceae, Proteaceae, Winteraceae, Lauraceae, and Nothofagaceae share disjunct distribution patterns. These patterns are also found in many insects and other vertebrates among them several groups of carabid beetles and the Onychophora (Darlington, 1965; Raven and Axelrod, 1972).

Darlington (1965) claims that "... the pattern of geographic relationship within the cold temperate zone occurs only among plants and invertebrate animals, not among terrestrial or freshwater vertebrates (except some salt-tolerant fishes)".

Although the present distribution pattern of marsupials (they are found in Australia, New Guinea and on the American continent) may suggest another disjunct southern hemisphere distribution pattern, there is evidence from the fossil record that these animals evolved more than 100 million years ago in the Northern Hemisphere (Raven and Axelrod, 1972). The fossil record of these animals and their absence from New Zealand lead Darlington (1965) to the conclusion that marsupials reached South America and Australia separately from the north, after the break up of the continents. Based on fossil evidence he also suggested a post Gondwanan dispersal of frog species and turtles into Australia. Raven and Axelrod (1972) agreed that most of the modern birds, lizards, and snakes may have attained their present distribution in the late Tertiary and subsequently.

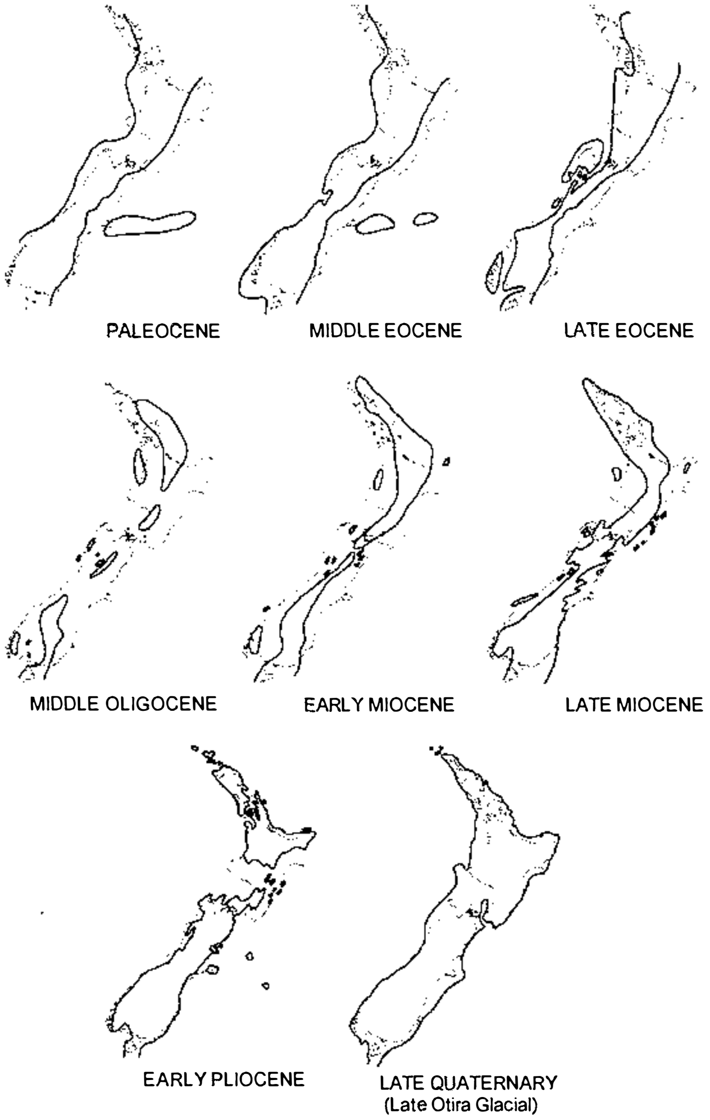
In contrast, the extant species of ratite birds - found in Africa, South America, Australia, New Guinea and New Zealand - may reflect a Gondwanan distribution. Raven and Axelrod (1972) suggest that these birds may have migrated from South America across Antarctica to populate other southern hemisphere lands before the Mid-Cretaceous. Molecular evidence corroborates this view by indicating that this group may have evolved 100 million years ago (Fridolfsson *et al.*, 1998). Recent molecular studies on ratite evolution published by Cooper *et al.* (2001) substantiate the hypothesis of a vicariant evolution of these ancient birds. They show that New Zealand ratites have diverged from other lineages during the Cretaceous - presumably before the break-up of Gondwana.

During the past decades it has been commonly accepted that continental drift explains many of the biogeographic patterns across the Southern Hemisphere. Raven and Axelrod (1972) state that many disjunct distribution patterns in the Southern Hemisphere reflect the geography of the region in the middle Cretaceous, when Africa, South America and Australia were connected with Antarctica. However, considering fossil and molecular evidence from a number of plant and animal species, the possibility of frequent trans-oceanic dispersal may not be disregarded.

1.3 THE HISTORY OF THE NEW ZEALAND LANDMASS

1.3.1 The Geological History of New Zealand

New Zealand is an archipelago situated in the south-west Pacific almost 2 000 km to the east of Australia, 12 000 km west of South America and more than 2 500 km north of the Antarctic continent. Its main territory lies between 33° and 47° of southern latitudes, with two major islands (North and South Islands). A few outlying islands extend the territory to 29° in the north-east (Kermadec Islands) and to beyond 52° in the south (Campbell Islands). The Chatham Islands to the east of the mainland are also part of New Zealand.



PALEOCENE MIDDLE EOCENE LATE EOCENE

MIDDLE OLIGOCENE EARLY MIOCENE LATE MIOCENE

EARLY PLIOCENE LATE QUATERNARY (Late Otira Glacial)

Figure 1.1: The changing outline of the New Zealand Archipelago during the Cenozoic (modified from Suggate *et al.*, 1978 by Winkworth, 2000).

New Zealand itself is situated on a system of relatively shallow submarine rises and plateaux between the south-west Pacific Ocean and the Tasman Sea. For over 60 million years the archipelago, now consisting of two large and many small islands, has been isolated from its nearest continental neighbours Australia and Antarctica. In the Upper Cretaceous, about 80 million years ago, the islands started moving away from Gondwanaland (Hayes & Ringis, 1975). During that time, erosion lowered the landscape and by the end of the Cretaceous much of the land was reduced to a peneplain of low relief. In the following Cenozoic stages only minor sea level changes were necessary to set parts of the New Zealand land mass under water (Fleming, 1979).

Much of the New Zealand landscape was submerged during the Oligocene. With the beginning of the Miocene, an increase of the topographical relief let the seawater retreat from formerly submerged areas (see Figure 1.1).

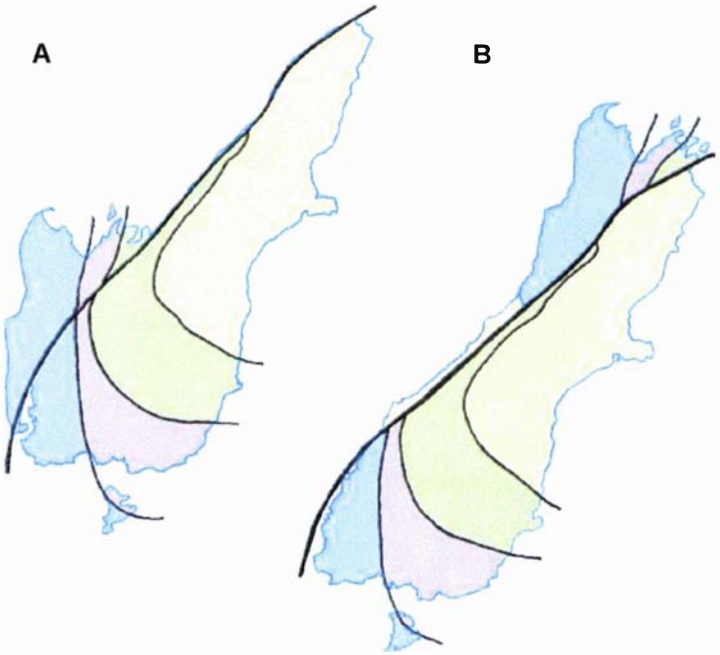


Figure 1.2: Model of the New Zealand South Island including the Alpine Fault sector. **A**: ca. 100 million years ago; **B**: at present. Colours represent different rock formations (after Stevens *et al.* 1995).

The New Zealand region straddles the boundary between the Pacific and Australian Plate. In the north of the archipelago, the eastern Pacific plate is proposed to be subducted beneath Australian plate to the west. In the south, the Australian plate is being subducted beneath the Pacific plate. The subduction zone forms the New Zealand Alpine Fault. In the Miocene the landmasses began to be shifted along this fault (see Figure 1.2). North-west Nelson and Westland, still attached to the south-west coast of the South Island and the present Fjordland regions started moving north (Heads, 1989). These movements were accompanied by increased volcanic activity during the Miocene and could be observed by a more extensive sedimentation of lava (Fleming, 1979).

Along the active transcurrent Alpine Fault, the Southern Alps were formed. A rapid land uplift began in the Pliocene, 5 million years ago - a process that is still going on (Batt *et al.*, 2000).

Table 1.1 summarises the geological history of New Zealand.

epoch	time	events
Lower Cretaceous	135 m.y.	NZ is attached to the super continent Gondwana
Upper Cretaceous	90 m.y.	opening of the Tasman sea: NZ starts to drift north
Palaeocene	65 m.y.	NZ is separated from Australia
Eocene	54 m.y.	reduction of landmass
Oligocene	37 m.y.	'Oligocene drowning'
Miocene	25 m.y.	slow land uplift
Pliocene	7 m.y.	mountain building: formation of the Southern Alps
Pleistocene	2 m.y.	ice ages: glacial-interglacial cycles

Table 1.1: Overview over the geological history of New Zealand. Abbreviations: **NZ:** New Zealand; **m.y.:** million years.

1.3.2 New Zealand Palaeoclimates

Studies on the fossil marine fauna suggest that the New Zealand climate has been changeable to a minor extent after its isolation of the Gondwanan landmass. In the Cretaceous slightly warmer temperatures dominated which suggest subtropical conditions on the islands. Measurements in oxygen-isotope contents of corresponding sediments show that temperatures declined steadily until the end of this era (Clayton & Stevens, 1968). Until the Pliocene, warmer temperatures (with an average 5-7 degrees above present values) were followed by slightly cooler periods with warm temperate climates (Devereux, 1967; Hornibrook, 1971).

Two and a half million years ago, with the beginning of the Pleistocene, global climate changes initiated the most recent period of glacial-interglacial cycles. In New Zealand there is evidence for at least 20 glacio-eustatic cycles during the last two million years (Beu & Edwards, 1993). The Otiran ice-age was the latest cold period, which ended approximately 10 000 years ago and had its maximum circa 18 000 BP (Fleming, 1979). Average temperatures during the coldest phase of the Otiran Glaciation were about 4.5° C below current temperatures and the permanent snow line was at least 1 000 m lower than at present (Willet, 1950). At present, an interglacial warm period dominates the global climates and New Zealand shows a temperate to warm temperate climate with a significant rain fall gradient from the west to east coast on both the North and the South Islands.

1.3.3 Effects of Pleistocene Glaciation on the Landscape

During glacial periods the polar ice sheets expanded and glaciers formed in alpine and mountainous regions. The glacial cycles were accompanied by sea level changes; during the colder periods the seawater level dropped and uncovered former seabeds (Webb and Bartlein, 1992).

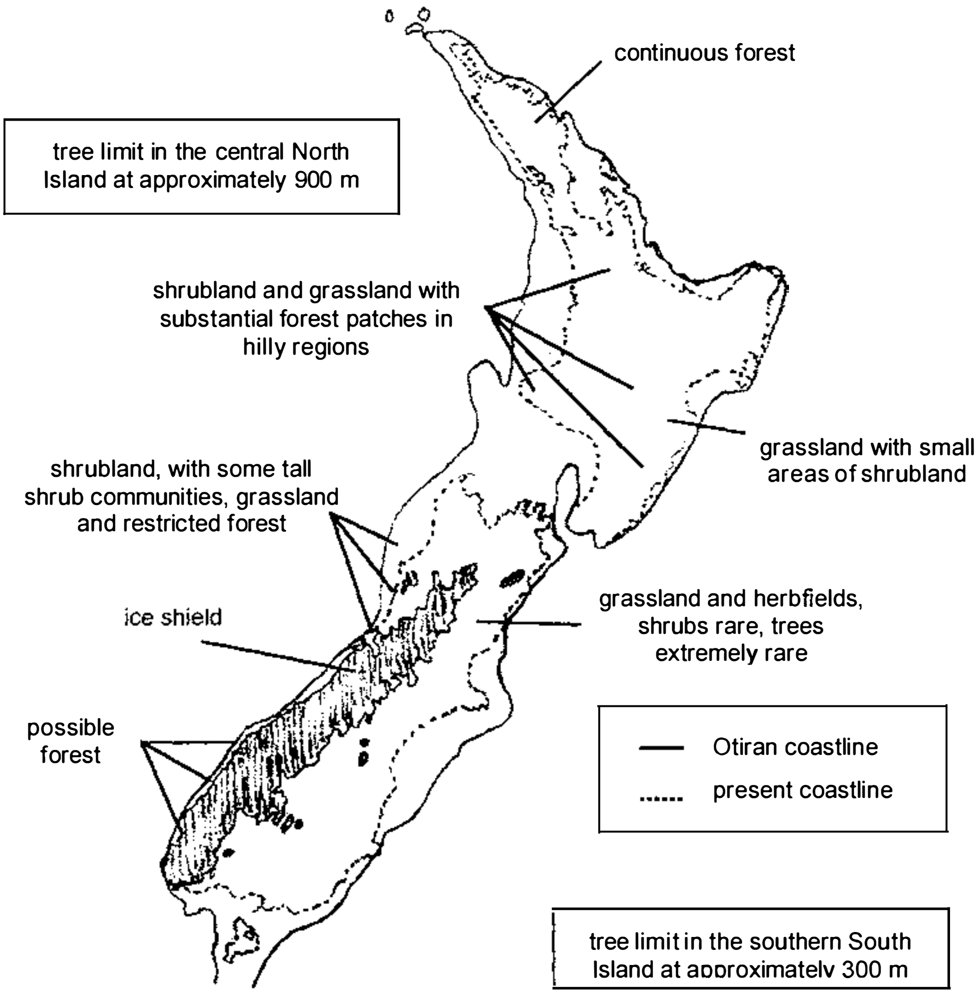


Figure 1.3: Distribution of glaciers and vegetation in New Zealand during the Otiran glaciation (after Suggate *et al.*, 1978 and McGlone, 2001).

In New Zealand, the eustatic cycles also contributed to sea level changes and the sea withdrew to more than 100 m below present levels. Much more of the continental shelf was exposed so that both Cook and Foveaux Strait were bridged. A land connection was established between north-west Nelson and Taranaki while a seawater inlet separated Marlborough from the southern Tararuas. The northern offshore islands were also attached to the mainland; only the Poor Knights Islands remained separated (Fleming, 1979). Figure 1.3. visualises the shape of the New Zealand landmass as well as the distribution of ice and vegetation during the Otiran glacial.

An ice shelf covered the Southern Alps and ice tongues carved glacial valleys on either side of the main divide (Gage, 1980). Signs of glaciation can be found on the South Island from Fjordland, where the glaciers dropped into the sea on the south-west coast, to north-west Nelson, where many small glaciers covered higher inland terrain. On the North Island, glaciation occurred to a minor extent in the Tararua ranges, on the Volcanic Plateau and on the top of Mount Taranaki.

Not only the protruding of glaciers, but also the melting of ice masses caused erosion on steep to moderate slopes throughout the South Island up to the Waikato Basin. Large alluvial plains and terraces were created by debris filled rivers and outwash fans can still be seen in most of the major water systems. Much of the eastern sides of both islands show loess depositions as a result of erosions in upland regions (McCraw, 1975).

1.4 THE FLORA OF NEW ZEALAND

The present New Zealand vegetation is temperate and predominantly evergreen and perennial. Deciduousness is found in varying degrees e.g. in species of *Plagianthus* and *Hoheria* (both Malvaceae), *Fuchsia* (Onagraceae) and *Sophora* (Fabaceae), annual herbs are uncommon.

The native vegetation includes forest communities, tussock grasslands and alpine areas. Before the first human settlements on the islands about 1 500-1 000 years ago (Sutton, 1994), most areas below the timberline, excluding zones of volcanic eruption and extensive swamps in the lower reaches of rivers, were covered with a variety of forms of forests or tall shrubland. Shrubs also extended to well above the timberline and a few shrubs extended into swamp areas (Poole and Adams, 1994). Recent studies suggest that during the Polynesian occupation, half of the forests and tall shrublands were burned. Much of the destroyed vegetation would have been replaced by secondary shrubland and tussock grassland. The European settlers who arrived in the 19th century continued the destruction of forest in a very rapid process.

Besides human modifications of the environment by burning of forest area for farming or hunting, the introduction of exotic plant and animal species has had a great impact on the New Zealand plant cover. Australian opossums, red deer, goats and other browsers have caused great damage to the vegetation. Adventive plants, introduced by human settlers, are now a major component of the New Zealand flora. Native plants, however, dominate in undisturbed areas.

Many botanists have noticed that New Zealand hosts a comparatively high number of small-leafed shrubs many of which have a divaricating branching morphology (Cockayne, 1919; Wardle, 1963). Cockayne suggested that extremely arid Pleistocene climates may have evoked these morphological features which seem characteristic of the New Zealand vegetation. Although Wardle (1963) disagreed with Cockayne that Pleistocene climates were responsible for this characteristic; he also considered the small leaves to be a xeromorphic adaptation. A different hypothesis suggests that these particular growth forms may be of ancient origin and were an advantageous adaptation against moa browsing (Stevens *et al.*, 1995).

1.4.1 Biogeographic Elements

Botanists and zoologists have agreed on the main elements in the biota, which have been classified according to their geographic relationships and possible origins. The classification presented is after Fleming, 1979.

1. Malay-Pacific: This element includes plants of tropical and subtropical derivation (excluding Australia). Plants from about 90 genera in 28 families show distinct subtropical or tropical affiliations. Examples are *Beilschmiedia* (Lauraceae), *Metrosideros* (Myrtaceae), *Elytranthe* (Loranthaceae), and *Macropiper* (Piperaceae). Some plants without an adequate fossil record are also attributed to this element. These are e.g. *Meryta* (Araliaceae) and *Cordyline* (Liliaceae), the cabbage tree.

2. Australian: The New Zealand flora is closely linked with that of Australia. About 80% of New Zealand angiosperm and gymnosperm genera and 300 species are held in common. However, the compositions of both floras are significantly different. Many plant groups that have a large number of species in Australia are represented by a comparatively small number of species in New Zealand. Examples are Fabaceae, Myrtaceae and Proteaceae.

a) Austral: This group includes taxa with a circumpolar or partially circumpolar distribution in the cool-temperate latitudes of New Zealand, eastern Australia, Tasmania and southern South America. There are also about 10 circumpolar lowland species on islands like Macquarrie, South Georgia and the Kerguelen which are separated by 11 000 km of ocean. Examples are *Phyllocladus* and other

Podocarpaceae, *Aristotelia* (Elaeocarpaceae), *Weinmannia* (Cunoniaceae), *Nothofagus* (Nothofagaceae) and *Fuchsia*.

3. Endemic: This group comprises plants that have no close relations in other countries to indicate their place of origin. These species have either developed in New Zealand (primary endemics) or their ancestors became extinct elsewhere (secondary endemics).

New Zealand holds a high number of endemic species: 40% of ferns and fern allies, all 20 species of gymnosperms, almost 90 % of Dicotyledons and about 60% of Monocotyledons are endemic to the archipelago (Godley, 1975). However, a much lower number of endemics are found on generic level. Only 9% of all genera present on the islands are endemic. This is an unexpectedly low ratio for a country, which has been isolated from the nearest continental landmass for over 60 million years. In Cape Peninsula of South Africa, for example, over 30% of all genera are endemic, in Hawaii over 20% and Australia has over 50% of endemic genera including several endemic families. Endemic genera include forest taxa like *Hoheria* and *Entelea*, as well as alpine genera like *Haastia* (Asteraceae) or *Hectorella* (Caryophyllaceae).

In the New Zealand alpine regions, the number of endemic species is particularly high (Smith, 1986). Approximately 93% of alpine species are endemic and this makes a significant contribution to the overall level of species endemism in New Zealand (Mark & Adams, 1995).

4. Holoarctic: A few plants are related to groups known from temperate regions of the Northern Hemisphere and not from the tropical regions, like the genus *Euphrasia* (Scrophulariaceae).

5. Cosmopolitan: Plants with a wide distribution range in the Northern and Southern Hemispheres. Many liverworts, mosses and ferns belong to this category; higher plants include many Asteraceae like *Senecio* or *Sonchus*.

1.4.2 Vegetation Types

The native flora of New Zealand has been classified by regional and climatic differences and by the predominant type of vegetation (Fleming, 1979, Wardle, 1991).

1. Mangrove forests: Very special ecosystems in the northern area of New Zealand are the mangrove forests that grow around the coast between mean sea level and high tide. New Zealand has only one species of mangrove, *Avicennia resinifera* (Avicenniaceae) - the most southerly mangrove in the world.

2. Conifer mixed hardwood forests: The New Zealand conifer mixed hardwood forests are composed partly of Podocarpaceae, Cupressaceae, and Araucariaceae and

partly of angiosperm trees and shrubs. Ferns are also a common component of these plant communities.

Common podocarps are *Dacrycarpus dacrydioides* (kahikatea) *Dacrydium cupressinum* (rimu), *Prumnopitys ferruginea* (miro), *Lagarostrobus colensoi* (silver pine) *Podocarpus totara* (totara) and *P. hallii* (Hall's totara). Depending on latitude, altitude and soil type, the composition of dominating species and understory species may vary.

In the northerly parts of New Zealand, mainly in Northland and the Auckland area, mixed conifer forest contain a component of *Agathis australis* (kauri, Araucariaceae) together with podocarps, a large amount of angiosperms e.g. *Metrosideros robusta* (northern rata) and *M. albiflora* (Myrtaceae), *Knightia excelsa* (rewarewa, Proteaceae), *Quintinia serrata* (Escalloniaceae), *Beilschmiedia tawa* (tawa, Lauraceae), *Rhopalostylis sapida* (nikau palm, Palmaceae), and various fern species, e.g. Aspleniaceae (*Asplenium flaccidum*), Cyatheaceae, and Dicksoniaceae.

In more southerly areas of New Zealand podocarps are the dominating element; warm-loving plants like *Agathis australis* or *Metrosideros albiflora* disappear while other species e.g. *Weinmannia racemosa* are abundant.

In the wetter areas and on the poorly drained soil types of the western South Island the conifer dominated communities may contain large amounts of *Metrosideros umbellata* together with *Weinmannia racemosa* and *Laurelia novae-zelandiae* (Monimiaceae). In these areas, a special conifer forest ecosystem is the kahikatea dominated swamp forest. These grow on the fertile soils of river floodplains.

Conifer forests in the rain shadow of the Southern Alps may contain many species that are common through much of the North Island but scarce or absent from the western South Island. Species like *Cyathea daelbata* (Cyatheaceae), *Hoheria angustifolia* (Malvaceae) or *Sophora microphylla* (Fabaceae) benefit from summer warmth and tolerate drought.

At higher altitudes of both the North and the South Island, lowland plants are subsequently replaced by species like *Phyllocladus alpinus* (on leached soils) or *Pseudowintera colorata*, *Schefflera* spp. and *Dicksonia squarrosa* (on more fertile soil types).

3. *Nothofagus* forests: *Nothofagus* species (southern beech; Nothofagaceae) form continuous subalpine forests, but with decreasing altitude become increasingly confined to shallow soils on the crests of ridges and spurs. Beeches also occur at low altitude on river bans or recent terraces that were seeded with beech before competing species could dominate. *Nothofagus* forests are found in the central North Island, in the Tararua ranges in the lower North Island, in the north-west Nelson region, in Fjordland and southern Otago (mainly Catlins National Park).

The four native species are *Nothofagus fusca, N. solandri, N. truncata* and *N. menziesii*. These can grow as monotypic stands or form mixed beech forests with two or three different beech species. In the cooler regions and at higher altitudes *Nothofagus* forests tend to be dominated by one species only. Here, southern beeches form the tree line. In warmer parts of New Zealand but at lower altitudes southern beeches may be part of a mixed forest community e.g. with *Dacrydium cupressinum*, *Dacrycarpus ferruginea* or *Weinmannia racemosa*. Shrubs and small trees common in many areas of beech are *Pseudowintera colorata*, *Coprosma foetidissima*, *C. pseudocunneata* (Rubiaceae), *Pseudopanax simplex*, *P. colensoi* (Araliaceae) or *Myrsine divaricata* (Myrsinaceae).

Similarly structured *Nothofagus* forests are also found in Australia, South America, New Caledonia and New Guinea.

4. Tussock grasslands: Two types of tussock grassland are distinguished on the New Zealand mainland: short-tussock grasslands and tall-tussock grasslands.

Short tussock grasslands cover areas of the eastern South Island from sea-level to between 800 and 900 m. This floristically diverse vegetation type is dominated by monocotyledon species like *Festuca novae-zelandiae*, *Poa caespitosa* and *P. colensoi* (Poaceae). The areas of tussock grassland in the South Island have been covered by forest communities until natural fires which occurred before humans arrived in the country destroyed large areas of forest (McGlone, 1980; Stevens *et al.*, 1995). From about 1 000 years ago, when humans arrived, fires destroyed more and more forests. Changes in microclimates, together with the small amount of rainfall in this region possibly prevented re-establishment of tree species. Tussock grassland replaced the former forest vegetation.

Tall tussock grasslands are found above the timberline to a greater or lesser extent along the axial range of the North Island and are widespread in the east of the South Island. The vegetation predominantly consists of several species of *Chionochloa* (Poaceae).

5. The alpine region: In New Zealand approximately 600 angiosperm species – representing about 30% of the vascular plant flora – occur in alpine areas (Mark & Adams, 1995).

Consistent with the greater extent of alpine habitats in the South Island, most of the alpine plant diversity is found in these mountains. In particular the mountains of the Nelson-Marlborough and Otago-Southland have high diversity, these areas supporting many locally endemic species, or plants with disjunct distributions (Wardle, 1963; Burrows, 1965; McGlone, 1985). By comparison the North Island mountains support an alpine flora of smaller size and lower diversity, which may be a result of the younger age and smaller extent of mountain habitats (McGlone, 1985; Wardle, 1988a).

6. Off-shore island vegetation: The Kermadec Islands, situated ca. 1 000 km north east of New Zealand have a little over 100 species of ferns and flowering plants 12% of which are endemic, 76 % found in New Zealand. There are no endemic genera and no gymnosperms. The dominating species of dry lowland forests is *Metrosideros kermadecensis* (Myrtaceae). In higher wetter forests *Ascarina lanceolata* (Chloranthaceae) and *Rhopalostylis cheesemanii* (Palmaceae) are common.

The Chatham islands, situated ca. 600 km to the east of New Zealand, hosts ca. 300 species of ferns and flowering plants of which 14% are endemic and 85% found in New Zealand. There are no large trees on the islands and much of the original forest vegetation has been destroyed by humans. Original coastal forests contained *Corynocarpus laevigata* (Corynocarpaceae) and *Rhopalostylis sapida*; at higher altitudes the endemic species *Dracophyllum arboreum* (Epacridaceae) was dominant.

In the subantarctic islands between 500 and 1 000 km south of the mainland there are about 193 species of which 31% are endemic and 63% found in New Zealand. The one endemic genus is *Pleurophyllum* (Asteraceae). Grasslands with *Poa litorosa* and *P. foliosa* or *Chinochloa antarctica* are common, also shrublands dominated by *Dracophyllum longifolium*. Auckland Islands also have a narrow belt of coastal forests with *Metrosideros umbellata* and shrublands with *Myrsine divaricata*, *Coprosma* spp. and *Pseudopanax simplex*. On Macquarrie Island there is one woody plant species, only 31 species of flowering plants, 3 ferns and 1 lycopod.

An interesting feature of the flora of the subantarctic Islands are the so-called 'megaherbs'. These have particularly large foliage and bright flowers while their close relatives on the New Zealand mainland generally have small leaves and inconspicuous flowers. Examples of megaherbs are *Stilbocarpa polaris* (Araliaceae) or *Pleurophyllum criniferum* (Asteraceae).

1.4.3 History of the New Zealand Flora: Evidence from the Fossil Record

Palynological and palaeobotanical studies indicate that early New Zealand vegetation consisted of gymnosperm forests including several species of Podocarpaceae and Araucariaceae (Fleming, 1979; Mildenhall, 1980). Fossil gymnosperm plant communities are known from Jurassic deposits excavated in Curio Bay (Southland) and Port Waikato (Pole, 1998). The dominant forest trees were members of the extinct genus *Podozamites* (c.f. Araucariaceae) and *Elatocladus* (c.f. Podocarpaceae).

Subsequently, the gymnosperms were joined by more and more angiosperm taxa and with the beginning of the Cenozoic, Nothofagaceae, Myrtaceae and Proteaceae (*Banksiaeidites*) had become important members of the New Zealand flora. Species of *Nothofagus* played a leading role.

During the Oligocene and Miocene eras *Nothofagus* forests of *brassii*-type species remained dominant but podocarps and *Nothofagus fusca*-type forests were also well represented. During this time, New Zealand was drifting northward and came under subtropical influence.

As judged by the fossil pollen record, a large number of new plant species appeared in New Zealand during the Oligocene and Early Miocene (Fleming, 1979; Mildenhall, 1980; Macphail, 1997b). Presumable Oligocene arrivals were e.g. *Lycopodium*, *Myrsine*, *Coprosma*, *Astelia* (Liliaceae) and *Pittosporum*. With the beginning of the Miocene, additional plant genera show their first recognisable appearance in the New Zealand fossil record. These are among others: *Elaeocarpus* (Elaeocarpaceae), *Macropiper* (Piperaceae), *Muehlenbeckia* (Polygonaceae), *Fuchsia*, *Cordyline* and *Phormium* (Liliaceae).

During the Oligocene and Miocene, warm loving plant genera known from tropical environments were recorded in high numbers. These include *Bombax* (Bombacaceae) and *Cocos* (Palmaceae), which were identified from their fossil nuts (Fleming, 1979). However these genera disappeared from New Zealand records during the mid-Miocene, possibly as a response to cooler climates.

first evidence of fossil pollen	family	genus	habitat
Lower Cretaceous	Araucariaceae	*Agathis* (1)	forest
	Podocarpaceae	*Podocarpus* (6) *Dacrydium* (1) *Phyllocladus* (2)	
Upper Cretaceous	Winteraceae	*Pseudowintera*	
	Nothofagaceae	*Nothofagus* (4)	
	Caryophyllaceae	*Scleranthus* (3)	
Palaeocene	Myrtaceae	*Metrosideros* (11)	
Eocene	Araliaceae	*Schefflera* (1)	
	Proteaceae	*Knightia* (1)	
Oligocene	Palmae	*Rhopalostylis* (2)	
	Myrsinaceae	*Myrsine* (9)	
Miocene	Agavaceae	*Cordyline* (5)	
	Violaceae	*Melicytus* (5)	
	Apiaceae	*Gingidia* (3)	montane grasslands
		Hydrocotyle (8)	damp areas
Pliocene	Brassicaceae	*Cardamine* (6)	
	Gentianaceae	*Gentiana* (24)	lowland and montane
	Scrophulariaceae	*Hebe* (79)	
Pleistocene	Boraginaceae	*Myosotis* (35)	open areas
	Ranunculaceae	*Ranunculus* (43)	lowland and montane
	Apiaceae	*Aciphylla* (39) *Anisotome*(16)	mainly montane

Table 1.2: Examples from the pollen record of New Zealand plant families (after Mildenhall, 1980). The values in brackets indicate the number of extant species.

By the early Pliocene, New Zealand had arrived at latitudes near to its present day position. Pliocene plant fossils give evidence of cooling climates, although *Agathis* and *Knightia* still appear south of their present limits. *Brassii*-type beech forests extended south to Otago and podocarps remained abundant. Only a few newcomers are recorded in the pollen record e.g. *Quintinia* (Grossulariaceae), *Hebe* (Scrophulariaceae) and *Taraxacum* (Asteraceae) (Fleming, 1979).

Table 1.2. shows when some important plant groups first appeared in the New Zealand geological strata.

1.4.4 The Importance of Trans-Oceanic Dispersal

Palynological (Pole, 1994; Macphail, 1997b), morphological and biochemical studies (Martin and Dowd, 1993; Swenson and Bremer, 1997), chromosome analyses (Godley, 1967; Sykes and Godley, 1968) and an increasing number of DNA sequencing studies (Winkworth *et al*. 2001 and references therein) provide convincing evidence that long distance dispersal has been important for the establishment of many Australasian and Pacific plant distributions. Examples for these are listed in Table 1.3.

Palynological studies substantiate dispersal events from Australia to New Zealand after the break-up of the continent. The fossil record discussed by Macphail (1997b) and Mildenhall (1980) shows the appearance of many taxa first in Australia and then in New Zealand. Dispersal from Australia to New Zealand is consistent with the hypothesis of west wind mediated trans-oceanic dispersal (Carlquist, 1972). Recent DNA studies also indicate that westerly winds have been an important mechanism for dispersing the ancestors of many extant plant taxa (e.g. Wagstaff and Garnock-Jones, 1998; Hurr *et al*., 1999).

The circum-polar westerly winds, which developed following the separation of Australia from Antarctica in the Palaeocene (ca. 55 m.y.a.), are generally considered the primary mechanism for the dispersal of plant propagules over great distances in the South Pacific (Mildenhall, 1980). Particle sediment analysis indicates that these prevailing westerly winds blew with great intensity during the Pleistocene (Stewart and Neall, 1984). Meteorological observations suggest that even today the westerly dispersal of seeds, spores and pollen from Australia to New Zealand is a relatively common phenomenon (Mildenhall, 1980).

Raven and Axelrod (1972) have argued emphatically that dispersal could not occur against the direction of west wind drift namely from New Zealand to Australia. However, recent molecular systematic studies on *Ranunculus* and *Myosotis* show that during the Pliocene/Pleistocene dispersal was not only in the direction of west wind drift but occurred multi-directionally within the South Pacific and adjacent regions (Lockhart *et al*. 2001; Winkworth *et al*., 2001).

plant family	plant group	direction of dispersal	reference
Nothofagaceae	*Nothofagus*	AU → NZ	Martin & Dowd, 1993
Asteraceae	Gnaphalieae	AU → NZ, AU → NGuin	Breitwieser *et al* 1999
Fabaceae	*Sophora*	NZ → Pacific/ SAtl/ Sub	Hurr *et al.* 1999
Viscaceae	*Korthalsella*	AU → NZ	Molvray *et al.* 1999
Fabaceae	*Montigena*	AU → NZ	Wagstaff *et al.*, 1999
Fabaceae	*Carmichaelia*	NZ → AU/ Lord Howe	Wagstaff *et al.*, 1999
Stylidiaceae	*Oreostylidium*	AU → NZ	Laurent & Bremer, 1999
Asteraceae	*Microseris*	AU ↔ NZ	Vijverberg *et al.* 1999
Corynocarpaceae	*Corynocarpus*	NGuin → NZ	Wagstaff & Dawson, 2000
Myrtaceae	*Metrosideros*	NZ → Pacific	Wright *et al.* 2000
Brassicaceae	*Cardamine*	Aus → NZ	Mitchell & Heenan, 2000
Ranunculaceae	*Ranunculus*	NZ → AU, NZ → Sub	Lockhart *et al.*, 2001
Tetrachondraceae	*Tetrachondra*	SA → NZ	Wagstaff *et al.*, 1999
Scrophulariaceae	*Hebe*	NZ → Sub, NZ → SA	Wagstaff & Garnock-Jones, 1998
Scrophulariaceae	*Parahebe*	NZ → NGuin	Wagstaff & Garnock-Jones, 1998
Scrophulariaceae	*Chionohebe*	NZ → Au	Wagstaff & Garnock-Jones, 1998
Apiaceae	*Aciphylla*	AU ↔ NZ, NZ → Chat	Winkworth *et al.* 2001
Apiaceae	*Anisotome*	AU ↔ NZ, NZ ↔ Sub	Winkworth *et al.* 2001
Gentianaceae	*Gentianella*	SA → NZ, NZ → AU	Winkworth *et al.* 2001
Boraginaceae	*Myosotis*	NZ → AU, NZ → Sub NZ → NGuin, NZ → SA	Winkworth *et. al.* 2001

Table 1.3: Suspected dispersal events for New Zealand plant groups (after Winkworth *et al.*, 2001). **Abbreviations**: **AU**: Australia; **Chat**: Chatham Islands; **Lord Howe**: Lord Howe Island, **NGuin**: New Guinea; **NZ:** New Zealand; **SA**: South America; **SAtl**: South Atlantic; **Sub**: Subantarctic Islands.

1.4.5 Effects of Past Environmental Changes on the New Zealand Flora

Geological research and studies on New Zealand's palaeoclimates have shown that the archipelago has had a very changeable past. Geological and climatic changes are very likely to have altered ecological conditions of biotopes and influenced the evolution of the biota (Fleming, 1979).

The 'Oligocene drowning': Geological changes such as continental movements, land uplifts or transgressions have altered the landscapes slowly. Geological research shows that the most dramatic reduction of New Zealand landscape occurred during the Oligocene (see Figure 1.1). Biologists agree that this event may have severely reduced species diversity (Fleming, 1979). Recent DNA analyses on New Zealand ratites support the existence of a widespread genetic bottleneck during the mid-Cenozoic

(Cooper and Cooper, 1995). Re-examination of fossil and palynological data led Pole (1994) to the conclusion that the Oligocene drowning event may have completely eliminated the ancestral vegetation on the archipelago. Indeed, palynological data (Mildenhall, 1980; Macphail, 1997b) and recent molecular data (Wagstaff and Garnock-Jones, 1998; Winkworth, 2000) suggest that many taxa have arrived in New Zealand following the Oligocene. However, no convincing evidence has been brought forward that supports the complete extinction of the New Zealand biota during the Oligocene.

Mountain building: Another incisive incident in New Zealand's geological history was the event of mountain building, which began during the Pliocene and has continued in the Holocene. With the uplift of the Southern Alps, subalpine and alpine environments offered new biotopes and ecological niches for plants and animals. With 5 million years of age these alpine habitats are a very recent development (Batt *et al.* 2000), and thus, the origin of New Zealand's high altitude species has been of particular interest to biologists (Wardle, 1978; Winkworth, 2000).

Pleistocene ice ages: Ice age cycles changed living conditions relatively rapidly for plants and animals in all biotopes (Webb and Bartlein, 1992). During cold periods, the polar ice sheets expanded and glaciers formed in the alpine regions of New Zealand (see 1.3.3). As vegetation zones shifted, species with adaptations to warmer climates were eliminated from their former habitats by the cooler climate conditions. The range of some species contracted whilst others may have migrated to populate new environments or may have survived in refugia with favourable microclimates (McGlone, 1985; Wardle, 1988a).

Evaluating the effects of glacial/interglacial activity on the flora and fauna is challenging. Since the beginning of this century numerous investigations concerning this issue have been undertaken (e.g. Cockayne, 1919; Wardle, 1988a; Comes and Kadereit, 1998).

Volcanism and earthquakes: Local disturbances have been frequently caused by volcanic eruption and earthquakes throughout the history of New Zealand. These disturbances have caused local extinction of plants and animals and have abruptly changed the living conditions in these particular areas (Stevens *et al.*, 1995).

1.4.6 Distribution Patterns within the New Zealand Flora

At the beginning of the century, Cockayne (1919) proposed a comprehensive scheme for dividing New Zealand into floristic and ecological districts. The scheme takes climate and soil conditions of a region into account in explaining the distributions of characteristic plant associations.

Three main 'centres of endemism' and floristic richness are described within New Zealand (Wardle, 1963; Burrows, 1965). Southland, Otago and the Subantarctic

Islands represent the Southern Centre, the Nelson-Marlborough Centre covers the northern regions of the South Island and the Northern Centre comprises Northland, the Waikato region and East Cape. The three 'centres' are separated by two 'gaps', the Southern Gap - Westland and Canterbury - and the Northern Gap, found in the upper Wellington region in the south of the North Island.

Only a low number of species span the gap regions, although there have been no significant differences observed in either the ecology or in the geology of these areas. Gap regions also show similar climates and microclimates to centre regions (Wardle, 1988a; Leathwick, 1998). In both gaps and centres, rainfall gradients exist in east west directions and offer a range of biotopes. This apparent lack of ecological constraints on the observed distribution patterns suggests that historical events are likely to be responsible for the centre and gap structure within the New Zealand flora (Wardle, 1963,1985, 1988a; Burrows, 1965; McGlone, 1985).

1.4.6.1 Glacial Refugia: A Hypothesis on the Survival of Species

Regarding the distribution patterns and the structure of the New Zealand flora, Cockayne (1919) suggested that Pleistocene glaciation has had a significant influence on the vegetation of the islands. In 1950, Willet developed this idea to explain the absence of beech from central Westland. Over a decade later, Wardle formalised a hypothesis to explain the discontinuity in the distribution of a number of New Zealand species. He suggested that the centres of endemism in the South and North Island represent centres where species survived the cold and harsh climates of the Pleistocene glacial periods. He proposed that since the last glacial, many taxa have not fully reoccupied their possible habitats (Wardle, 1967, 1988a).

From geological investigations there is evidence for only minor glaciation in areas throughout Nelson and Marlborough as well as in Otago-Southland regions (Fleming, 1979). With regard to its climate and relief, it has been suggested that Nelson harboured plants that were adapted to moist and mild conditions whereas Marlborough was a refugium for lowland and montane plants with more rupestral habitats (Wardle, 1963). Proposed characteristics would have made a Northland refugium more suited for species with a demand for warmer climates (Wardle, 1963). In the North and South Islands, alpine, subalpine and montane plants may have migrated to lower altitudes as glaciers protruded and migrated to former habitats during interglacial warm periods (McGlone, 1985; Wardle, 1988a).

In contrast to these predications, isozyme studies on New Zealand hard beech (*Nothofagus truncata*) have supported the possibility of survival of this less cold tolerant forest species in the South Island during Pleistocene glacial/interglacial cycles (Haase, 1992). Nevertheless, compared to North Island populations, significantly lower genetic differentiation within the South Island (south Westland) has been found and this suggests that reduction of population size in the South Island during ice ages may have been accompanied by loss of genetic variation (bottleneck effect; Haase, 1992).

Postglacial Reforestation

Pollen analyses suggest that the reforestation of the North Island with podocarp-hardwood communities occurred within 300 years during the early post-glacial between 14 000 and 13 000 BP (McGlone, 1980). Beech forests were present as scattered patches and seem to have expanded at a slower rate. Reforestation of the South Island did not take place until the mid post-glacial (10 000 to 5 000 years ago) and was a result of climate amelioration during this time. Pollen diagrams of lowland regions representing this era show a remarkable increase in pollen of podocarps and give rise to the idea of a very fast expansion of woody species occurring about 10 000 years ago (Moar, 1970). The data indicate that the reoccupation of suitable habitats by podocarp-dominated lowland forests might have taken place within less than 500 years.

It has been suggested that rapid recolonisation can be explained if plant species survived in numerous pockets of favourable microclimate within the South Island (McGlone, 1985). Survival in many small refugia scattered over the species' present ranges would have made it possible for forest plants to reoccupy former habitats within a short period of time (McGlone, 1985). If this were true, the number and location of microrefugia would be expected to vary for different species.

1.4.6.2 Tectonic Movements: A Hypothesis on the Disjunction of Populations

There seems no doubt that, as with the analogous situation in the Northern Hemisphere (Comes and Kadereit, 1998), Pleistocene climate change has greatly influenced the evolution and distribution of the New Zealand flora. However, other events during the history of the islands cannot be completely disregarded. In 1985 McGlone introduced a complex proposal on the possible influence of Miocene geological changes on the flora and fauna of New Zealand. He suggested that Miocene marine transgression and land rafting might have induced disruption of the Miocene flora. He pointed out that the increase of topographical relief from the Late Miocene onwards altered microclimates, soil conditions and changed the composition of habitats and is likely to have had effects on the New Zealand biota. Until then, lowland habitats with leached and poorly drained soils dominated the landscape of the archipelago. The tectonic movements created steep hills and low mountain ranges. The evolution of new soils was then a result of the geological reorganisation of the landscapes. McGlone considered that forestation of the newly formed Miocene habitats occurred slowly in succession following grass and shrub associations. The Miocene flora had been adapted to more infertile lowland conditions. Remnants of the typical plant associations of this period are presumed to be now concentrated in the north of the North Island where the environment resembles that under which they developed.

The increase in topographical relief created air turbulence and higher winds. The consequence of Luv and Lee conditions across the mountain barriers increased the east-west rainfall gradient in New Zealand. Cooler and windier climates may have eliminated cold intolerant plants from the South Island (McGlone, 1985). This may be true for *Agathis australis* Salisbury which is now confined to areas north of about latitude 38°S (Ecroyd, 1982). During the Miocene this species or a close relative was also present on the South Island (Fleming, 1979). The extinction of Kauri populations on the South Island could be a consequence of the reported changes in microclimate, soil and relief during the late Miocene and the Pliocene (McGlone, 1985). Species that are presently associated with *Agathis australis* forest on the North Island can still be found on the South Island and show disjunctions within their habitats. McGlone proposes that *Nothofagus truncata* (Col.) Ckn. is one such example for the presence of *Agathis australis* associates in other than the shared northern habitat. Hard beech prefers old leached soils and warm microclimates that are likely to have dominated during the Miocene. McGlone suggests that the disjunction of *N. truncata* populations has occurred because of a loss of South Island habitats as a consequence of the tectonic changes during the late Miocene and Pliocene.

The environmental changes caused by the Cenozoic land uplift in New Zealand can be summarised as follows:

1. The evolving mountains created barriers for the lowland species that dominated the New Zealand Miocene flora. Subtropical plants were restricted in sheltered lowland regions and did not survive in higher altitudes.

2. Soil and ground water conditions changed in most areas throughout the landmass. Plants with specific preferences for poorly drained lowland soils were out competed in areas with newly evolved soil types.

3. Microclimates altered as a result of the increased topographical relief. With the mountains stronger local winds came up and the west-east rain fall gradient increased.

McGlone proposed that some disjunctions of New Zealand plant populations can be explained by these late Miocene environmental changes.

1.5 OUTLINE OF THE THESIS

In this thesis DNA analyses have been used to investigate the history of extant plant taxa native to New Zealand forest. Genetic data was determined for three plant genera: *Agathis, Myrsine* and *Nothofagus*. Analyses of these data have been presented alongside results and re-analyses of genetic data determined for other New Zealand forest plant groups. The materials and methods used in carrying out this comparative approach have been described in Chapter 2.

Chapter 3 describes the results from DNA sequence data determined during this PhD project, alongside results from re-analyses of molecular data available from the NCBI GenBank. Phylogenetic analyses are presented for Araucariaceae, Podocarpaceae, Myrtaceae (genus *Metrosideros*), Myrsinaceae, Nothofagaceae, Proteaceae and Winteraceae, all of which have extant species in New Zealand. In this study, emphasis is placed on examining the question of lineage and species origins in the context of molecular, geological and palynological evidence. Hypotheses are proposed to explain the origin and evolution of these taxa in New Zealand. The purpose of this section of the thesis is to synthesise recent work and provide a framework for further molecular investigations into the origins of the New Zealand forest flora.

In Chapter 4, results are reported on the use of AFLP (a novel DNA fingerprinting technique) as a method for obtaining molecular markers in the study of natural plant populations. AFLP was used to detect and isolate variable genome regions in *Myrsine* and *Nothofagus*. These markers provided useful tools for studying closely related taxa and populations.

Results on both novel AFLP and more widely used molecular markers (nuclear ITS, plastid *ndh*F and *trn*L regions) are described in Chapters 5 and 6. Chapter 5 presents investigations of a group of closely related taxa of Myrsinaceae in New Zealand. Chapter 6 concentrates on population genetics to investigate disjunct distributions of the New Zealand silver beech (*Nothofagus menziesii*) and the Australian antarctic beech (*Nothofagus moorei*.)

Chapter 7 summarises the main findings of this thesis, and outlines directions for future research.

CHAPTER 2

Material and Methods

2.1 MATERIAL

2.1.1 Laboratory Equipment

	article	source
	Microcentrifuge	Eppendorf, Germany
	Transilluminator (wavelength: 302 nm)	UVP Incorporated
	Video Camera	Panasonic
	Agarose Gel Chamber Horizon ® 11.14	Life Technologies, U.S.A.
	S2 Casting Boot for Polyacrylamide Gel Electrophoresis	Life Technologies, U.S.A.
	S2 Shark Tooth Combs (5.7 mM point-to-point spacing)	Life Technologies, U.S.A.
	S2 0.4 mM Plastic Spacers	Life Technologies, U.S.A.
	Microwave Oven	Panasonic
	Mechanical Shaker	Labline Instruments, U.S.A.
	PTC-200 DNA Engine	MJ Research, U.S.A.
	PTC-150 Minicycler	MJ Research, U.S.A.
	Milli-Q System	Milli-Q

2.1.2 Plastic and Glass Ware

	article	source
	beaker (250 ml)	Kimax, U.S.A.
	glass tubes	Kimax, U.S.A.
	glass pasteur pipette	John Poulton Ltd., England
	screw lid bottles (50 - 100 - 250 - 500 - 1000 ml)	Schott, Germany
	plastic ware	Payless Plastics, NZ

2.1.3 Consumables

	article	source
	microcentrifuge tube (1.6 ml)	Axygen Scientific, U.S.A.
	reaction tubes (0.2 ml)	Sørenson BioScience, U.S.A.
	pipette tips (10µl-20µl -200µl)	Life Technologies, Australasia
	pipette tips (1000µl)	Greiner Labortechnik, Germany
	pipette tips with filter (10µl-20µl -200µl -1000µl)	Molecular Bioproducts, U.S.A.
	paper towels (Kimwipe)	Kimberley-Clark, U.S.A.
	Chromatography paper 3 mM	Whatman, U.S.A.

2.1.4 Chemicals and Reagents

	article	source
	ammonium persulphate (APS)	Sigma, U.S.A.
	Ampicillin	Sigma, U.S.A.
	Bacto-Agar	Life Technologies, U.S.A.
	Bacto-Trypton (peptide No. 5)	Difco, U.S.A.
	Bind-Silane	Pharmarcia, Germany
	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal)	Sigma, U.S.A.
	bromophenol blue	Serva, Germany

2.1.4 Chemicals and Reagents (continued)

article	source
cetyl trimethyl ammonium bromide (CTAB)	Sigma, U.S.A.
chloroform	BDH, England
dNTP	Boehringer Mannheim, Germany
ethanol p.A. 99.8-100%	BDH, England
Ethylenediaminetetra-acetic acid [EDTA	BDH, England
Ficoll Ty 400	Pharmarcia
formaldehyde (37%)	Sigma, U.S.A.
glacial acetic acid and 0.05% (v/v)	BDH, England
isopropanol	BDH, England
isopropylthiogalactoside (IPTG)	Sigma, U.S.A.
LE agarose	Boehringer Mannheim, Germany
magnesium acetate (MgOAc,),	Sigma, U.S.A.
orthoboric acid	BDH, England
potassium acetate (KOAc,),	Sigma, U.S.A.
polyvinyl-pyrrolidone (PVP)	Sigma, U.S.A.
Polyethylenglycol (PEG 8000)	BDH, England
RAIN-X®	UNELKO, U.S.A.
silica gel (6-8 mesh)	BDH, England
silver nitrate ($AgNO_3$,)	BDH, England
sodium acetate	BDH, England
sodium carbonate (Na_2CO_3 anhydrous, AnalaR)	BDH, England
sodium chloride (NaCl)	BDH, England
sodium hydroxide (NaOH)	Sigma, U.S.A.
sodium thiosulphate ($Na_2S_2O_3$,).	BDH, England
NNN'N'-tetramethylethylenediamine (TEMED)	BDH, England
Tris-acetate	Life Technologies, U.S.A.
Triton x - 100	BDH, England
urea	BDH, England
xylene cyanol	Sigma, U.S.A.
yeast extract	Difco, U.S.A.

2.1.5 Kits and Ready-to-Use Products

article	source
Dye Terminator Sequencing Mix	Perkin-Elmer Corporation, U.S.A.
Long Ranger gel solution	FMC BioProducts, U.S.A.
MAX efficiency DH5∝™ competent cells	Life Technologies, U.S.A.
PCR-Kit	Qiagen, Germany
pGEM-T Easy	Promega, U.S.A.

2.1.6 Enzymes

enzyme	source
Eco RI (10 U/μl)	New England Biolabs, U.S.A.
T4 Ligase (2000 U/μl)	New England Biolabs, U.S.A.
Lysozyme	Sigma, U.S.A.
Mse I (10 U/μl)	New England Biolabs, U.S.A.
RNase A	Sigma, U.S.A.

2.1.7 DNA markers

	article	source
	100 bp DNA Ladder	Life Technologies, U.S.A.
	1kb DNA Ladder	Life Technologies, U.S.A.
	1kb plus DNA ladder	Life Technologies, U.S.A.
	Low DNA Mass™ Ladder	Life Technologies, U.S.A.

2.1.8 Software

	software program	source
	ClustalX 1.8	ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/
	Chromas 1.42	http://www.technelysium.com.au/chromas.htm
	ImagePC software	Scion Corporation, U.S.A.
	Genetyx-Mac 7.3	Software Development Co. LFD., U.S.A.
	Paup*	Sinauer Associates, Inc. Publishers, U.S.A
	Popgene 3.2	http://www.ualberta.ca/~fyeh/fyeh/
	Spectrum 2.3.0	http://www.taxonomy.zoology.gla.ac.uk/~mac/mike.htm
	SplitsTree 3.1	ftp://ftp.uni-bielefeld.de/pub/math/splits/

2.1.9 Standard Molecular Markers Used for Phylogenetic Analyses

	gene region	genome	reference
	Internal transcribed spacer regions (ITS 1 and ITS 2) and 5.8S of nuclear rDNA	nuclear genome	Baldwin, 1992
	*ndh*F gene region (partial sequence)	plastid genome	Neyland and Urbetsch, 1996
	*rbc*L gene region (partial sequence)	plastid genome	Hasebe *et al.*, 1994
	*trn*L - *trn*F intergenic spacer	plastid genome	Taberlet *et al.*, 1991
	*trn*L intron	plastid genome	Taberlet *et al.*, 1991

2.1.10 Oligonucleotides

Oligonucleotides, which were used throughout this study as PCR primers and AFLP linkers, were obtained from Life Technologies. Each primer was rehydrated to a final concentration of 1 nmol/µl and stored at –80°C.

2.1.10.1 Oligonucleotides Used for PCR and DNA Sequencing

name	DNA sequence	orientation	application
ITS4	tcc tcc gct tat tga tat gc	ITS2, 3'	amplification and sequencing of ITS1, ITS2 and 5.8S of nuclear rDNA
ITS5	gga agt aaa agt cgt aac aag g	ITS1, 5'	
ITS2	gct acg ttc ttc atc gat	ITS1, 3'	sequencing of ITS1
ITS3	gca tcg atg aag aac gta	ITS2, 5'	sequencing of ITS2
***ndh*F 972F**	gtc tca att ggg tta tat gat g	*ndh*F 5'	amplification and sequencing of the *ndh*F gene
***ndh*F 2110R**	ccc yab ata ttt gat acc ttc kcc	*ndh*F 3'	
***ndh*F MR**	gta ctc att cca ttt ctc tat gg	*ndh*F 5'	sequencing of the *ndh*F gene

2.1.10.1 Oligonucleotides Used for PCR and DNA Sequencing (continued)

name	DNA sequence	orientation	application
Nm-af2AF	ctg gta ttc tgt aaa gaa aat g	Nm-af2 5'	amplification and sequencing of an AFLP derived marker Nm-af2 for *Nothofagus* spp.
Nm-af2AR	cat gtt cga tgt acc atg act tc	Nm-af2 3'	
Nm-af7AF	ttc aat aag cca caa agg ca	Nm-af7 5'	amplification of the AFLP derived marker Nm-af7 for *Nothofagus* spp.
Nm-af7AR	cag tgt att agt tgc gtg g	Nm-af7 3'	
Ma-af1AF	gtt cag aaa tag gtg tc	Ma-af1 5'	amplification and sequencing of the AFLP derived marker Ma-af1 for *Myrsine* spp.
Ma-af1AR	caa tgc ctt agc caa ttt c	Ma-af1 3'	
Ma-af1BF	ctg aaa ttg taa atg atc g	Ma-af1 5'	nested PCR and sequencing of the AFLP derived marker Ma-af1 for *Myrsine* spp.
Ma-af1BR	gca tta ttc aat ttt tgc	Ma-af1 3'	
Md-af3AF	ctc gtg gta tgg gaa aga ctg g	Ma-af3 5'	amplification of the AFLP derived marker Md-af3 for *Myrsine* spp.
Md-af3AR	gct cca ggc cgt gaa cct tcc g	Ma-af3 3'	
Md-af9AF	cgt aat aca ata ttg tcc g	Md-af1 5'	amplification and sequencing of the AFLP derived marker Md-af9 for *Myrsine* spp.
Md-af9AR	gca cat gtt tta cca tta cc	Md-af1 3'	
Md-af9BF	gct ttg ggc aaa gtt ata gc	Md-af1 5'	nested PCR and sequencing of the AFLP derived marker Md-af9 for *Myrsine* spp.
Md-af9BR	gtt tta cca tta cca cak aag	Md-af1 3'	
Md-af11AF	tac gat act tgg tat gcg	Md-af11 3'	amplification of the AFLP derived marker Md-af11 for *Myrsine* spp.
Md-af11AR	cac taa tta ttg cat tca tg	Md-af11 5'	
M13 Forward	ccc agt cac gac gtt gta aaa cg	3' end of plasmid cloning site	amplification and sequencing of DNA fragments cloned into the pGEMT vector
M13 Reverse	agc gga taa caa ttt cac aca gg	5' end of plasmid cloning site	
***rbc*L af**	atg tca cca caa aca gag act aaa gc	*rbc*L 5'	amplification and sequencing of the *rbc*L gene
***rbc*L-cR**	gca gca gct agt tcc ggg ctc ca	*rbc*L 3'	
***rbc*L 90F**	cat ggt atg cac ttt cgt g	*rbc*L 5'	sequencing of the Araucariaceae *rbc*L gene
***rbc*L 42F**	gat cta cga att cct cct tc	*rbc*L 5'	
***rbc*L48R**	cac ccc aac aag gga cgg	*rbc*L 3'	
***rbc*L102R**	gtt cac ctt caa gtt tac c	*rbc*L 3'	
TabC	cga aat cgg tag acg cta cg	*trn*L intron 3'	amplification and sequencing of the *trn*L intron
TabD	ggg gat aga ggg actt gaa c	*trn*L intron 5'	
TabE	ggt tca agt ccc tct atc cc	*trn*L - *trn*F spacer 3'	amplification and sequencing of the *trn*L - *trn*F intergenic spacer
TabF	att cga act ggt gac acg ag	*trn*L - *trn*F spacer 5'	

2.1.10.2 Oligonucleotides Used for AFLP

name	DNA sequence	description	application
Mse adapter I	5'-GACGATGAGTCCTGAG-3'	the double stranded adapter is ligated to DNA *Mse* I restriction ends	serves as priming site for '*Mse*' amplification primers
Mse adapter II	3'- TACTCAGGACTCAT-5'		
Mse -Preamp	5'-GATGAGTCCTGAGTAAC-3'	primers based on *Mse* adapter sequence	preselective AFLP amplification
Mse–CAG	5'-GATGAGTCCTGAGTAACAG-3'		selective AFLP amplification
Eco adapter I	5'-CTCGTAGACTGCGTACC-3'	the double stranded adapter is ligated to DNA *Eco*RI restriction ends	serves as priming site for '*Eco*' amplification primers
Eco adapter II	3'-CTGACGCATGGTTAA-5'		
Eco-Preamp	5'- GACTGCGTACCAATTCA-3'	primers based on *Eco* adapter sequence	preselective AFLP amplification
Eco-ATA	5'- GACTGCGTACCAATTCATA-3'		selective amplification
Eco-ACG	5'- GACTGCGTACCAATTCACG-3'		

2.3 METHODS

2.3.1 DNA Extraction from Plant Material

DNA was extracted from fresh and dry plant tissue using a protocol modified from Doyle and Doyle (1987). The method uses cetyl trimethyl ammonium bromide (CTAB) to bind DNA. After the DNA-CTAB complex is separated from cell debris by chloroform extraction the DNA can be precipitated from a liquid phase using standard methods. As the extraction pellet often contained high amounts of polysaccharides, a selective precipitation of polysaccharides was performed following a protocol modified from B. Hurley (pers. comm.). This protocol uses a high NaCl concentration to keep DNA in solution while polysaccharides are pelleted by centrifugation.

Two different CTAB extraction buffers were used for DNA extractions from plant material. For plants with high carbohydrate contents (e.g. *Myrsine* spp.), polyvinyl-pyrrolidone (PVP, Mw=25 000,) was used for DNA extraction. With other taxa (e.g. *Nothofagus* spp.) PVP was exchanged for polyethylene glycol (PEG, Mw=8000) which seemed to increase the yield. In each case CTAB was added to a sterile buffer stock solution immediately before use and allowed to dissolve at 60°C.

	CTAB buffer A	CTAB buffer B
	1% [w/v] PVP 25 000	1% [w/v] PEG 8000
	1.4 M NaCl	1.4 M NaCl
	100 mM Tris-HCl [pH 8]	100 mM Tris-HCl [pH 8]
	20 mM EDTA	20 mM EDTA
	1% [w/v] CTAB	1% [w/v] CTAB
	(added immediately before use)	(added immediately before use)

Table 2.1: Composition of two CTAB-based buffers used for DNA extraction.

Procedure: Prior to processing the dried plant samples, it was necessary to prepare a tool for grinding the tissue. A glass pasteur pipette was heated in a bunsen flame, close to where the pipette narrowed, until the two sections separated. The thicker section of the pipette was again passed through the flame to create a small bulb at the end that could be used to grind the frozen tissue sample in a 1.6 ml microcentrifuge tube. One tool was prepared for each sample to be processed.

A small amount (ca. 30 mg) of dried tissue was transferred, using clean tweezers, to a fresh 1.6 ml microcentrifuge tube. The tube was then suspended in liquid nitrogen to rapidly freeze the dried tissue fragments. Once the contents were frozen the tube was removed from the liquid nitrogen and the tissue sample crushed using the glass tool. In order to thoroughly disrupt the plant material it was usually necessary to dislodge the sample from the bottom of the tube, and refreeze the sample before continuing to grind. A 600 µl aliquot of freshly prepared CTAB extraction buffer (see above) was added and mixed with the finely ground tissue using the glass tool and gentle agitation. The tube was then placed in a heating block at 65°C, and incubated for approximately 40-60 minutes, with occasional inversion and gentle agitation.

Following incubation, the tube was allowed to cool for approximately 5 minutes, so that the temperature of the solution dropped below 65°C. A 600 µl aliquot of chloroform was then added to the extraction mixture. The two solutions were mixed by repeated inversion and then left to stand for 2-3 minutes to allow separation of the aqueous and organic phases. The tube was then centrifuged at 10 000 × g and room temperature for 1 minute, to completely separate the phases and compact the interface. After the upper aqueous phase of the extraction was transferred to a clean 1.6 ml microcentrifuge tube. A wide bore tip was used to minimise possible mechanical shearing of the DNA due to pipetting.

The DNA was then collected by precipitation with one volume of isopropanol. To the collected aqueous solution 600 µl isopropanol were added; the two solutions were mixed by gentle inversion and the tube incubated on ice for 5 minutes. After incubation, the solution was centrifuged at 12 000 x g at room temperature for 1 minute to collect the DNA at the bottom of the tube. The supernatant was discarded and the pellet was washed twice with 1 ml of 70% (v/v) ethanol to remove salt residues. The DNA pellet was air dried before being resuspended in 400 µl sterile water and allowed to dissolve over night at 4°C.

After the DNA was completely dissolved, 100 µl of 5 M NaCl and 80 µl of 80% (v/v) ethanol were added to the DNA. The solution was mixed by inversion and left on ice for five minutes. After a centrifugation at 20 000 x g at 4°C carbohydrates collected at the bottom of the reaction tube while DNA was still in solution. The supernatant was carefully removed using a wide bore 1000 µl pipette tip and transferred into a new reaction tube. The DNA was again precipitated by adding one volume of isopropanol and washed twice with 70% (v/v) ethanol as described above. The DNA pellet was air dried before being resuspended in 25-40 µl of TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]). A 5 µl aliquot of each solution was electrophoresed as described in Section 2.3.2.

The amount of DNA obtained with this extraction method depended on the age and quality of the dried plant material. Good quality plant material produced 100 - 200 µg of DNA.

2.3.2 Gel Electrophoresis and Visualisation of DNA

Gel electrophoresis can be used to separate DNA molecules based on their size. DNA has a negative charge in solution at pH 8.0, so it migrates to the cathode in an electric field. In a gel matrix, large pieces of DNA move slower than small pieces of DNA and the different sized fragments are separated from each other. Several techniques can be used to visualise DNA in the gel matrix.

2.3.2.1 Agarose Gel Electrophoresis

The standard method used to separate and identify DNA fragments and genomic DNA was electrophoresis through agarose gels. Horizontal agarose gel electrophoresis was performed in a Horizon® 11.14 mini gel apparatus following standard protocols as described by Sambrook *et al.* (1989).

Typically, agarose was added to 1 x TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH=8.0), to a concentration of 0.8-1% (w/v) depending on the size of the fragments to be resolved. The mixture was heated in a microwave oven until the agarose was completely dissolved, and cooled down to approximately 55°C before being poured into the gel tray.

DNA samples were combined with 1-2 µl of 10× loading buffer (27.5% [w/v] Ficoll Ty 400, 0.44% [w/v] bromophenol blue and 0.44% [w/v] xylene cyanol) and the mixture loaded onto the agarose gel. Either the 1kb™ or 1kb plus™ DNA were routinely included on electrophoretic gels as size standards. Samples were electrophoresed at approximately 3-5 V/cm in 1× TAE buffer.

Following electrophoresis the DNA samples were visualised by ethidium bromide fluorescence on a UV transilluminator (wavelength 302 nm,). Bands of DNA in the gel are stained with the intercalating dye ethidium bromide; as little as 1 ng of DNA can be detected by direct examination of the gel in ultraviolet light (Sharp *et al.* 1973). A digital photograph of the illuminated gel was taken using a video camera and ImagePC software.

2.3.2.2 Denaturing Polyacrylamide Gel Electrophoresis

Polyacrylamide has been used as matrix for gel electrophoresis of DNA and proteins. As it allows the separation of DNA fragments which differ in size by only one bp polyacrylamide has been used for DNA sequencing (see Section 2.3.8). Crystallised acrylamide can be dissolved in water and its polymerisation initiated by addition of a strong oxidant like ammonium persulphate (APS) and a catalyst (TEMED). DNA gel electrophoresis is often performed under denaturing conditions, which occur when urea is present in the matrix.

Preparation of Denaturing Poylacrylamide Gels

Prior to pouring the polyacrylamide gel, the glass plates were cleaned with detergent under hot running water and then repeatedly rinsed to remove any soapy residue. One side of the long glass plate was treated with a 'repel-agent' to prevent the polyacrylamide matrix from affixation. This was achieved by applying a small amount of RAIN-X®, 2-3 ml, to the glass surface. The plate was washed once again under hot water.

Both plates were then rinsed with Milli-Q H_2O and dried with a clean paper towel. One side of each was thoroughly wiped with ca. 2 ml ethanol and polished with a dry paper towel. This procedure was repeated twice more for each plate to completely remove dust particles from the surfaces.

The cleaned face of the shorter glass plate was then treated with a Bind-Silane solution to ensure that the polyacrylamide matrix remains attached to the glass plate after gel electrophoresis. The Bind Silane solution consisted of 2 ml 95% (v/v) ethanol, 0.5% (v/v) glacial acetic acid and 0.05% (v/v) Bind-Silane and was allowed to dry for 4-5 minutes. Excess Bind-Silane was removed by cleaning the plate three times with 2 ml of 95% (v/v) ethanol. Once dry, both plates were assembled with a pair of 0.4 mM spacers in position, to form a glass sandwich. The assembly was then clamped together using a S2 casting boot. It was important during this procedure not to allow the treated faces of the glass plates to come into contact, as the resulting interaction caused the gel to bind to both surfaces.

A 5% (w/v) denaturing polyacrylamide gel solution was prepared which contained 8 M urea, 10% (v/v) Long Ranger gel solution, and 1× Tris borate EDTA (TBE) buffer (90 mM Tris-borate and 1 mM EDTA [pH 8.0]) in a total volume of 70 ml. The solution was then filtered through 2 pieces of Whatman 1 filter paper. Before pouring the gel, 350 µl of 10% (w/v) ammonium persulphate (APS) and 35 µl NNN'N'-tetramethylethylenediamine (TEMED) were added and the solution mixed with gentle agitation. Using a 250 ml beaker, the gel solution was slowly dispensed into the glass sandwich taking particular care not to introduce air bubbles. Once the gel was poured, a pair of sharktooth combs (5.7 mM point-to-point spacing) were inserted, in reverse orientation, to form the sample well. The combs were clamped into position with several large bulldog clips. To allow complete polymerisation the gel was left at room temperature for at least 1 hour. The polyacrylamide gel was routinely covered loosely with plastic film and stored overnight.

Preparation of DNA Samples for Polyacrylamide Gel Electrophoresis

DNA samples that were to be loaded onto the gel (typically PCR products or AFLP selective amplifications) were combined with 1/2 a volume of formamide loading dye (98% [v/v] formamide, 10 mM EDTA, 0.05% [w/v] bromophenol blue and 0.05% [w/v] xylene cyanol). The solutions were prepared in a 0.2 ml reaction tube and denatured at 94°C for 4 minutes in the thermocycler, then rapidly cooled to 4°C. Depending on the DNA concentration of the sample 3-7 µl were loaded onto the polyacrylamide gel.

The DNA size standard (100 bp DNA Ladder) electrophoresed with AFLP reactions was prepared in volumes of 100 µl by diluting 1 µl of the ladder (1 µg/µl) in 99 µl of formamide loading dye. The probe was denatured as described above and placed on ice until loaded. 5 µl of this solution were applied. After initial use, the ladder was stored at –20°C; denaturation was not repeated prior to subsequent applications.

Loading and Running of Denaturing Polyacrylamide Gels

Prior to gel electrophoresis the casting boot and bulldog clips were removed and the outside of the glass plates cleaned. The gel sandwich was then placed into the S2 electrophoresis apparatus, and each of the buffer tanks filled with approximately 600 ml of 1× TBE buffer (90 mM Tris-borate and 1 mM EDTA [pH 8.0]). After removing the combs, the gel was pre-run at 35-40 W for 30 minutes to obtain optimal running temperature between 45 and 50°C.

After the pre-run the sample well was flushed out using a syringe filled with 1× TBE. The pair of sharktooth combs were inserted into the sample well, taking care not to trap any loose polyacrylamide fragments and ensuring that the points of the teeth were resting on the gel surface rather than piercing it. Before dispensing a sample onto the gel, the well into which it was to be loaded was again flushed with a 20 µl pipette which was also used for loading the DNA sample. Samples were usually loaded side by side, although gaps of one or a few wells were routinely left between sets of reactions for ease of interpretation.

The gel was typically run at 35-40 W until the bromophenol blue dye front was 3-6 cm from the bottom of the polyacryamide gel. After completion of the electrophoresis procedure the gel sandwich was removed from the apparatus and the glass plates immediately separated by inserting a scalpel blade between them. Optimally, the polyacrylamide matrix adhered to the small glass plate which had been treated with

Bind Silane (see 2.3.2.2.1). The gel was immediately taken through the silver staining procedure.

Silver Staining of Denaturing Polyacrylamide Gels

The silver staining method of DNA is a highly sensitive detection of nucleic acids in the nanogram range, which is achieved by the specific chemical reduction of silver ions. Bassam *et al*. (1991) have optimised a silver stain method originally introduced by Merril *et al*. (1981) which uses formaldehyde to selectively reduce silver ions to metallic silver under alkaline conditions. The limit of detection upon visual inspection of double-stranded DNA was approximately 1 pg/mm. This is about 1 000 to 10 000 times more sensitive than ethidium bromide staining (see Section 2.3.1).

Procedure: The short glass plate, to which the acrylamide gel adhered, was transferred to a tray containing 4 l 10% (v/v) acetic acid solution. This was agitated on a mechanical shaker for 2 hours to remove the urea from the gel. After the wash, 2 l of the acetic acid solution were retained and stored at –20°C for later fixing of the gel. The gel was transferred to a second developing tray that contained 2 l Milli-Q H_2O and agitated for 2-3 minutes. This water rinse step was repeated twice more, each time transferring the gel to a developing tray containing clean Milli-Q H_2O to remove acetic acid. After the final water rinse, the gel was placed in 3 l of stain solution consisting of 6 mM silver nitrate ($AgNO_3$) and 0.15% (v/v) formaldehyde and gently agitated for 1 hour.

As the developing step required the use of chilled solutions, 4 l Milli-Q H_2O and 4 l of a 280 mM sodium carbonate (Na_2CO_3 anhydrous, AnalaR,) solution had been placed at –20 °C. Following the staining the gel was quickly rinsed, (approximately 5 seconds, in 2-3 l of chilled Milli-Q H_2O) and immediately transferred to a developing tray containing 2 l of cold developing solution, which was made immediately before use by adding 0.16% (v/v) formaldehyde (37%) and 50 µM sodium thiosulphate ($Na_2S_2O_3$) to the chilled 280 mM sodium carbonate solution. The formaldehyde reduces silver ions attached to the DNA particles while thiosulfate dissolves access silver ions from the gel surface, and thus decreases nonspecific staining (Blum *et al.* 1987). The sodium carbonate acts as a buffering agent.

The tray was vigorously agitated to disperse reduced silver ions and remove them from the gel surface. After approximately 5 minutes DNA bands became visible, at which point the gel was transferred to a second tray containing the remaining 2 l of chilled developing solution. Agitation was continued until the bands were clearly discernible, taking care not to over-develop the gel which resulted in a dark background. The developing reaction was then stopped by the addition of the chilled acetic acid solution. To achieve the best results, the gel was usually removed from the developing solution, the acetic acid added and the two solutions mixed vigorously by hand before the gel was returned to the tray. The gel was mechanically agitated for a further 2-5 minutes while carbon dioxide was released. The gel plate was removed from the tray when the gas production weakened. A 2 minute rinse in Milli-Q H_2O was then used to remove the acetic acid from the gel. The gel was allowed to air dry overnight at room temperature.

2.3.3 Polymerase Chain Reaction (PCR)

The Polymerase Chain Reaction (PCR) is an *in vitro* method for the enzymatic synthesis of specific DNA sequences (Mullis and Faloona, 1987). It uses two oligonucleotide primers that hybridise to opposite strands and flank the region of interest in the target DNA. A repetitive series of cycles involving template denaturation, primer annealing, and the extension of the annealed primers by DNA polymerase results in the exponential accumulation of a specific fragment whose termini are defined by the 5' ends of the primers. Because the primer extension products synthesised in one cycle can serve as a template in the next, the number of target DNA copies approximately doubles at every cycle. Thus, 20 cycles of PCR yields about a million-fold amplification. The introduction of the thermostable DNA polymerase (*Taq* polymerase, Saiki *et al.*, 1988) isolated from *Thermus aquaticus* transformed the PCR into a simple and robust reaction which could be automated by a thermal cycling device.

The PCR is effected by varying the reaction parameters (e.g. enzyme, primer and Mg^{2+} concentration as well as the temperature cycling protocol). Although, for any given pair of oligonucleotide primers, an optimal set of conditions can be established, there is no single set of conditions that will be optimal for all possible reactions.

2.3.3.1 Amplification of DNA Regions Using a One-Step PCR Protocol

PCR reactions were performed in a final volume of 20 µl containing 1 U *Taq* polymerase, 1 x PCR buffer (Tris-HCl, KCl, $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, (pH 8.7); supplied by the manufacturer), 10 pmol of each primer, 250 µM of each dNTP and 10-100 ng DNA template. PCR reactions were routinely supplemented by 1 x Q-solution as recommended by the manufacturer.

To minimise the possibility of cross-contamination between reactions when preparing multiple amplifications the DNA template was always added last. One negative control, containing no DNA template, was always run with each reaction set.

All the thermocycling programs included an initial hold for 2 minutes at 94°C to completely denature the double stranded DNA templates and a final 5 minute extension time at 72°C to ensure that the reaction products were fully extended. The ramping rate, the speed at which the thermocycler moves between temperatures, was the default value for the thermocycler (approximately 1-2°C/second) except during amplification of the JSAD locus when this was limited to 1°C/second. Thermocycling was carried out using a heated lid to prevent evaporation of the reaction mixture.

The combinations of oligonucleotide primers and amplification protocols used for amplification of specific marker loci are detailed in Table 2.2.

Following thermocycling, a 2-4 µl aliquot of each amplification was electrophoresed on a 1% (w/v) LE agarose gel as described in Section 2.3.2.1 to determine the efficacy, by comparison to the DNA ladder, of individual reactions.

amplified region	primer pair	template	PCR program
ITS1, 5.8S and ITS2 of nuclear rDNA)	ITS5 - ITS4	genomic DNA	96 °C for 2 min [(94°C for 1 min; 43 °C for 1 min; 72 °C for 1 min 10 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
ndhF gene, partial product	*ndh*F 972F - *ndh*F2110R	genomic DNA	96 °C for 2 min [(94°C for 1 min; 43 °C for 1 min; 72 °C for 1 min 10 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
Ma-af1	Ma-af1 AF - Ma-af1 AR	genomic DNA	96 °C for 2 min [(94°C for 1 min; 47 °C for 1 min; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
Md-af9	Md-af9 AF - Md-af9 AR	genomic DNA	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
Nm-af2	Nm-af2 AF - Nm-af2 AR	genomic DNA	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 50 sec) 34 cycles] 72 °C for 5 min; hold at 4 °C
***trnL* intron**	TabC - TabD	genomic DNA	96 °C for 2 min [(94°C for 1 min; 50 °C for 50 sec; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
trnL . trnF- intergenic spacer	TabE - TabF -	genomic DNA	96 °C for 2 min [(94°C for 1 min; 50 °C for 50 sec; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
***rbcL* gene, partial sequence**	*rbc*L aF - *rbc*L cR	genomic DNA	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 1 min 10 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C

Table 2.2: Thermocycling protocols for amplification of 6 gene regions from genomic DNA

2.3.3.2 'Nested' and 'Semi-Nested' PCR

When DNA regions amplified only weakly, it was necessary to use a nested or semi-nested PCR approach to amplify the target region. Both, nested and semi-nested PCR use templates produced in a one-step PCR protocol described above in Section 2.3.3.1. While semi-nested PCR uses one of the original amplification primers and a second primer, with a binding site within the target fragment, nested PCR uses two internal primers to amplify a portion of the desired locus.

Template produced by the one-step PCR protocol was diluted 1:10 in sterile water and added to a PCR cocktail, prepared as described in Section 2.3.3.1. Regions amplified with nested and semi-nested PCR, primer combinations and conditions for PCR are listed below in Table 2.3; primer sequences are presented in Section 2.1.10.1.

Following thermocycling, the efficacy of the amplifications was determined, as previously, by electrophoresing a 2 µl aliquot of each reaction on a 1% (w/v) LE agarose/TAE gel.

amplified region	primer pair	template	PCR program
Ma-af1	Ma-af1 AF - Ma-af1 BR Ma-af1 BF - Ma-af1 AR	Ma-af1	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
Md-af9	Md-af9 AF - Md-af9 BR Md-af9 BF - Md-af9 AR	Md-af9	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C
Md-af9	Md-af9 BF - Md-af9 BR	Md-af9	96 °C for 2 min [(94°C for 1 min; 50 °C for 1 min; 72 °C for 50 sec) 35 cycles] 72 °C for 5 min; hold at 4 °C

Table 2.3: Thermocycling protocols for amplification of 3 gene regions from PCR products (semi-nested and nested PCR approaches).

2.3.4 Purification of PCR Products and Plasmid DNA

PCR products were purified using the Concert™ PCR Purification kit according to the instructions of the manufacturers. Briefly, the DNA sample was diluted in the respective binding buffer, mixed, and bound to the silica-gel membrane in the presence of chaotropic salts by centrifugation. One washing step was performed using the solutions supplied by the manufacturer, and the bound DNA was eluted using low salt solutions, e.g., TE buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA) or sterile water.

2.3.5 Cloning Procedures

2.3.5.1 Ligation of PCR Products into a Plasmid Vector

PCR products obtained by amplification with *Taq* polymerase have an A overhang at the 3'end which is added by the enzyme. This overhang can be used efficiently for cloning PCR products into so-called T vectors. In this study the pGEM®-T Easy Vector System supplied by Promega was used for cloning.

Procedure: PCR products were purifed using the Concert™ PCR Purification kit as described in Section 2.3.4. Ligations of PCR products into the pGEM®-T Easy Vector System were performed in 10 µl volumes using 1 x Ligation buffer, 3 Weiss units T4 DNA Ligase, and 50 ng pGEM®-T Easy Vector supplied by the manufacturer. The molar ratio of insert:vector was 3:1 and the amount of PCR product needed for ligation reaction was calculated as recommended by the manufacturer by the following equation:

$$\frac{\text{ng of vector x kb of insert}}{\text{3.0 kb vector}} \text{ x insert:vector molar ratio} = \text{ng of insert}$$

Ligations were performed overnight at 4°C.

2.3.5.2 Transformation of Ca^{2+} Competent Cells

After treating *Escherichia coli* bacteria (*E. coli*) with $CaCl_2$, the cell walls are permeable for DNA fragments, and the cells are able to incorporate plasmid DNA (Cohen *et al.* 1972). Treated bacteria (Ca^{2+} competent cells) can be stored at -80 °C until required for transformation.

Procedure: Per transformation, 30 µl of high efficiency competent *E.coli* DH5α cells were thawed on ice for 30 min. The whole ligation reaction prepared as described in the previous Section 2.3.5.2, was added to the cells and incubated in ice for 45 minutes to allow the plasmid DNA to anneal to the cell surface. Incorporation of the non-chromosomal DNA was obtained by 'heat shock treatment' (90 sec at 42°C). The reaction was placed on ice for additional 2 minutes and transferred into 600 µl of Luria-Bertani (LB) broth (composition per litre: 10 g bacto-tryptone, 5 g yeast extract, 10 g NaCl, pH=7; Sambrook *et al.*, 1989). After incubation at 37°C on a shaker at 200 rpm for 45 min, the whole cell suspension was plated on selective agar plates (composition per litre: 10 g bacto-tryptone, 5 g yeast extract, 10 g NaCl, 15 g bacto-agar, 0.6 mg Ampicillin, 0.8 mg Xgal, 0.8 mg IPTG, pH=7; Sambrook *et al.*, 1989), and incubated at 37°C overnight.

2.3.5.3 Selection of Transformed Cells

After transformation cells were plated on selective agar plates containing ampicillin (see Section 2.3.5.2). The plasmid vector contains a gene coding for ampicillin resistance, which allows cells containing the vector to grow on the selective medium. Untransformed cells do not grow in the presence of the antibiotic.

Recombinant cells (cells containing the plasmid with DNA insert) can then be identified with the blue-white screening (Sambrook *et al.* 1989). The method is based on a colour detection: recombinant colonies show a white colour while non-recombinant cells are blue. With self-religation of the plasmid a *lac*Z gene which encodes a ß-galactosidase enzyme will be reassembled. This gene is expressed in the presence of IPTG and cleaves X-gal. The blue cleavage product is stored within the cells. In cells with a recombinant plasmid the ß-galactosidase enzyme is inactive, cells remain white.

2.3.5.5 Isolation of Plasmid DNA from *E. coli*

Plasmid preparations were performed by the method described by Holmes and Quigley (1981). A single colony of transformed *E. coli* DH5α cells was grown overnight in LB broth supplemented with ampicillin (final concentration 80 µg/ml). An aliquot of 1.5 ml was pelleted by centrifugation at 12 000 x g in a microcentrifuge at room temperature (RT) for 1 min, and subsequently resuspended in 400 µl STET buffer (10 mM Tris-HCl (pH 8.0) 0.1 M NaCl, 1 mM EDTA, 5% (w/v) Triton X-100). The suspension was treated with 0.2 mg/ml RNase A and 2 mg/ml lysosyme, left for 5-10 min at RT and then boiled for 40 sec, chilled on ice for 60 sec and immediately spun at 12 000 x g for 10 min at RT. The cell debris was removed using a sterile toothpick and plasmid DNA precipitated for 5 min at RT by the addition of 40 µl 3 M sodium acetate (pH = 5.2) and 420 µl 2-isopropyl alcohol. DNA was pelleted by centrifugation at 18 000 g for 10 min at RT, washed with 70% ethanol, and resuspended in 60 µl TE buffer (10 mM Tris-HCl [pH 8.0] / 1 mM EDTA)

The DNA concentration of the solution was measured by gel electrophoresis by visual comparison to a molecular weight marker in the presence of ethidium bromide.

2.3.5.6 Analytical Restriction of Plasmid Vectors

To confirm that PCR amplified fragments of the appropriate size had been successfully cloned into the pGEM-T® Easy vector, an analytical restriction digest was performed. During this procedure the cloned fragment was excised and separated from the vector sequence by agarose gel electrophoresis (see Section 2.3.2.1).

Procedure: The restriction digests were carried out in 1.6 ml microcentrifuge tubes using a total volume of 10 µl containing 10 U *Eco* RI,1 × NEBuffer for *Eco* RI and 200-300 ng plasmid DNA. Before incubation the mixtures were gently agitated and briefly centrifuged to collect the contents at the bottom of the tube. Digests were performed at 37°C for 30 minutes. The reactions were terminated by adding 1/10 volume of loading dye and placed on ice before being analysed by gel electrophoresis on a 1% (w/v) agarose/1× TAE gel. The size of cloned PCR fragments was determined by comparison to the DNA ladder.

2.3.5.7 Characterisation of Cloned Inserts by PCR

As an alternative to the analytical restriction digest (Section 2.3.5.6) a PCR based approach was used to screen transformed colonies. The region of the plasmid containing the insertion site was amplified directly from bacterial cells using specific PCR primers.

Reaction cocktails were prepared as described in Section 2.3.3 and contained 10 pmol M13 forward and 10 pmol M13 reverse primers. A clean 20 µl pipette tip was used to transfer cells from one cell colony to the PCR reaction tube. Following this step, the same tip was dropped into a glass tube to inoculate 5 ml of selective LB broth (containing 80 µg/ml ampicillin) with the transformed cells.

Thermocycling consisted of an initial hold at 94°C for 3 minutes to ensure cell lysis and to denature the double stranded DNA templates, followed by 30 cycles of 30 seconds at 94°C, 30 seconds at 65°C and 1 minute at 72°C. A final extension period of 72°C for 5 minutes was also incorporated. After thermocycling the reactions were held at 4°C. A 1-2 µl aliquot of each PCR amplification was analysed by gel electrophoresis on a 1% (w/v) LE agarose/TAE gel as described in Section 2.3.2.1 to determine the efficacy, by comparison to the DNA marker, of the PCR. Colonies carrying inserts of the appropriate size were identified and plasmids were isolated as described in Section 2.3.5.3.

2.3.6 Amplified Fragment Length Polymorphism (AFLP)

The AFLP technique, originally known as selective restriction fragment amplification (SRFA) (Zabeau and Vos, 1993) is a novel DNA fingerprinting method. It amplifies restriction fragments obtained by endonuclease digestion of target DNA using 'universal' AFLP primers complementary to the restriction site and adapter sequence. However, not all restriction fragments are amplified because AFLP primers also contain selective nucleotides at the 3' termini that extend into the amplified restriction fragments. These arbitrary terminal sequences result in the amplification of only a small subset of possible restriction fragments.

The use of endonuclease combinations that vary in their restriction frequency can also be used to vary the number of fragments. Generally, the abundant restriction fragments produced from complex genomes require AFLP primers with longer selective regions. Conversely, analysis of small genomes requires only few arbitrary nucleotides added at the primer 3' termini. AFLP fingerprints are usually a rich source of DNA polymorphisms that can be used in mapping and general fingerprinting endeavours.

The AFLP protocol can be divided in the following steps: (1) DNA digestion with two different restriction enzymes, (2) ligation of double-stranded adapters to the ends of the restriction fragments, (3) optional DNA pre-amplification of ligated product directed by primers complementary to adapter and restriction site sequences, and (4) DNA amplification of subsets of restriction fragments using selective AFLP primers and (5) detection of amplified products. An overview of the procedure is shown in Figure 2.1.

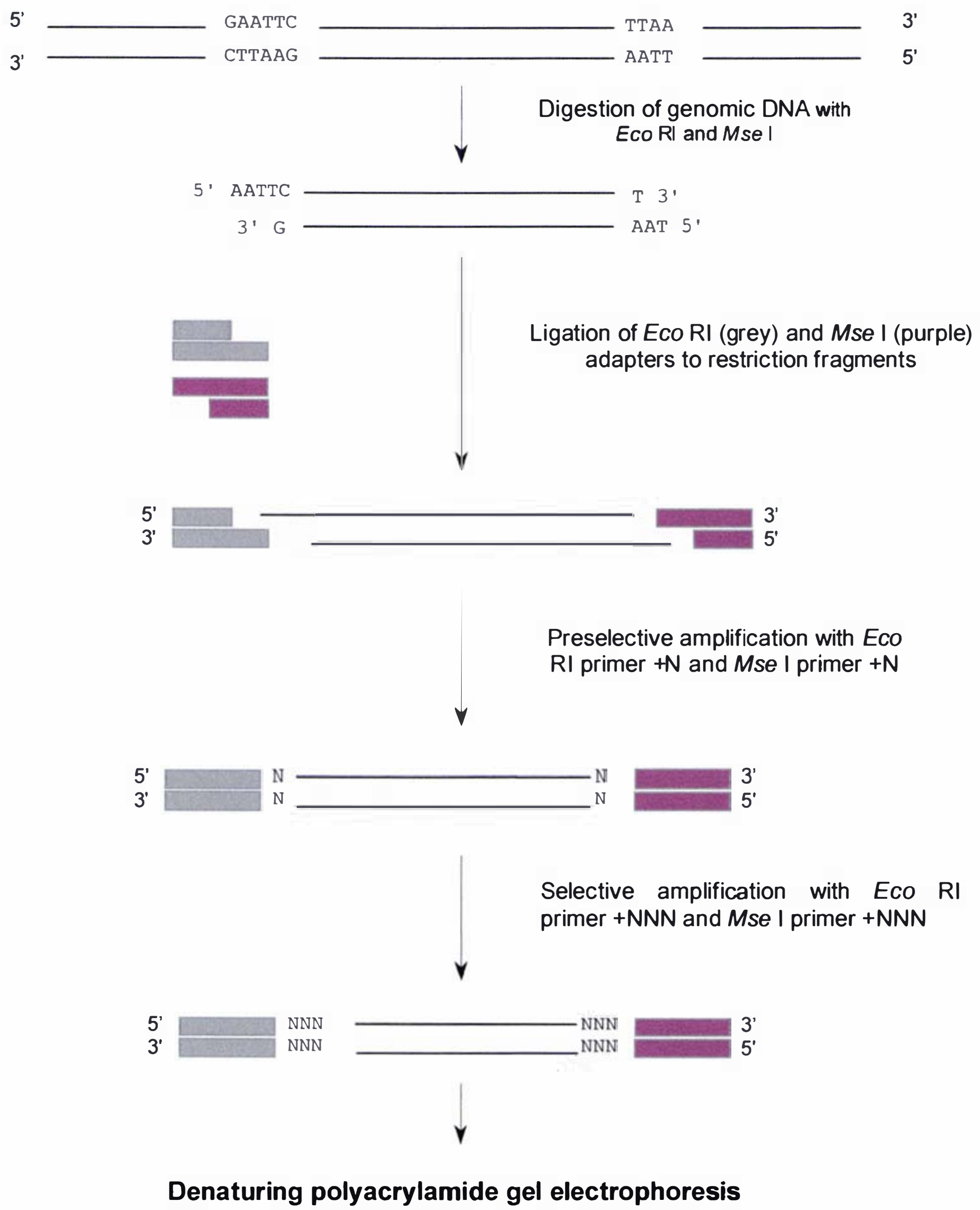


Figure 2.1: Overview of different steps performed under the protocol for Amplified Fragment Length Polymorphism (AFLP). Modified from Winkworth (2000).

In contrast to the small genomes of e.g. micro-organisms, plant genomes are complex and thus require the use of more than 2 selective bases in one or both primers to reduce the number of DNA fragments amplified. It is also advantageous to carry the amplification in two consecutive steps (preamplification and selective amplification) to increase specificity and the amount of initial template. The protocol used in this thesis work was modified from the GIBCO BRL AFLP Analysis System I manual with the modifications of Lockhart and McLenachan (1997; [http:// www2.lifetech.com/ focus_page.html]).

2.3.6.1 AFLP Step 1: Restriction of Genomic DNA

150 –200 ng of genomic DNA was restricted with 4 U *Mse* I and 1 U *Eco* RI in a total volume of 25 µl at 37°C for 2 hours. The 1 x reaction buffer contained 50 mM potassium acetate (KOAc), 10 mM magnesium acetate (MgOAc), 10 mM Tris-HCl (pH 7.5). To ensure complete digestion the tubes were agitated and then briefly centrifuged at intervals during the incubation. After incubation the tubes were heated at 70°C for a further 15 minutes to denature the restriction enzymes. Digestions were either used immediately or stored at -20°C until use. The efficacy of genomic DNA digestion was determined by gel electrophoresis: 5 µl aliquots of each digestion reaction were electrophoresed on 1% (w/v) agarose/TAE gel as previously described in Section 2.3.2.1. Typically, a smear of low molecular weight DNA, between 100 base pairs (bp) and 1.5 kilobase pairs (kb) was observed.

2.3.6.2 AFLP Step 2: Ligation of Adapter Sequences

A pair of double-stranded adapters, which provided targets for the pre-selective and selective amplification primers, were prepared by annealing two single-stranded oligonucleotides (see Figure 2.1). A reaction cocktail was prepared for each type of adapter. As the adapter concentration depends on the restriction frequency of the enzyme more *Mse* I adapter than *Eco* RI adapter was used. The *Eco* RI adapter cocktail contained 45% (v/v) TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]), 5 pmol *Eco* RI adapter I and 5 pmol *Eco* RI adapter II in a total volume of 100 µl. The *Mse* I adapter mixture consisted of 45% (v/v) TE buffer, 50 pmol *Mse* I adapter I, and 50 pmol *Mse* I adapter II again in a total volume of 100 µl. Sequences of the oligonucleotides used to prepare AFLP adapters are given in Section 2.1.10.2. Both cocktails were incubated at 94°C for 4 minutes before being allowed to cool slowly, on the benchtop, to room temperature.

The oligonucleotide adapters were ligated to the restricted ends of the genomic DNA fragments (see AFLP step 1) in a reaction volume of 20 µl. The reaction cocktail contained 60-80 ng of digested genomic DNA, 50 pmol *Mse* I adapter, 5 pmol *Eco* RI adapter, 1 Weiss U T4 DNA ligase and 1 x Ligation buffer supplied by the manufacturer. The tubes were gently agitated to mix and then briefly centrifuged to collect the contents at the bottom of the tube. Ligation reactions were incubated at 4°C overnight; the following day reactions were either used immediately or transferred to –20°C for storage until use.

2.3.6.3 AFLP Step 3: Preamplification

All pre-selective amplifications were performed using oligonucleotide primers that had a single base overhang – an adenosine on the *Eco* RI primer and a cytosine on the *Mse* I primer (see Figure 2.1). Reactions were carried out in 1× Q solution, 1× PCR buffer, 250 µmol of each dNTP, 10 pmol *Eco* RI+A pre-selective primer, 10 pmol *Mse* I+C pre-selective primer, and 1 U of *Taq* DNA polymerase in a total volume of 20 µl including template. Primer sequences are given in Section 2.1.10.2. The template, 1 µl 1:5 dilution of ligation product, was added to the reactions last. Thermocycling was performed in 20 cycles of 30 seconds at 94°C, 1 minute at 56°C and 1 minute at 72°C. The ramping speed for the program was limited to 1°C/second; after completion of thermocycling, the reactions were held at 4 °C.

Agarose gel electrophoresis of half the reaction volume was routinely used to determine the efficacy of the amplifications. Sample electrophoresis and detection of amplified fragments was as previously described in Section 2.3.2.1. Typically, a smear was observed between 100 bp and 1 kb in size.

2.3.6.4 AFLP Step 4: Selective Amplification

Selective amplifications used primers with a three base overhang, which included the single adenosine or cytosine of the pre-selective primers (see Figure 2.1). Selective amplifications were carried out in a total volume of 20 µl, including template, and contained 1× PCR buffer, 50 mM $MgCl_2$, 250 µmol of each dNTP, 10 pmol *Eco* RI+ANN selective primer, 10 pmol *Mse* I+ANN selective primer, 1 U of *Taq* DNA polymerase and 1 µl of diluted pre-selective amplification product (dilution 1:10). Several selective primer combinations were used, primer sequences are given in Section 2.1.10.2. The thermocycling program consisted of an initial hold at 94°C for 2 minutes, 6 cycles of 30 seconds at 94°C, 30 seconds at 65°C and 1 minute at 72°C, 6 cycles of 30 seconds at 94°C, 30 seconds at 60°C and 1 minute at 72°C and 24 cycles of 30 seconds at 94°C, 30 seconds at 56°C and 1 minute at 72°C. A final extension period of 72°C for 5 minutes was also incorporated. After completion of thermocycling the reactions were held at 4°C.

Agarose gel electrophoresis was used to determine the efficacy of the amplifications prior to acrylamide gel electrophoresis. An 8 µl aliquot of each selective PCR amplification was loaded onto a 2% (w/v) Seakem LE agarose/TAE gel. Sample electrophoresis and the determination of the amplification efficiency was as described previously in Section 2.3.2.1. Typically, a series of discrete bands, of between 100 bp and 1 Kb in size, were observed.

2.3.7 Conversion of AFLP Bands into Sequence Specific DNA Markers

2.3.7.1 Isolation of DNA Fragments from Polyacrylamide Gels

If it was intended to isolate polymorphic bands from the polyacrylamide gel (see Section 2.3.2.2), a single well gap was left between sample loadings to reduce the likelihood of cross-contamination during loading, electrophoresis or later excision of bands.

Polymorphic loci of interest were identified on polyacrylamide gels soon after completion of silver staining (see Section 2.3.2.2.3). These were usually DNA bands between 250 bp and 1 kb in size that occurred in two or more, but not all, of the samples that had been electrophoresed. A representative of each identified band size was excised from the gel using a clean sterile scalpel blade and transferred to a clean labelled 1.6 ml microcentrifuge tube. A 20 µl aliquot of Milli-Q H_2O was then added to each tube. The tubes were stored overnight at 4°C to allow diffusion of the DNA from the polyacrylamide matrix. The DNA template was suitable for reamplification for 2-3 days.

2.3.7.2 Reamplification and Cloning of Excised AFLP Fragments

Reamplification of polymorphic bands isolated from AFLP profiles was regularly carried out using PCR. Each reaction mixture contained 1× Q solution , 1× PCR buffer , 250 µmol of each dNTP, 10 pmol *Mse* Preamp primer, 10 pmol *Eco* Preamp primer, and 1 U of *Taq* DNA polymerase in a final amplification volume of 20 µl. Routinely two parallel reactions were performed using 1 µl aliquot of the AFLP band eluate and a 1:10 dilution as template DNA. Thermocycling consisted of an initial hold at 94°C for 2 minutes followed by 35 cycles of 1 minute at 94°C, 1 minute at 50°C and 2 minutes at 72°C. A final 5 minute extension time at 72°C was used as before. At the completion of the program reactions were held at 4°C.

A 2 µl aliquot of each amplification was electrophoresed on a 1% (w/v) LE agarose/TAE gel as described in Section 2.3.2.1 to determine the efficacy of the PCR.

The amplified products from single reactions were purified using the Concert™ PCR Purification kit as described in Section 2.3.4. After purification, a 2 µl aliquot of each cleaned PCR product was quantified by agarose gel electrophoresis (Section 2.3) using a Low DNA Mass ladder.

The reamplified polymorphic AFLP bands were cloned into a pGEMT® vector as described in Section 2.3.5 and sequenced (see Section 2.3.8). The sequence of the

AFLP derived fragment was screened against the EMBL/GENBANK databases using the BLAST program (website: http://www.ncbi.nlm.nih.gov/blastest/blast.cgi).

2.3.7.3 Primer Design

Primers were designed internally at both the 3' and 5' ends of cloned fragments. Care was taken to match melting temperatures between primers and to avoid strings containing repetitive sequences or consecutive guanine or cytosine nucleotides.

2.3.8 DNA Sequencing

Purified PCR fragments and cloned amplification products purified as described in Section 2.3.4 were sequenced using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit. The technique used by this protocol is based on a method introduced by Sanger and Coulson (1975) which uses polymerisation of DNA under presence of dideoxynucleotide triphosphates (ddGTP, ddATP, ddTTP, or ddCTP). When a ddNTP is incorporated into an extending fragment, the chain is simultaneously broken. The different sized products can be separated by polyacrylamide gel electrophoresis and visualised. In the chemistry used by the ABI PRISM™ each ddNTP is labelled with a different and specific fluorescent dye which can be detected by the ABI PRISM™ DNA sequencer.

Procedure: Automatic sequencing reactions were typically carried out using 25-45 ng of purified PCR product or 100–200 ng plasmid DNA in a total volume of 10 µl. For each reaction a 4 µl aliquot of Terminator Ready Reaction Mix was dispensed to a labelled 0.2 ml reaction tube. To this was added 1.6 pmol of the appropriate primer, an appropriate volume of template and if necessary Milli-Q H_2O to bring the volume up to 10 µl. For automatic sequencing reactions, oligonucleotide primers were always used freshly diluted to a concentration of 1 pmol/µl from the 10 pM/µl working stock. Reactions were then thermocycled in a PTC-200 DNA Engine or PTC 150 Minicycler using a program which consisted of 25 cycles of 10 seconds at 96°C, 5 seconds at 50°C, and 4 minutes at 60°C. Ramping between temperature steps was limited to 1°C/second throughout the program. At the completion of thermocycling the reactions were held at 4°C.

Extension products were subsequently purified by ethanol precipitation. For each reaction 1 µl 3M NaOAc (pH 5.2) was added to a labelled 1.6 ml microcentrifuge tube. The entire 10 µl reaction volume was added, followed by a 25 µl aliquot of ice-cold absolute ethanol. The solutions were mixed by gentle agitation and the tubes incubated on ice for approximately 5 minutes before being centrifuged for 15 minutes at 13 000 × g and 4°C. After centrifugation, the supernatant was carefully removed from each tube using an automatic pipette. A 700 µl aliquot of ice-cold 80% (v/v) ethanol was added to each tube, which was then repeatedly inverted to rinse the pellet. The tubes were then centrifuged as before for 15 minutes. The ethanol was decanted and the DNA pellet was dried at 37°C for approximately 30 minutes. Dried extension products were stored at -20°C until required.

The reactions were resuspended and run by the MUSeq Massey University DNA Analysis Service (Massey University, Palmerston North, New Zealand) on an ABI PRISM™ DNA sequencer.

2.3.9 Collection and Processing of Sequence Data

Typically, both strands of a DNA marker locus were characterised for sequence analyses. Generated data were visualised as chromatograms using the Chromas software program version 1.44 available at [http://www.technelysium.com.au/-chromas.html]. The program also allowed editing of the DNA sequences.

Complementary sequences of one DNA region were aligned to each other using the multiple sequence alignments option of the software program ClustalX version1.8 (Thompson *et al.*, 1997).

Sequences obtained in this way were searched against the database GENBANK to identify similar sequences using the BLAST search engine (Altschul *et al.*, 1997) located at http://www.ncbi.nlm.nih.gov/.

2.4 PHYLOGENETIC ANALYSES OF DNA SEQUENCE

Phylogenetic DNA analysis typically describes the inference of evolutionary relationships between taxa or sequences. The process by which relationships are inferred from sequences can be characterised by several components: the data alignment, any required correction for unobserved changes under a substitution model, the tree selection criteria and the evaluation of the recovered tree or trees (Penny *et al.*, 1992). This section briefly reviews these components and describes analytical methods used in this investigation.

2.4.1 The Data Alignment

Phylogenetic analysis of sequence data usually requires inference of the positional homology between the sequences – the nucleotides at a given position in the DNA sequence should trace their ancestry to a single character that occurred in a common ancestor of those sequences. To create a sequence alignment that satisfies the requirement for positional homology it may be necessary to insert gaps of one or more nucleotides into certain sequences, these gaps corresponding to insertions or deletions (indels) (Morrison and Ellis, 1997). With increased taxon sampling it may become difficult to ensure positional homology between sequences of different length. Choices are made about the number, length and position of introduced gaps.

The computer program ClustalX (Thompson *et al.* 1997) was used for sequence alignment in this thesis. The program is freely available at ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/.

The software uses a Neighbor Joining algorithm to produce a guide tree for determining the order in which DNA sequences are progressively aligned. This procedure does not always recover a correct alignment of indel positions (Hickson *et al.* 2000), hence the alignments were typically inspected visually and corrected manually where necessary. In principle, the alignment order is known to bias the outcome of subsequent evolutionary tree building. This issue has been shown to be particularly important when evolutionary distances between sequences are large and when indels are common (Lake, 1991; Lockhart *et al.*, 1996b).

2.4.2 Pair Wise Distance Analyses

The number of nucleotide sites at which two sequences differ are the basis for investigating the amount of evolutionary change between two sequences (p distances = observed substitutions/total sequence length). The expected accumulation of substitutions between two sequences in a given time is a linear function when $p< 0.75$ for 4 character states and $p < 0.5$ for 2 character states. However, empirical studies show that as mutations accumulate above this level, p-distances can seriously underestimate the actual number of substitutions that have taken place (Page and Holmes, 1998). Convergence, multiple substitutions at the same site or back substitutions result in the observed number of sequence differences being smaller than the actual number of substitutions. In this case the sequences are said to be saturated. Correction formulae can be applied to the p-distances assuming a substitution model, in an attempt to estimate the total number of substitutions between pairs of sequences. Such corrected distances are called 'objective distances'.

2.4.3 Models of Sequence Evolution

To implement algorithms for phylogenetic interpretations, assumptions need to be made on the processes of sequence evolution. Correction formulae approximate natural evolutionary processes. Implemented in Paup* 4.05 and used in the present study are three series of nested substitution models which estimate objective distances.

2.4.3.1 Objective Distance Models

JC - Jukes Cantor Model

The Jukes Cantor model was one of the first models proposed (Jukes and Cantor, 1969) and it is perhaps the simplest model of sequence evolution. It assumes that the four bases have equal frequencies and that all substitutions are equally likely.

K2P - Kimura 2 Parameter Model

The DNA-nucleotides guanine and adenine (A and G) are purines (derivatives of purine); thymine and cytosine (T and C) are pyrimidines (derivatives of pyrimidine). Substitutions that exchange a purine for another purine or a pyrimidine for another pyrimidine are called 'transitions'. Substitutions that exchange a pyrimidine for a purine or vice versa are called 'transversions'. It has been empirically observed that transitions are more frequent than transversions (Wakely, 1996). For this reason, sequences are often saturated faster for transitions than for transversions, which seem to accumulate with a more linear time correlation. Under the assumption that all nucleotides are equally likely Kimura's 2 parameter model incorporates the observation that the rate of transitions per site may differ from the rate of transversions (Kimura, 1980).

F81 - Felsenstein Model

One reason that some sequences might saturate more quickly than others is variation in base composition (Lockhart *et al.*, 1992). Base composition can vary among genes, and in the same gene among different species. If some bases are more common than others some substitutions might be more frequent than others. The model by Felsenstein (1981) addresses this concern to some extent by allowing the frequencies of the four nucleotides to vary. However, the method requires that the frequency of each nucleotide is the same in all sequences being considered.

HKY85 - Hasegawa, Kishino and Yano Model

The HKY85 model (Hasegawa *et al.*, 1985) essentially merges the K3P and F81 model. It allows transitions and transversions to occur at different rates and allows base frequencies to vary.

General Reversible Model (REV)

The general reversible model (Rodriguez *et al.* 1990; Yang *et al.* 1994) allows each possible substitution to have its own probability of change. As such it uses a matrix with 6 parameters.

Additional parameters can be added to all of these models. The most common ones being the assumption of invariable sites (sites that cannot change) in the data and different and different rates of evolution for different sequence positions. The latter are usually modelled by assuming a gamma distribution of rates across sites (Yang *et al.* 1994).

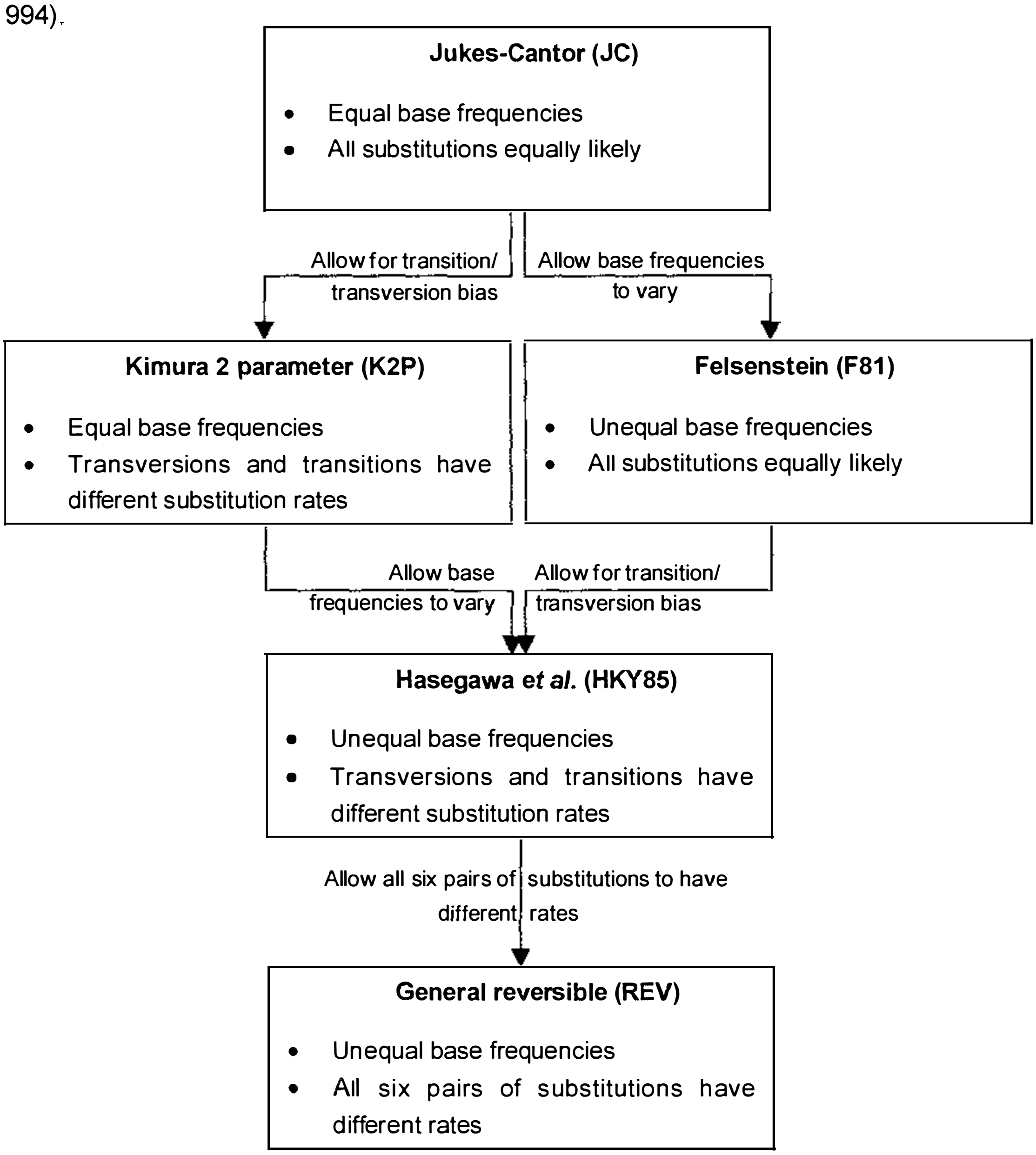


Figure 2.2: Interrelationship among five models for estimating the number of nucleotide substitutions among a pair of sequences. The JC, K2P, F81, HKY85models can all be generated by constraining various parameters of the REV model.

2.4.3.2 Substitution Models and Likelihood Ratio Tests

In the present work, the choice of substitution models was made using a likelihood ratio test (LRT; Page and Holmes, 1998; Pasada and Crandall, 2001). This test compares the ratio of the likelihood scores of one model (i.e. the 'model' includes the tree T1 and substitution model) to the likelihood score of another model (i.e. tree T1 and a nested substitution model). When models are part of a nested series the LRT statistic Δ is χ^2; distributed with degrees of freedom = n-1 (i.e. n = number of parameters differing between the two models). Reference to χ^2 tables is made to determine if the score for the specified tree is significantly better under one substitution model or under another. This approach provides an objective means to choose the simplest sequence substitution model that could explain evolution of the data. Likelihood ratio tests can also be used to evaluate whether data evolves under a molecular clock (see Section 2.4.6). In this latter case the models compared are: (tree T1 + specified substitution model + constrained edge lengths) and (tree T1 + specified substitution model + unconstrained edge lengths).

2.4.3.3 A General Model for Additive Paths

LogDet

The LogDet correction (Lockhart *et al.*, 1994) is a further correction formula designed to circumvent the problem of variable base composition among sequences of an alignment. Unlike the other models described, the LogDet does not require symmetry of substitution processes across an alignment (i.e. uniform base compositions). If sequences have different base compositions tree building methods tend to group these sequences together on the basis on similarity of base composition rather than evolutionary relationship. However, under some conditions, i.e. equal rates across all sites, the LogDet correction will recover a set of additive pathlength between sequences even when base composition is variable (Steel, 1994). For each pair of nucleotide sequences a 4 x 4 matrix is constructed, (a 20 x 20 matrix is constructed in case of amino acid sequences). The entries are the proportion of all possible nucleotide pairs in the 2-taxa sequence alignment. The proportions are calculated from the number of times a nucleotide type in DNA sequence 1 matches each of the nucleotide types in sequence 2 and vice versa. The determinant of each matrix is then calculated and the log is taken. The resulting values are then used as the LogDet distances.

The correction will fail to give additive path lengths when substitution rates vary among sites (Yang, 1996). However, in some instances LogDet has been shown to be effective when implemented as part of an invariable site model to accommodate positional rate heterogeneity (Steel *et al.*, 2000; Waddel *et al.* 2000).

2.4.4 Tree Building Methods

An 'evolutionary tree' ('phylogeny') is a mathematical graph to visualise historical relationships between groups. The 'topology of the graph' (tree shape) is usually calculated by an algorithm. A tree consists of 'nodes' (also called 'vertices') which are connected by 'edges' or 'internodes'. A phylogeny has 'internal edges' which connect 'internal nodes' and 'external edges' which connect an 'internal' and an 'external' or 'terminal node' ('tip'). The sum of edges connecting two taxa is called a path, the sum of mathematical values assigned to these edges being the path length.

Molecular phylogenetics is generally based on DNA or amino acid sequence data. To construct evolutionary trees algorithms may use the number of differences between individual sequences or 'taxa' (discrete methods) or genetic distances (distance methods). The 'genetic distance' is a measure of the amount of evolutionary change between two taxa.

A 'split' or 'partition' is a signal in the data matrix that partitions a set of taxa into two mutually exclusive subsets. However, a split separating a single taxon from the rest is not considered under parsimony criteria (see below) to be 'phylogenetically informative' – since these splits do not provide support under this method for distinguishing between trees.

2.4.4.1 Maximum Likelihood and Maximum Parsimony

Maximum likelihood (ML) and maximum parsimony (MP) are discrete tree building methods implemented typically to evaluate the relative number of site patterns in a multiple alignment of sequences. Both use optimality criteria to assign global scores of fit to possible bifurcating trees. Among all possible trees the one with the best score is chosen. However, with increasing amounts of contradictory patterns within given data it may not be possible to find the optimal tree (or trees).

Maximum likelihood criteria are usually employed to find the greatest probability of obtaining the observed data given a specified substitution model and evolutionary tree. That is, maximum likelihood attempts to identify the tree, or phlyogenetic hypothesis, which is most likely to have given rise to the data (Felsenstein, 1981; Swofford *et al.*, 1996). In calculating the probability that any given unweighted tree evolved the data, additional parameters, such as edge length, ti/tv ratio, etc., are estimated or assumed (Steel and Penny, 2000).

Most implementations of ml require 'stationary' substitution models (see Section 2.4.3) that assume a uniform process across the underlying phylogeny. Sequence positions are expected to evolve identically in process – although perhaps at different rates – and independently (the 'iid' assumption). The log-likelihood score for all possible trees is calculated at every sequence position, and the overall score for any tree is the sum of this value for all positions. The tree (or trees) with the best overall score is chosen as the optimal solution. Maximum likelihood methods perform well in simulations, although, like parsimony and other methods of phylogeny reconstruction they may also become inconsistent with increasing complexity of biological data (e.g. Lockhart *et al.*, 1996a; Steel *et al.*, 2000).

Maximum Parsimony searches for the bifurcating unweighted tree (or trees) that requires the fewest evolutionary changes. Typically, characters are considered as unordered. Although no explicit substitution model is assumed, parsimony tree building can become inconsistent unless the number of substitutions between taxa is very low.

The search strategies employed by ml and MP become difficult as the number of taxa increases. With greater than 10-12 taxa it may be impossible to make the measurements of fit on all possible trees. With up to 20 taxa 'branch and bound' searches can be employed, for larger numbers of sequences heuristic methods are generally used.

2.4.4.2 UPGMA and Neighbor Joining

UPGMA (Unweighted Pair Group Method with Arithmetic Means) is a greedy algorithm used on path length data for clustering taxa into a bifurcating tree. The algorithm describes a series of steps that are repeated until all relationships are represented as bifurcating. Pairwise distance values (d_{ij}-distances) are used to determine the order of (agglomerative) clustering of the least dissimilar sequences or nodes. The algorithm assumes an 'evolutionary clock' so that sister taxa are equidistant from their ancestral node. On the basis of this assumption the method averages distances between sequences and clusters. Hence, the true branch lengths from the root of the tree to each tip are the same: the expected amount of evolution in any lineage is proportional to elapsed time.

The Neighbor Joining method by Saitou and Nei (1987) also uses measures of path length to identify which sequences should be clustered as closest neighbours in a reconstructed graph. However, rather than directly interpreting d_{ij}-distances (as done in UPGMA), the decisions of which sequences and nodes to cluster are based on values in a modified distance matrix, the entries of which are called S_{ij} values:

$$S_{ij} = (N-2)\, d_{ij} - R_i - R_j$$

with

N = the number of sequences in the d_{ij} matrix
R_i = the sum of pairwise distances of sequence i from other sequences in the d_{ij} matrix
R_j = the sum of pairwise distances of sequence i from other sequences in the d_{ij} matrix

Neighbor Joining and S_{ij} values are particularly useful when the relative edge lengths in the true underlying tree are irregular in their lengths. Branch lengths are estimated as the nodes and sequences are joined. Scores are not evaluated for rearrangements in the tree. As the method does not assume an evolutionary clock, the algorithm produces an unrooted tree.

2.4.4.3 Quartet Puzzling

Quartet based tree building methods analyse sequences into subsets of four (quartets). For each quartet three possible unrooted bifurcating trees (or unresolved trees), are identified that best describe the data. Quartet Puzzling (QP) uses criteria such as maximum likelihood or parsimony to evaluate the highest quartet tree score. In a successive step the method assembles the quartet trees into a final tree containing all taxa. Under the Paup* default setting for QP a global heuristic assemblage of quartets is made 1 000 times. A consensus of these trees is then calculated. Given perfect data, the assembly step is trivial, as there will be only one tree that all subtrees agree on. However, if there are many contradictory patterns in the sequence data it may be that not all quartet trees can be combined into one unique tree. In this case the QP-tree will try to accommodate the greatest possible number of subtrees. The incompatibilities in the data will be reflected in low QP reliability values, which are assigned to each tree edge.

2.4.4.4 Spectral Analysis and Split Decomposition

Spectral analysis (Penny *et al.*, 1992) and split decomposition (Bandelt and Dress, 1992) use similar but different criteria for evolving the support for possible splits (edges) in a data set.

While split decomposition displays as a 'splitsgraph' the most strongly supported splits in the data, spectral analysis calculates the support for all splits possible between n taxa. However, only the support for splits of interest may be represented (Steel *et al.*, 2000). In its simplest form, spectral analysis plots the frequency of each split in a data set. External splits are trivial splits as they have to occur in any tree. Non-trivial splits are the internal splits. A phylogenetic tree can be built from a spectrum by finding the biggest set of mutually compatible splits.

In calculating the support for splits, spectral analysis uses a Hadamard matrix to add and subtract path length between pairs and higher order subsets of taxa. In contrast, split decomposition estimates support for splits using only quartets of taxa.

Split decomposition first evaluates support for internal splits. To do this it examines bifurcating trees from all combinations of four taxa in a data set. For every quartet, three bifurcating trees are possible. Support is evaluated for the internal splits of these bifurcating trees. The value of support is called the isolation index score. Parsimony, distance, maximum likelihood or other criteria can be used in deciding which bifurcating trees are most strongly supported by the data.

The two splits which have the highest isolation index score for each combination of four taxa are said to be weakly compatible and they are used to next obtain a split system that describes the relationship between all taxa in the original data set.

Since some of the quartets that identify the same split between more than four taxa are not excluded in the first step of the quartet study, a decision needs to be made as to which isolation index value will be used. From all the quartet splits that are compatible with a split in the split system split decomposition (in its standard implementation) chooses the *smallest* isolation index score to be represented in the split system.

The split system also must contain the splits, which correspond to the external edges or branches in the final tree or network. Support for these is calculated using all combinations of three taxa. Split decomposition also chooses the smallest value for that split.

Once the splits that will be present in a split system have been identified a graph is constructed. If a split system contains no contradictory patterns a treelike structure will result. Incompatible splits will produce alternative internal connections between two nodes. The resulting graph will visualise contradictions in the sequence data as a network. A score called the 'isolation index' indicates how well the splitsgraph represents pairwise distances calculated from the sequences.

The fit of sequence data to a splitsgraph is described by a fit statistic (also see Section 2.4.5.3) which is the sum of path lengths in the splitsgraph divided by sum of path lengths between all pairs of taxa calculated independently of any graph. In the program SplitsTree, path lengths used to obtain the length of splits ('d-splits') can be calculated using either (i) 'p', also called hamming distances (i.e. observed number of differences over the total number of positions), or (ii) Kimura 3 parameter model (K3St) distances or (iii) Jukes Cantor model (JC) distances or (iv) LogDet distances.

2.4.5 Measures for the Evaluation of Phylogenetic Trees

In evolutionary tree building there are no tests for consistency, there are only tests for 'convergence' (Penny *et al.*, 1992). Such tests indicate whether the topology of the reconstructed tree, made using a certain method, is expected to be stable, given more data of the type already analysed. Tests used to evaluate the fit of data onto evolutionary trees have been discussed by Swofford *et al.*, (1996) and by Page and Holmes (1998).

2.4.5.1 Bootstrap and Jacknife Analyses

Statistical resampling techniques, such as non parametric 'bootstrapping' and 'jacknifing', can be implemented to estimate the stability of an evolutionary tree by resampling the original data set (Swofford *et al.*, 1996). Jackknifing tests the stability of a certain phylogeny given less information. The method can be employed either by randomly sampling proportions of the sequence information or by jackknifing the number of taxa.

Bootstrapping, in contrast, uses the sequence data and can be defined as 'sampling with replacement'. For evaluating phylogenetic trees, the algorithm randomly resamples the original data set, with replacement, to create multiple new data sets with the same number of characters as the original. A tree is produced from each of the replicate data sets, of which the properties are recorded and used to construct a consensus tree. Values are assigned to the internal edges of the consensus tree that indicate the proportion of times a given edge was recovered from tree building on the replicate data sets. Bootstrapping does not establish that a given phylogeny is correct, it rather provides a measure of the relative support in the data for a tree given a specified tree building procedure. High bootstrap values indicate that relatively large numbers of uncontradicted patterns support the arrangement. Furthermore, they suggest that the addition of further data, with a similar structure, should not markedly alter the choice of optimal tree shape for the given tree building method.

There are two possible explanations for low bootstrap values: 1. There are relatively few informative - possibly uncontradicted - patterns which support the relationship. 2. Similar numbers of conflicting pattern types are present in the data (Efron, 1982; Swofford *et al.*, 1996).

2.4.5.2 Quartet Puzzling Values

With quartet puzzling 'reliability' values (Strimmer and von Häseler, 1996) are given for the internal branches in the reconstructed tree. QP values indicate the proportion of times during the puzzling step the edge (split) was found in the tree built from quartets. Although they can sometimes be similar to bootstrap values (see Section 2.4.5.1), the quartet puzzle scores for edges in trees have an advantage over bootstrapping (2.4.5.1) when the data shows relatively few substitutions. An edge will receive a high puzzle value – suggesting strong support for that edge – when it is compatible with the remaining data, even if there is only one nucleotide position which suggests the relationship. Bootstrapping would give a low value in this example (Lockhart *et al.*, 2001).

2.4.5.3 Global Fit Statistics on Graphs and Trees

Sequence spectra, observed and objective distances, calculated from pairwise comparisons, rarely perfectly fit reconstructed trees. This can occur because the true number of evolutionary changes that lead to the observed data is difficult to estimate. In practice, the true amount of genetic change between taxa may be greater or less than that represented on the tree. Measures of fit can be used to both select and compare trees. 'Global' criteria, such as implemented under parsimony or maximum likelihood methods, identify the best fit of observed patterns in the data to a tree – i.e. the so called optimal tree. Maxmimum likelihood methods for comparing the relative fit of sequence patterns to different trees used in this thesis have been described in section 2.4.5.4

Distance measures can be also used to select and compare the relative fit of sequence data to particular reconstructed trees i.e. the optimal tree will be that which minimises the difference between the distances in the distance matrix and the distances represented in the tree. Different trees can also be compared in terms of how well they minimise this difference. In the case of split decomposition, because it represents not just one bifurcating tree, the fit statistic associated with each splitsgraph gives an indication of whether or not the strongest incompatible signals in the data can be represented in a splitsgraph – thus even if the data are reticulate the fit statistic can still be 100% so long as the extent of reticulation can still be visualised in the splitsgraph. The measure used is the sum of all path length through the graph divided by the sum of pairwise distances in the d_{ij} matrix.

Other commonly used measures of fit in tree building include 'CI, RI, and HI' indices (Swofford *et al.*, 1996) but have not been used in this study.

2.4.5.4 Comparing Evolutionary Trees Using Global Scores

Evolutionary trees can be compared using tests which examine the difference between global scores (e.g. log likelihood values) for the trees (hypotheses). If the difference in scores is significantly greater than that expected by chance - given the stochastic variation in the data, then the null hypothesis can be rejected that the score of one tree on the data is significantly different from the score of another tree on the same data.

2.4.5.5 Kishino Hasegawa (KH) Test

This test is typically used when the two trees to be compared are specified without reference to the data (Goldman, *et al.*, 2000). In principle, a distribution can be calculated for the score difference between the two trees. This can be done analytically, or by bootstrapping the data and calculating the difference in the tree scores for each bootstrap sample. These score differences can then be used to plot a distribution of scores centred around zero. Such a centred distribution represents the null hypothesis that the difference in the scores between the two trees is zero. If the score difference obtained from comparing the trees on the original data set is outside the centred distribution of scores e.g. obtained from comparing the trees on the bootstrapped data, then it assumed that the two trees are significantly different. An analytical version of this test is implemented in PAUP*.

2.4.5.6 Shimodaira Hasegawa (SH) Test

This is also a non parametric test for comparing two trees and can be used when the two trees compared are not directly specified with reference to the data – the only proviso is that the two trees must be within a set of trees (set 'M') that are possible explanations of the data – and these trees in set M should be chosen without reference to the data (Goldman *et al.*, 2000). The null hypothesis for this test is that all trees in set M are equally good explanations of the data - once stochastic variation is taken into account (either by bootstrapping or calculation). That is, the score differences between any two trees in set M will be zero. In practise, the test becomes a more conservative version of the KH test and the p values obtained with the SH test are approximately greater or equal to p/2 values obtained for the KH test. An analytical version of this test is also implemented in PAUP*.

2.4.6 Molecular Clock Analysis

The suggestion that genes evolve at some constant rate is called a 'molecular clock assumption'. This assumption allows estimation of divergence times by comparing the relative divergence of molecular sequences. Although a generalised molecular clock is unlikely, the existence of 'local' molecular clocks has been advocated. That is, between closely related taxa with similar life histories and molecular processes, a consistent rate may be expected at an individual gene locus. (Hillis *et al.*, 1991; Page & Holmes, 1998).

As mentioned briefly previously, one method available to test for the applicability of a molecular clock is the 'likelihood ratio test'. In this approach, maximum likelihood trees are inferred for a data set both with and without the assumption of a molecular clock. The overall likelihood scores for the two trees are used to derive a measure of similarity, which is distributed as χ^2 with (n-2) degrees of freedom, where n is the number of sequences (Hillis *et al.*, 1996; Page and Holmes, 1998). Although the test is sensitive, it may be subject to violations of the evolutionary model specified (i.e. both the tree and the assumed model of sequence substitution).

A more general test is the relative rates test described by Steel *et al.*, (1996). This examines triplets of sequences and tests whether the distance between sequences on either side of an ancestral node are statistically equivalent (see Figure 2.3). If they are, a 'molecular clock' is inferred.

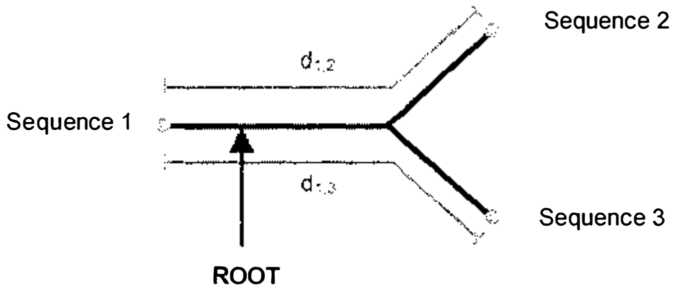


Figure 2.3: Path length comparison for the relative rates test of Steel *et al.*, (1996). This procedure tests whether distances across the root ($d_{1,2}$ and $d_{1,3}$) are statistically equivalent, an assumption that is required for the existence of clock-like change between the three sequences.

2.5 APPROACHES IN POPULATION GENETICS

2.5.1 Alleles and Allele Frequencies

Approaches in population genetics attempt to describe the structure of genetic variation within and between subgroups of species i.e. subspecies, variants, or populations. Traditionally, measurements have concentrated on evaluating the relative frequencies of different forms of genes ('alleles').

Allelic variation effects the genotype of an individual and may also cause phenotypic differentiation. The frequencies of different alleles in a population can be measured directly and are often shown as percent values. If changes in allele frequency occur as a response to environmental forces the process is called 'natural selection'. 'Genetic drift', in contrary, is a result of random variation (Sewall Wright effect; Page and Holmes, 1998). Population genetic measures have been used in the present work to investigate intraspecific variation in *Nothofagus.*

In a diploid organism, two alleles are present for each gene situated on either homologous chromosome. An individual is said to be homozygous at a particular gene locus, if both homologous alleles are identical. The individual is said to be 'heterozygous' if both homologous alleles differ from each other. In some cases, one allele may be expressed preferentially over the other, in which case it is the 'dominant allele', while the other one is the 'recessive allele'. 'Codominant alleles' are simultaneously expressed in an organism.

2.5.1.1 The Hardy-Weinberg Theorem

The Hardy-Weinberg theorem is named after the co-discoverers G.H. Hardy and W. Weinberg and describes the balance in the relative number of alleles that is maintained within a large population over a period of time. The theorem requires a stable population which fulfils the following four assumptions:

1. mating is random
2. there is no natural selection
3. there is no migration
4. there is no mutation

For a population with two different alleles *A* (dominant) and *a* (recessive), three genotypes are possible: *AA* (homozygous dominant), *Aa*, (heterozygous), and *aa* (homozygous recessive). The frequency of the three genotypes can be expressed by the Hardy-Weinberg equation: $p^2 + 2pq + q^2 = 1$ with:

p^2 = frequency of *AA* individuals
$2pq$ = frequency of *Aa* individuals
q^2 = frequency of *aa* individuals

The equation can also be applied when there are more than two alleles at one locus. Deviations from the expected frequencies are important for studying the evolution of genes in populations and effects of natural selection.

2.5.2 Wright's Statistics and Gene Flow

2.5.2.1 The Fixation Index

In 1951, Wright defined what has come to be called the fixation index, to quantify the inbreeding effect within populations. This index equals the reduction in heterozygosity which is expected under the assumption of random mating between all individuals and the Hardy-Weinberg theorem, respectively. If the assumption of random mating between populations is fulfilled, the fixation index will approximate zero.

The measurements may include all populations of one species or a subsets of 'nested hierarchical levels', which generally reflect different local populations within one larger geographical area. Wright distinguished between subpopulations (S) (or groups of individuals (I)) regional aggregates (R) and the total population (T). Depending on the hierarchical levels included in the statistics, the fixation index is called F_{SR}, F_{RT} or F_{ST} (or F_{IT}) and the proportion of heterozygotes (H) of the lower hierarchical level is compared to the higher level.

In this study the fixation index is referred to as F_{ST} and the formulae is:

$$F_{ST} = \frac{(H_T - H_S)}{H_T}$$

with: H_T = observed heterozygosity of the total population

H_S = observed heterozygosity of the subpopulation

2.5.2.1 Genetic Exchange among Populations

Genetic exchange (gene flow) is the transfer of genetic material between populations of one species of between individuals within populations. Gene flow between populations results from movements of individuals or their gametes. Usually, gene flow is expressed as a migration rate 'm' defined as the proportion of alleles in a population that is of migrant origin.

Wright showed that for neutral alleles i.e. without selection, equilibrium expectations are:

$$F_{ST} = \frac{1}{(1 + 4Nm)} \qquad \text{or} \qquad Nm = \frac{(1 - F_{ST})}{4\, F_{ST}}$$

Values calculated under this equilibrium are irrespective of population size. The average exchange of one individual per generation and between populations corresponds to a migration rate of Nm = 1 and a Fixation index of F_{st} = 0.20. In theory the exchange of one individual per generation is sufficient to prevent genetic drift.

CHAPTER 3

Origins and Evolution of the New Zealand Forest Flora

3.1 INTRODUCTION

The origins and evolution of New Zealand's flora have been of international interest since the earliest botanical explorations of the Southern Hemisphere. Hypotheses have ranged from those suggesting that the present New Zealand flora evolved in situ with direct and unbroken ancestry from Gondwana, to those suggesting that the entire flora of New Zealand arrived by transoceanic dispersal.

In his work on the New Zealand flora, Hook. (1853) dismissed the possibility of long-distance dispersal. He was more impressed by the idea that the present plant cover represented the remnants of migration over larger landmasses. To explain transoceanic distributions he speculated on the existence of ancient land bridges that connected islands and continents. For decades, by many scientists this hypothesis was favoured over those invoking other explanations. Even as recently as the 1960s, van Steenis (1962) explained botanical affiliations of the modern flora by an hypothesis suggesting the former existence of land bridges between now isolated landmasses. He found it particularly intriguing that the 'latitudinal zonation' of floristic elements includes all climate zones and 'concerns almost all families of both Gymnosperms and Angiosperms'. However, no geological evidence has been brought forward to support these ideas.

With the introduction of the plate tectonic theory (Wegener, 1912), many scientists became sceptical of land bridge hypotheses. However, a real change of paradigm did not come until 1972, when Raven and Axelrod reinterpreted Australasian biogeography almost entirely as a result of plate tectonics. They stated that biotic relationships between extant landmasses reflected the common origin and the break-up of the Australasian region. They regarded that much of the present lowland flora of New Zealand resembled that of temperate Gondwanaland 80 million years ago. Differences in the floristic makeup of the Australasian region were suggested to be at least partly due to a different climate between New Zealand and Australia during the Cretaceous.

When geological studies excluded the possible existence of recent continental land connections for a number of islands including 'continental islands' like New Zealand as well as volcanic landmasses like Hawaii (Fleming, 1979; Seibold and Berger, 2000), more scientists began to think that trans-oceanic dispersal of plants and animals had to be reconsidered. A strong advocate for dispersed origins of New Zealand's plants and animals was the palaeontologist and biogeographer Fleming (1949, 1979). His early arguments centred around the absence of land dinosaurs, early mammals and snakes. Later in his career it was discovered that New Zealand did have a land dinosaur record (Molnar, 1980). However, Fleming did not alter his views on dispersal. The absence of geological evidence for land bridge hypotheses as well as the palynological available at that day led him to argue strongly that the arrival of many species in New Zealand must have occurred long after the break-up of Gondwanan.

In New Zealand, analyses of plant micro- and macrofossils have provided evidence against direct Gondwanan ancestry of the flora (Mildenhall, 1980; Pole, 1994; Macphail, 1997b). Some plant lineages first appear in the New Zealand fossil record long after the separation of the southern hemisphere continents. Further, having observed the poor match between pre-Pleistocene fossil pollen and modern pollen types, Pole (1994) proposed that Gondwanan plant cover may not have survived the "Oligocene drowning" of New Zealand's landmasses, and that all extant lineages represent late Tertiary and Quaternary immigrants.

Palynological studies have identified many plant groups that have their first appearance in the New Zealand fossil record subsequent to the Oligocene (Pole, 1994; Macphail, 1997b). Many recent molecular systematic studies verify this observation and show that during the Pliocene/Pleistocene trans-Tasman dispersal occurred between Australia and New Zealand (Winkworth *et al*, 1999 and references therein). Interestingly, most of these molecular studies, which also identify the Quaternary as an era of extensive radiation in many plant groups, have concentrated on the alpine flora. They provide evidence for dispersal and radiation into habitats that are themselves very young (see Sections 1.5.3 and 1.5.4). The findings are consistent with those from similar studies on alpine flora from northern hemisphere mountainous regions (Comes and Kadereit, 1998) and also from oceanic islands (Baldwin and Sanderson, 1998).

3.1.1 Aims and Expectations

As yet few molecular studies have been undertaken on plant groups that occupy habitats that have had the potential to exist in New Zealand for much longer periods of geological time. The aim of these studies was to test the possibility of direct Gondwanan ancestry for New Zealand species of *Agathis* (Araucariaceae), *Dacrycarpus*, *Podocarpus* and *Dacrydium* (all Podocarpaceae), *Nothofagus* (Nothofagaceae) and *Pseudowintera* (Winteraceae) using molecular phylogenetic analyses of DNA sequence data. Furthermore, the hypothesis of a continuous presence in New Zealand since their first appearance in the Early-Mid Tertiary fossil record was to be tested for the three plant groups *Metrosideros* (Myrtaceae), *Knightia* (Proteaceae), and *Myrsine* (Myrsinaceae).

The species chosen for this study are found in habitats that may have existed in New Zealand since the Cretaceous. It was expected from fossil evidence that some species, originating from distant southern hemisphere lands, were separated by significant genetic divergences and that at least some of these plant lineages have had a long, if not continuous, presence in New Zealand since the break-up of Gondwana (Mildenhall, 1980; Macphail, 1997b).

3.1.2 Data and Analyses

For the phylogenetic analyses presented in this chapter DNA sequence data were determined for species of *Agathis* (4 species), *Myrsine* (10 species), *Elingamita* (1 species) and *Rapanea* (3 species), *Nothofagus* (8 species) and *Metrosideros* (1 species). Herbarium vouchers for these species are listed in Appendix II. DNA amplification and sequencing were performed as described in Sections 2.3.3 and 2.3.8 of this thesis.

The sequence data were complemented by data already available from the NCBI GenBank (http://www.ncbi.nlm.nih.gov/) which allowed phylogenetic hypotheses and in particular species divergence orders and divergence times to be estimated.

Sequence data were aligned using the progressive multiple alignment procedure, ClustalX (Thompson *et al.*, 1997). These alignments were inspected visually and checked for possible misalignment. For use in phylogenetic programs, sequence alignments were transformed into nexus format. All nexus files used are reported in Appendix V and are provided on the accompanying CD. In phylogenetic reconstructions, and tests with sequence data, ambiguous/heteroplasmic characters were excluded from the data matrix.

Split decomposition (Huson, 1998, Lockhart *et al.*, 2001) was used to investigate the tree-like properties of the sequence data using observed distance estimates and the software program SplitsTree 3.1 (Huson, 1998).

To identify the model of sequence evolution that best describes the properties of a particular set of DNA sequences for ingroup taxa, model-fitting analyses were employed under maximum likelihood criteria. To do this, parsimony search strategies with TBR branch swapping were employed to identify one or several equally optimal unweighted trees. The parameters of sequence evolution were then described for the given tree(s) under nested models within the GTR, HKY and Fel84 series (Swofford *et al.*, 1996) and evaluated in likelihood ratio tests (Yang *et al.*, 1994). The parameter settings for three variables (base composition, proportion of invariable sites and transition/transversion ratio) were estimated in model fitting analyses.

Molecular clock tests were made using the relative rates method described by Steel *et al.*, (1996) and were carried out to further investigate substitution properties of a particular genome region. The test was made using the 'test_clock' option under SplitsTree 3.1 (Huson, 1998).

Estimates of divergence between two groups of taxa used the method described by Steel *et al.* (1996) which is implemented under the 'diverge option' of SplitsTree 3.1. In calculating a Jukes Cantor distance (Jukes and Cantor, 1969) between groups using this method, invariable sites were first estimated using the capture-recapture method of Steel *et al.* (2000) in SplitsTree 3.1 and excluded from the data set.

3.2 ARAUCARIACEAE HENKEL & W. HOCHST. 1865

3.2.1 Introduction

Araucariaceae are an ancient gymnosperm family in the order of Pinales. Some authors elevate the family to the level of an order (Araucariales). Today, the family has a primarily southern hemisphere distribution with species in South America, Australia, New Zealand, New Guinea, New Caledonia and other south Pacific islands.

About 35 species are known from three well-defined genera: *Araucaria* de Jussieu with 19 extant species (de Laubenfels, 1988), *Agathis* Salisbury with 13 extant species (Whitmore, 1980) and the monotypic genus *Wollemia* Jones *et al.* (Wollemi Pine). The Wollemi Pine was only discovered in August 1994, in the Wollemi National Park in New South Wales, Australia (Jones, 1995). The new genus has many morphological features in common with Cretaceous and early Tertiary fossil groups such as Araucarioides, and may be closer to these. Its pollen also rather resembles fossil examples than pollen of living Araucariaceae.

All species are evergreen trees with spirally arranged, narrow or broad leaves, which often have parallel veins. Plants are dioecious or monoecious. Male cones are relatively large, cylindrical, with numerous sporophylls and with ±12 inverted pollen sacs; pollen grains are wingless. Female cones are usually borne erect with one-seeded scales. While the number of cotyledons in seedlings may vary within the genus Araucaria (de Laubenfels, 1988), Wollemia and all species of *Agathis* have two cotyledons (Whitmore, 1980; Jones *et al.*, 1995).

Fossils of *Araucaria* are among the oldest fossils of extant coniferous genera and are found in strata dating from the Triassic or Jurassic eras (Miller, 1977, 1988; Stockey, 1982). The genus *Araucaria* is represented by good fossil material in both hemispheres as early as the Jurassic (Stockey, 1994). Cuticle studies of extant Araucariaceae have enabled accurate comparisons of fossil leaves to living taxa. *Araucaria* Section *Bunya* is represented by cones of several types in the Jurassic. In addition to these remains, a suite of araucarian cones showing affinities to several sections of the genus *Araucaria* have been described from England, Japan and North America (Stockey *et al.*, 1994).

Agathis is known only from the Southern Hemisphere. Fossils from Tertiary deposits are known from the Australasian region, which exceed the present range of the genus by 10 degrees to the south and include pollen from western Antarctica near the tip of South America (Whitmore, 1980). The oldest known fossils of *Agathis* are a collection of leaves from the Lower Cretaceous excavated in southern Victoria, Australia (Cantrill, 1992).

3.2.1.1 New Zealand Araucariaceae:

Agathis australis (D.Don) Lindl or 'kauri' is the only living member of the Araucariaceae in New Zealand. It is the only member of the genus that occurs outside the tropical belt. With up to 60 m height and 7m diameter it is probably the most well known member of the New Zealand forest. The oldest living tree has been estimated to be 2100 years old.

The thick leathery leaves with their parallel veins are sessile and arise on the branchlets either alternately or opposite to one another with conspicuous internodes. Leaves of young trees are lanceolate 5-10 cm long and 5-12 mm wide; adult leaves are more blunt and only 2-3.5 cm long. The species is monoecious and begins to bear fertile seeds with twenty-five to thirty years of age. The female cone releases ovate winged seeds, which are free from the cone scale and wind borne.

In New Zealand, *Agathis australis* is confined to areas north of 38° southern latitude (Ecroyd, 1982). When European settlers arrived on the islands, kauri covered much of the area of North Auckland and the Coromandel Peninsula with dense forests from sea level to 600 m (Salmon, 1996). During much of the last century the huge trees were heavily milled, they formed the basis of timber-logging industries in northern New Zealand. Before human times *Agathis* seems to have been more widely spread on the islands. Pollen of *Agathis* cf. *australis* is found in lacustrine sediments near Wellington that are between 340 000 and 520 000 years of age (Mildenhall, 1984).

Araucarian fossils have been described from ancient New Zealand sediments. *Araucaria* has been recorded from an Upper Cretaceous stratum in Southland (Bose, 1975 and references therein). As early as 1887 Ettinghausen described three species of *Dammara* spp. from upper Cretaceous leaf material, which were later assigned to genus *Agathis*. Pole (1994), however, was convinced that these fossil leaves reported by Ettinghausen were misinterpreted and illustrated up-side down. He surmised the fossils were 'in fact the sharp apices of *Araucaria* leafs' and stated that there was 'presently no fossil record of *Agathis* in New Zealand'. His view did not consider studies by Daniel (1989) who found fossil leaf material embedded in Mid-Cretaceous sediments of the Clarence valley in the New Zealand South Island, which he could clearly identify as remains of the genus *Agathis.* Daniel compared cuticular micro-morphological characters of the fossil remains with those of extant taxa. He found that the fossil remains were most similar to but distinct from *Agathis australis* and described the new species *as Agathis seymouricum*. He stated that 'since mid Tertiary times *Agathis australis* is the only remaining descendant chronospecies of a diverse assemblage of araucariaceaens' and concluded that *A. australis* has, most probably, directly evolved from *A. seymouricum*. In contrast to this hypothesis, Pole (1994) raised doubts whether this (or any other) genus has had a continuous presence on the archipelago. Rather he proposed that extant genera that appear in the New Zealand Cretaceous fossil record were re-dispersed to New Zealand after the 'Oligocene drowning' 30 million years ago.

3.2.2 Phylogenetic Analyses of Araucariaceae *rbc*L - Sequence

In order to elucidate the phylogenetic relationship between New Zealand and overseas Araucariaceae, an alignment of *rbc*L sequence data was analysed according to procedures described in Section 3.1.2. Analyses in this section were also performed to test the hypothesis, proposed by Pole (1994), that *A. australis* has not survived the 'Oligocene drowning' of the New Zealand landmass.

For this purpose, *rbc*L sequences were determined for *Agathis australis* (characterised in three accessions) and for the three Australian species *A. atropurpurea* Hyland (two accessions), *A. microstachya* Bailey & White (two accessions), and *A. robusta* Bailey (one accession). Additional Araucariaceae *rbc*L sequences were taken from the study by Setoguchi *et al.* (1998). All sequences available at the NCBI GenBank for genus *Agathis* and a representative number of taxa for genus *Araucaria* were included in the present study. *Taxus baccata* L. *x cuspidata* Siebold and Zucc., *Sequoia sempervirens* (D.Don.) Endl. and *Juniperus conferta* Parl. were used for outgroup placement. *Rbc*L sequences for these taxa were taken from Brunsfeld *et al.* (1994) and Gadek and Quinn (1993), respectively.

3.2.2.1 Split Decomposition

The suitability of the *rbc*L sequences for phylogenetic analysis of Araucariaceae species and the positioning of the New Zealand species *Agathis australis* was studied with SplitsTree 3.1 (Huson, 1998). The splitsgraph, based on observed distances from all three *rbc*L codon positions (1257 homologous sites), is shown in Figure 3.1. The graph indicates treelike (largely compatible) patterns in the sequences and their suitability for phylogenetic analysis. The data displays non-treelike properties only between some closely related species within the genus *Agathis*. These incompatibilities are visualised as a box-like structure. Tree building methods under maximum likelihood criteria (Figure 3.2) recover a similar phylogeny as with split decomposition.

Under split decomposition the information in the sequence data converges to a tree shape that separates the three genera *Agathis*, *Araucaria* and *Wollemia* by two strongly supported internal edges (bootstrap values ranged between 98 and 100). Figure 3.1 also suggests that the Wollemi Pine is genetically equidistant from both *Agathis* and *Araucaria*. This finding corroborates the observation that this species displays morphological characters which are intermediate between both other extant araucariaceous genera. The edge separating the Wollemi Pine from other taxa is rather short. Its position and placement in the *Araucariaceae* phylogeny suggests a relatively slow rate of *rbc*L-sequence evolution in the lineage leading to this species.

Agathis australis is distinct from other species of *Agathis* and placed on a unique edge in the unrooted graph. Its resolution from other species of *Agathis* as well as from *Wollemia nobilis* gains high bootstrap support of 99 and 100, respectively. Similar to the Wollemi Pine, the New Zealand kauri is placed on a short external edge that further suggests a non-clock-like DNA substitution rate within the *rbc*L gene of this species.

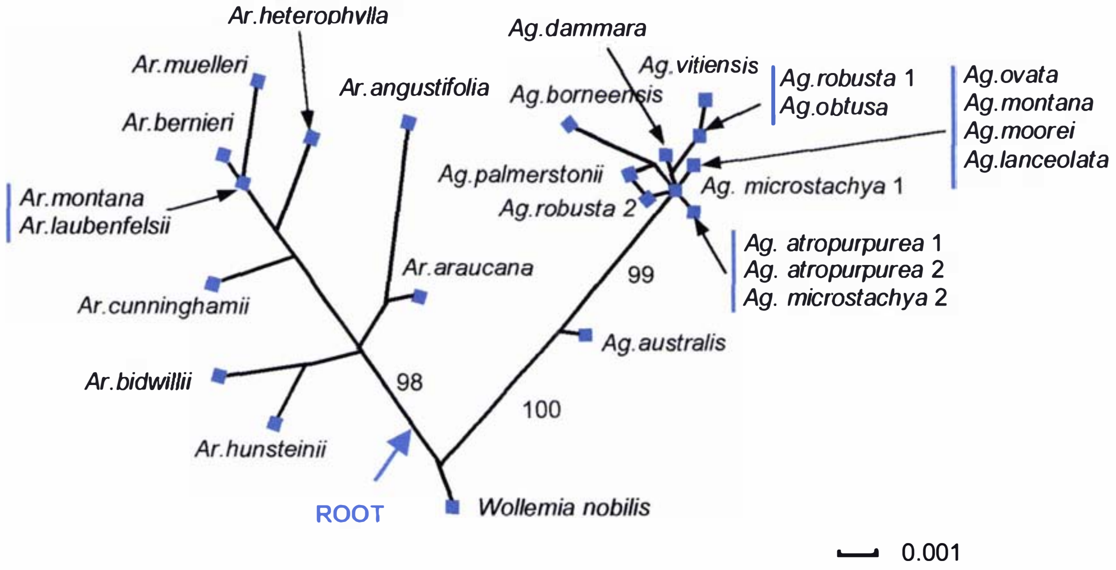


Figure. 3.1: Splitsgraph for Araucariaceae made using observed distances calculated from all three *rbc*L codon positions (1257 homologous sites) under SplitsTree (Huson, 1998), made without outgroup sequences. The position of the root as determined by outgroup placement with *Taxus*, *Sequoia* and *Juniperus* is indicated by a blue arrow. Bootstrap values taken from 100 replicates are shown on the figure. Scale: 0.001 substitutions per site. Details: Fit: 71.3, characters = 1257; missing 14 (off); constant = 1184; non-parsimony = 1209; dsplits-hamming. For splitsgraph details see section 2.4.4.4. Abbreviations: **Ag**: *Agathis*; **Ar**: *Araucaria*.

Inference on the position of the root was made from the placements with *Taxus baccata, Sequoia sempervirens* and *Juniperus conferta.* Split decomposition places the root (100% bootstrap support) basal to *Wollemia* and basal to all species of *Araucaria.* However, the root placement causes the internal edge separating *Wollemia* from the genus *Araucaria* to collapse. A splitsgraph, which includes the chosen, outgroup taxa fails to visualise the distance between these internal nodes.

3.2.2.2 Maximum Likelihood Analyses

The phylogenetic tree shown in Figure 3.2 was obtained with quartet puzzling under maximum likelihood criteria. To obtain the ML parameters to be used for tree building, a heuristic parsimony search was employed to find seven equally parsimonious unweighted trees of all ingroup taxa (*Araucaria, Agathis* and *Wollemia*). Evaluation under Shimodaira - Hasegawa sites test (Goldman *et al.*, 2000) showed that three out of the seven MP trees were equally likely. ML parameters estimated on one of these trees were chosen for quartet puzzling analyses.

Model fitting analyses identified the GTR model and parameter settings to provide the best explanation for the evolution of the Araucariaceae *rbc*L sequence data. GTR parameter settings, which allowed a proportion of sites to be invariable (GTR +I model), were used to obtain the phylogeny shown in Figure 3.2. In the phylogeny, all three genera *Araucaria*, *Agathis,* and *Wollemia* are monophyletic. The ML tree describes a closer relationship between *Agathis,* and *Wollemia.* The root position, as determined by outgroup placement with *Taxus baccata, Sequoia sempervirens* and *Juniperus conferta* was locally stable and joined between *Araucaria* and *Wollemia.*

Within the genus *Agathis*, *A. australis* is the earliest diverged lineage and it is well separated from all other members of the genus. Analyses using a LogDet/Invariable sites model (Steel *et al.*, 2000) under Neighbor Joining (Swofford, 1999) also supported the conclusion that the position of *A. australis* in the phylogeny is highly stable.

All *Agathis* species from Australia, the Malaysian region, Fiji, New Caledonia and Vanuatu form a monophyletic group. The low genetic distance between all members of this group suggests recent radiation and dispersal between Australia, the Malaysian region and some Pacific Islands. Within this phylogenetic cluster, maximum likelihood recovers three subgroups. One group comprises all New Caledonian species, the two other groups include species from different localities, suggesting more complex evolutionary origins. The two accessions included for *Agathis robusta* appear in two different groupings. No tree building method applied recovered a tree showing adjacent relationships between these two accessions.

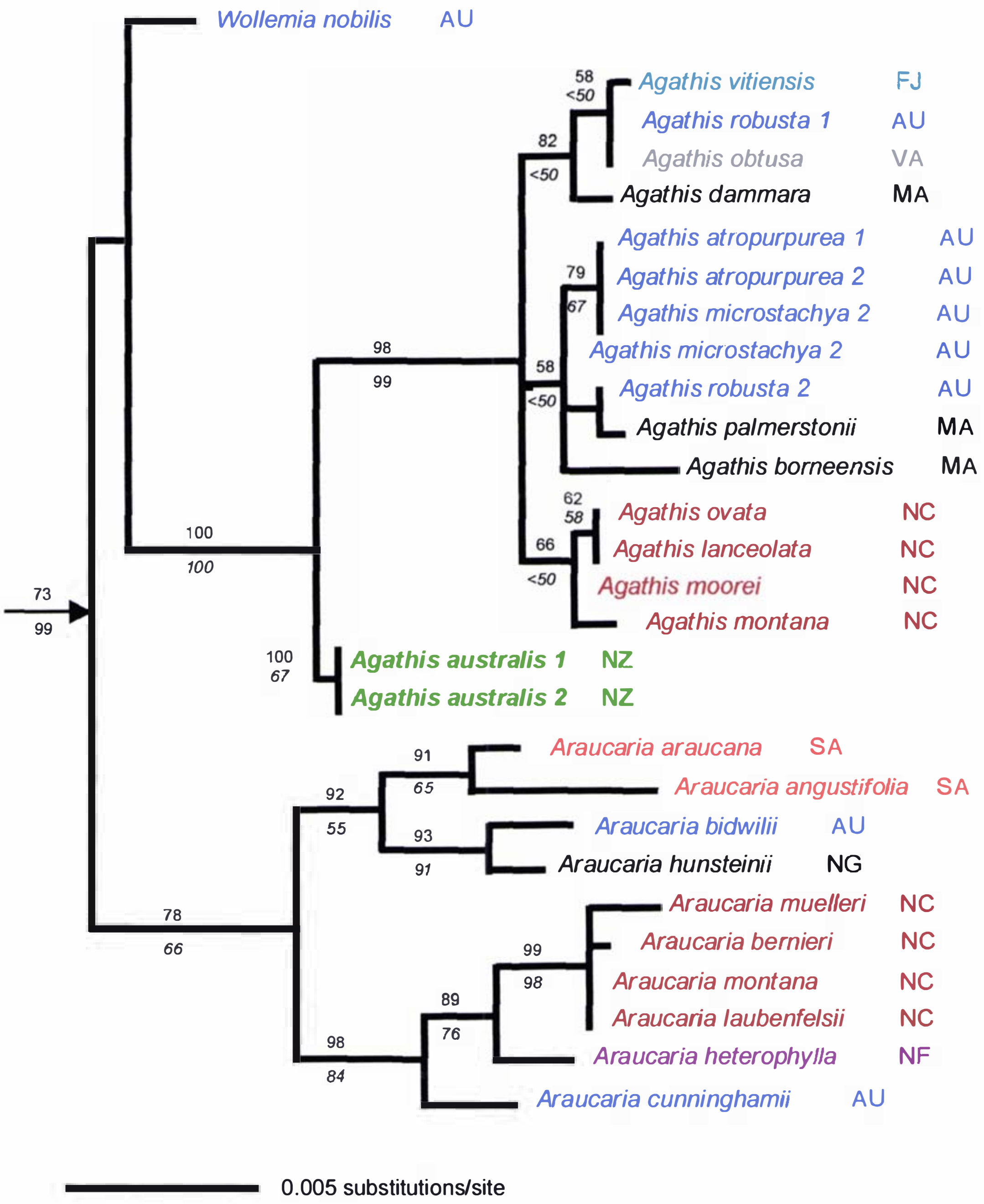


Figure 3.2: Phylogeny for Araucariaceae based on *rbc*L sequence data (1257 homologous sites). The position of the root is indicated by an arrow. Quartet puzzling tree made under Maximum likelihood using PAUP*. QP and bootstrap (italics) values are assigned to the edges. Substitution model: GTR + I. -Ln likelihood = 3195.47005. ML settings: number of substitution types = 6; nucleotide frequencies: empirical with: A = 0.26080 C = 0.20059 G = 0.24970 T = 0.28890; assumed proportion of invariable sites = 0.4708; distribution of rates at variable sites = equal. Abbreviations: **AU**: Australia; **FJ**: Fiji; **MA**: Malaysian region; **NC**: New Caledonia; **NG**: New Guinea; **NF**: Norfolk Islands; **NZ**: New Zealand; **SA**: South America; **VA**: Vanuatu.

Phylogenies obtained with quartet puzzling under the HKY and Fe84 models differ from the GTR tree in suggesting different relationships between closely related species of *Agathis*. However, the position of *Agathis australis* and the three major groupings recovered for the Austral-Pacific radiation of *Agathis* remains unchanged. Phylogenetic relationships within the genus *Araucaria* do not change when an alternative substitution model and different parameter settings are used. Under the Fe84 model the root position joins on *Wollemia*. The change in root position places *Wollemia* basal to *Araucaria* and *Agathis* (result not shown).

3.2.2.3 Molecular Clock Tests and *rbc*L-substitution Rate Estimates

A molecular clock test was performed with SplitsTree 3.1 using the method described by Steel *et al.* (1996). The Araucariaceae *rbc*L sequence alignment included *Araucaria* spp., *Agathis* spp., and *Wollemia nobilis*. The root position was determined by outgroup placement with *Taxus baccata, Sequoia sempervirens* and *Juniperus conferta* (see Figure 3.1).

Molecular clock tests failed for *Agathis australis* and *Wollemia nobilis*. Both the splitsgraph and the quartet puzzling tree place these taxa on short external edges, which suggests slower substitution rates in the evolution of the *rbc*L regions of *A. australis* and *Wollemia*.

An estimate was made for the rate of araucarian *rbc*L base substitution using the 'diverge' option under SplitsTree as described by Steel *et al.* (1996). For this, the divergence between the two genera *Araucaria* and *Agathis* was estimated. Taxa that failed to fulfil MC assumptions were not included in these estimates. From the original *rbc*L data set, all ambiguous characters were excluded from analyses. Invariable sites were estimated under SplitsTree 3.1 and 651 sites were removed.

Between both genera a genetic distance of DIST = 0.027 +/- 0.004 was calculated. To calibrate this rate of change with absolute time, a divergence time of 130 million years between genus *Agathis* and *Araucaria* was assumed from the fossil evidence of Cantrill (1992) and Stockey, (1982). Using this fossil date the substitution rate was evaluated for Araucariaceae *rbc*L gene sequences: $s_r = (0.208 \pm 0.017) \times 10^{-9}$ substitutions per site per year. Details and results of the estimate are given in Table 3.1.

number of base pairs in original data set:	1269 bp
molecular clock tests:	MC assumption failed for *Agathis australis* and *Wollemia nobilis*
excluded sites:	ambiguous sites: 14 sites
	invariable sites = 651
distance measure:	Genetic distance between *Agathis* spp. and *Araucaria* spp.
calculated distance:	DIST = 0.027 +/- SD = 0.004 [substitutions per site]
fossil date:	Divergence between Araucaria and Agathis in the Early Cretaceous: 130 million years (260 m. y. parallel evolution)
substitution rate	$s_r = (0.208 \pm 0.017) \times 10^{-9}$ [substitutions per site per year]

Table 3.1: Details and results of substitution rate estimate for the *rbc*L gene region of Araucariaceae.

3.2.2.4 Estimating the Age of the *A. australis* Lineage

While the placement of *Agathis australis* in reconstructed phylogenies suggests that it represents the earliest diverged lineage within the genus, molecular clock tests (2.4.6) show that the *rbc*L gene of *Agathis australis* has not evolved under molecular clock assumptions. This poses a problem for estimating the time of divergence for this species from other *Agathis* species. As observed from the splitsgraph and quartet puzzle trees, the relatively short branch leading to *A. australis* suggests a slow down of the substitution rate on the *Agathis australis* lineage. Nevertheless, using the *rbc*L substitution rate obtained from other araucarian species whose evolution of *rbc*L sequences is described by a molecular clock, it is possible to apply this rate to the *A. australis* lineage and obtain a lower bound estimate for the divergence time of *A. australis* from other species of *Agathis*.

The distances between internodes were taken from the quartet puzzling tree shown in Figure 3.2. This procedure could be employed as the distances between *Araucaria* and *Agathis* estimated under maximum likelihood criteria correspond to the SplitsTree 3.1 distance values. Estimates made using this procedure suggest that *A. australis* has diverged from the main *Agathis* lineage at least 70 million years ago (Figure 3.3). These results are consistent with *A. australis* having had a continuous presence on the New Zealand archipelago since the break up of Gondwana.

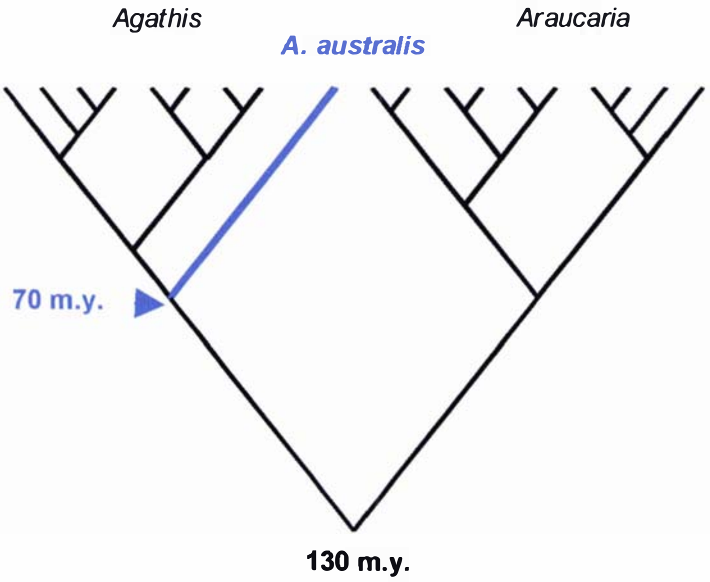


Figure 3.3: Scheme of divergence for Araucariaceae. The divergence time of 130 million years between the two genera *Araucaria* and *Agathis* is taken from fossil data (Cantrill, 1992; Stockey, 1982). The suspected divergence of *Agathis australis* from the main lineage (70 million years) was calculated using a substitution rate of sr = $(0.2077 \pm 0.0173) \times 10^{-9}$ substitutions per site per year.

3.2.3 Discussion: *Agathis australis* - a Living Fossil?

Investigations showed that the *rbc*L sequences of Araucariaceae have very treelike properties which make this marker highly suitable for phylogenetic analyses. A molecular clock assumption was fulfilled for most taxa allowing an average *rbc*L substitution rate to be calculated for the *Araucariaceae.*

In all phylogenies recovered from these data, the New Zealand endemic *Agathis australis* (kauri) is the first taxon to diverge from the main *Agathis* lineage. The New Zealand species is genetically most distinct from other *Agathis* species from Australia and the Polynesian Islands, which share closer phylogenetic relationships among each other. These findings demonstrate that *Agathis australis* is the sole member of an ancient lineage, which has been isolated from other species of the genus *Agathis* for a long period of time.

The New Zealand kauri being the sole member of an ancient lineage has parallels with the situation of *Wollemia* in Australia, which must have survived in isolation for a long period of time. Interestingly, molecular clock tests showed that for the *rbc*L region rates of sequence evolution in these taxa were much lower than in other araucarian taxa. A prevailing slow rate of sequence evolution may explain why *Wollemia nobilis* has been described as a 'living fossil' which means that the extant species seems more closely related to long extinct species than to other living taxa. Hence, molecular and morphological analyses are consistent in suggesting that this species - like the New Zealand *Agathis* - has retained ancestral characters.

The result from divergence time estimates for *A. australis* and other extant species of *Agathis* (Figure 3.3) would be consistent with a continuous presence of *Agathis* on the archipelago since the break-up of Gondwana. Thus, the molecular studies presented in this section reject the hypothesis proposed by Pole (1994) that the entire New Zealand flora went extinct during the Oligocene. Rather, the results support the hypothesis by Daniel (1989) that *Agathis australis* has directly evolved from the fossil species *A. seymouricum* which he had excavated from Cretaceous sediments in the New Zealand South Island.

3.3 PODOCARPACEAE ENDLICHER 1847

3.3.1 Introduction

Podocarpaceae are another gymnosperm family from the order of Pinales with a southern hemisphere distribution. The family includes about 15 genera and 150 described species.

The family is predominantly found in the Australasian region, with most taxa native to areas between New Zealand and south-east Asia. Most of the species are not of wide distribution, being confined to one or a few islands: namely, Tasmania, New Zealand, New Caledonia, New Guinea, Philippines, and Borneo. Some species of the genera *Dacrydium* Sol. ex Forst. f., *Lepidothamnus* Philippi, *Nageia* Gaertn., *Podocarpus* L'Heritier, *Prumnopitys* Philippi and *Saxegothaea* Silba are found beyond Australasia: in India, Japan, China, Africa, the Caribbean and in South America, from Mexico to Chile. Of these, *Saxegothaea*, a monotypic genus found in Chile and Argentina, is the only genus with no representatives in the Australasian region. Most members of the family are native to wet tropical or subtropical forests (often tropical mountains).

All species are evergreen trees or shrubs typically having a straight trunk and more or less horizontal branches. The leaves are usually spirally arranged, sometimes opposite, scale-like, needle-like, or more apart, flat and leaf-like, linear to lanceolate. Plants can be monoecious or dioecious. Seeds are completely covered by a fleshy structure referred to as an epimatium. Pollen grains are usually winged and airborne. Seedlings have two cotyledons with two parallel vascular bundles.

The Podocarpaceae are a very ancient family with a good fossil record. Although macrofossils have been described from the Northern Hemisphere, the majority of ancient podocarp remains have been described from the Southern Hemisphere (Hill and Brodribb, 1999). Coniferous fossil remains excavated from the uppermost Permian sediment in south China exhibit wood and leaf venation similar to modern Podocarpaceae and suggest that podocarpaceous conifers may have evolved by the end of the Palaeozoic (Yao *et al.*, 2000). This view is corroborated by leaf fossils assigned to Podocarpaceae recovered from Triassic strata in Antarctica and New Zealand (Axsmith *et al.*, 1998; Pole, 1995, 1997).

3.3.1.1 New Zealand Podocarpaceae:

17 podocarp species occur in New Zealand, Allan (1982) distinguishes three genera: *Podocarpus* L'Heritier, *Dacrydium* Sol. ex Forst. f., and *Phyllocladus* L.C. et A. Rich. Mixed podocarp-hardwood forests cover large areas on the West Coast of the South Island and the central North Island. Most species prefer warm-temperate and humid climates. The large trees have played an important role for the New Zealand logging industry at end of the 19th and the beginning of the 20th century.

The genus *Podocarpus* comprises over 70 species world wide, New Zealand has seven endemic tree and shrub species (Allan, 1982). Among these are *Podocarpus totara* G. Benn ex D. Don. in Lamb. (totara) and *P. hallii* Kirk. (mountain totara). Both species are tall, long living forest trees, which can acquire massive trunks. Two smaller tree or shrub species are *P. acutifolius* Kirk., occurring only in the South Island from Marlborough Sounds to Westland and *P. nivalis* Hooker, a subalpine scrub with wide spreading branches.

More and more authors have now excluded *P. spicatus* R. Br. ex Mirbel (matai or black pine) and *P. ferrugineus* G. Benn. ex D. Don. in Lamb. (miro or brown pine) from *Podocarpus* and placed both taxa in a separate genus: *Prumnopitys* Philippi (*Prumnopitys taxifolia* (Sol. ex D. Don) de Laub. and *Prumnopitys ferruginea* (D. Don) de Laub). Both species are trees of the lowland forest and widespread throughout New Zealand.

P. dacrydioides A. Rich., the white pine or kahikatea, is also often placed in a different genus: *Dacrycarpus dacrydioides* (A. Rich.) de Laub. With 60m height it is the tallest New Zealand tree species. It is found throughout the country and on Stewart Island in lowland forests up to 600 m, often dominating in swampy areas.

Allan (1982) also describes seven New Zealand species within genus *Dacrydium*. However, the genus has lately been subdivided into *Halocarpus* C. J. Quinn (three species), *Lepidothamnus* R. A. Philippi (two species), and *Lagarostrobos* C. J. Quinn (one species), leaving one species within the New Zealand endemic genus *Dacrydium*, which is *D. cupressinum* Lamb. with its common name rimu or red pine. Rimu is an attractive tall canopy tree of the New Zealand lowland and montane forest.

Lagarostrobos colensoi (silver pine) is a New Zealand endemic species. The closest relative is *L. franklinii* (houn pine) which is restricted to Tasmania. Molloy (1995), however suggested to exclude this species from genus *Lagarostrobos* because of striking morphological differences. He established the new monotypic genus *Manoao* to accommodate the New Zealand silver pine. Nevertheless, a close phylogenetic relationship between the two species is supported by molecular analyses based on *rbc*L sequence data (Conran et al., 2000).

The genus *Phyllocladus* comprises about six species with representatives in Tasmania, New Guinea, Borneo and the Philippines. Allan (1982) describes three species of this genus with its typical phylloclades of flattened branchlets in New Zealand. The three species are: *P. alpinus* Hooker a mainly subalpine scrub, *P. glaucus* Carr. a rather small tree found only in the North Island and *P. trichomanoides* D. Don. in Lamb. a member of the lowland forest in the North and South Island. In 1995, Molloy described the new species *P. toatoa* which had formerly been included in *P. glauca*. Traditionally the genus is placed within the family Podocarpaceae (e.g. Kelch 1998). More recently its separation into the monotypic family Phyllocladaceae Bessey has received some recognition (Molloy and Markham, 1999), although some morphological (Quinn, 1987) as well as molecular (Conran *et al.*, 2000) evidence suggests a phylogenetic positioning of *Phyllocladus* within Podocarpaceae.

Podocarpaceae are well represented in the New Zealand fossil plant record. Mildenhall (1980) reports the first appearance of pollen of the genus *Podocarpus* in Mesozoic sediments, *Dacrydium* in Cretaceous and *Dacrycarpus* as well as *Phyllocladus* in Upper Eocene sediments of New Zealand. Pole (1997) describes one extinct species of *Prumnopitys* excavated from Palaeocene strata in south Canterbury.

3.3.2 Phylogenetic Analyses of Podocarpaceae 18S Sequence

For phylogenetic analyses on Podocarpaceae, partial 18S sequences of nuclear ribosomal DNA (rDNA) were taken from Chaw *et al.* (1993) and Kelch (1997) and obtained from the NCBI GenBank. From 16 available sequences 15 were used in the present analyses. With five species, the New Zealand podocarps were well represented in the phylogeny. 18S rDNA sequences from three species of Araucariaceae were used for root placement analyses. Sequence data from *Agathis robusta* Bailey, *Araucaria columnaris* (Forst.) Hook. and *A. heterophylla* (Salisbury) Franco were taken from Graham *et al.*, (1996). The treelike properties of the data were investigated under split decomposition (Huson, 1998). Maximum likelihood analyses included model fitting and tree building with quartet puzzling as detailed in Section 3.1.3.

3.3.2.1 Split Decomposition

The splitsgraph shown in Figure 3.4 visualises the phylogenetic relationships between members of the Podocarpaceae as suggested by the 18S rDNA sequence. Contradictory patterns in the 18S data are visualised as box-like patterns in the centre of the splitsgraph.

The graph was made under a least squares option – which allows re-estimation of edge lengths in a splitsgraph using a least squares criterion. This approach may be helpful if the fit statistic value of a splitsgraph is less than 100% and the external edges in the reconstructed splitsgraph appear disproportionately long with respect to internal edge lengths. On the 18S data, split decomposition without re-estimated edge length shows a relatively low fit value (68.6) for the graph depicting 15 taxa and great differences in external edge length. This suggests that some distances in the distance matrix are not well represented by the graph.

However, in spite of edge length re-estimation, differences in relative edge length are still a prominent feature of the re-evaluated splitsgraph. The presence of long and short edges may indicate that the 18S genes of the analysed podocarp species evolve at different rates. This observation is corroborated by results from molecular clock analyses made under SplitsTree (Figure 3.4).

The differences in edge lengths, the low fit value and the network structure in the splitsgraph indicate that the 18S data has limited suitability for phylogenetic tree-building analyses. Nevertheless, some relationships can be inferred using the 18S data.

The phylogeny recovered with split decomposition separates two groups: group one includes the genera *Phyllocladus*, *Lagarostrobos*, *Parasitaxus* and *Prumnopitys* , group two ('*Podocarpus* group') comprises the genera *Podocarpus*, *Dacrydium*, *Dacrycarpus*, *Nageia*, *Microstrobos*, *Microcachrys* and *Saxegothaea*. The root - as determined by outgroup placement with Araucariaceae - joins between both subgroups.

The graph shows that within the second group, *Dacrydium cupressinum*, diverges from the main lineage at an early evolutionary stage. Split decomposition places this species closest to the root on the connecting edge between both podocarp groups. Here, ambiguities in the phylogenetic signal are visualised by a network structure.

Interestingly, *Dacrycarpus dacrydioides* the New Zealand kahikatea and the African species *Nageia nagi* share the same position in the phylogeny. They appear as rather basal taxa as on an internal node at the radiation point of *Podocarpus* group, indicating that both taxa have retained ancestral 18S sequence characters.

A New Zealand and a Malaysian species of *Phyllocladus* appear in the splitsgraph (and as will be seen also in the Quartet Puzzle tree) as relatively recently diverged taxa. However, the *Phyllocladus* lineage, which is represented by these two species, is significantly diverged from the remaining Podocarpaceae. Even more divergent is the *Prumnopitys taxifolia* lineage that derives from a similar common ancestor shared between *Phyllocladus*, *Lagarostrobos* and *Parasitaxus* species.

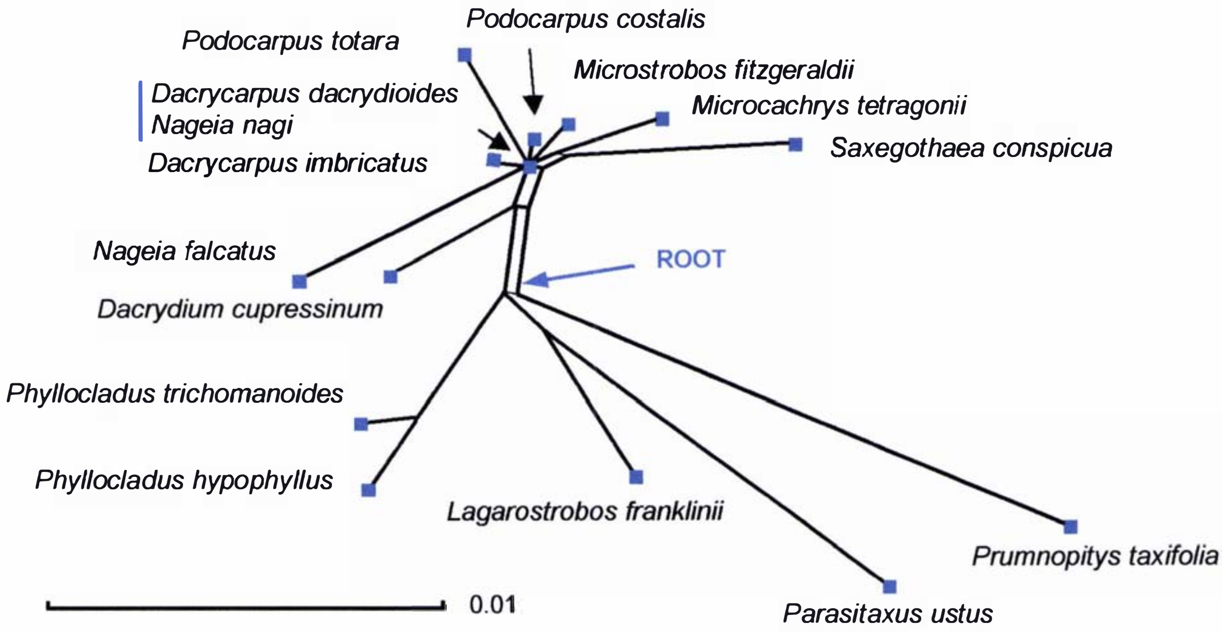


Figure 3.4: Splitsgraph with least squares estimated edge lengths for 15 species of Podocarpaceae based on the 18S of nuclear rDNA. made without outgroup sequences. The position of the root as determined by outgroup placement with *Agathis robusta*, *Araucaria columnaris* and *A. heterophylla* is indicated by a blue arrow. Scale: 0.01 substitutions per site. Details: Fit = 68.6 (determined on a splitsgraph without re-estimated edge length) taxa = 15, characters = 1403(of 1682), gaps = 66(off), missing = 213(off), constant = 1303, nonparsimony = 1371, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

3.3.2.2 Maximum Likelihood Analyses

The phylogeny of Podocarpaceae as shown in Figure 3.5 was constructed using quartet puzzling and maximum likelihood criteria. Model fitting analyses identified the GTR substitution model to best describe the evolution of the podocarp 18S data. The chosen parameter settings included an invariable sites model.

Root placement analyses using Araucariaceae 18S sequence gained high support under quartet puzzling and bootstrapping. The position is locally stable and joins on the long edge of the *Phyllocladus* lineage. Under alternative model and parameter settings maximum likelihood recovers the same root position as split decomposition which is situated between *Phyllocladus*, *Prumnopitys*, *Lagarostrobos*, *Parasitaxus* and the remaining taxa (result not shown).

Low support values were recovered by quartet puzzling and bootstrapping for phylogenetic relationship between ingroup taxa. This result corroborates analyses with split decomposition, and suggests limited suitability of these data for tree building purposes.

Bootstrapping under the chosen model and parameter settings recovers support greater 50% in only three places: the phylogenetic relationship of both *Phyllocladus* species gains 93% bootstrap support and a relationship between *Lagarostrobos* and

Parasitaxus is weakly supported by 82%. A split separating the four genera *Phyllocladus*, *Prumnopitys*, *Lagarostrobos* and *Parasitaxus* from the remaining taxa is also supported by 60% bootstrap support.

Maximum likelihood analyses on the 18S data recover some support for a phylogenetic relationship between *Lagarostrobos*, *Parasitaxus*, and *Prumnopitys*. An alliance between *Microstrobos, Saxegothaea*, and *Microcachrys* is also supported (QP value = 61). No support is gained for a phylogenetic relationship between both species analysed for genus *Podocarpus*. While the Australian species appears in a group with *Nageia* spp. the New Zealand member of this genus groups with genus *Dacrycarpus*. This result suggests a polyphyletic origin of genus *Podocarpus*.

Consistent with split decomposition, quartet puzzling suggests that *Dacrydium cupressinum* is an early diverged lineage.

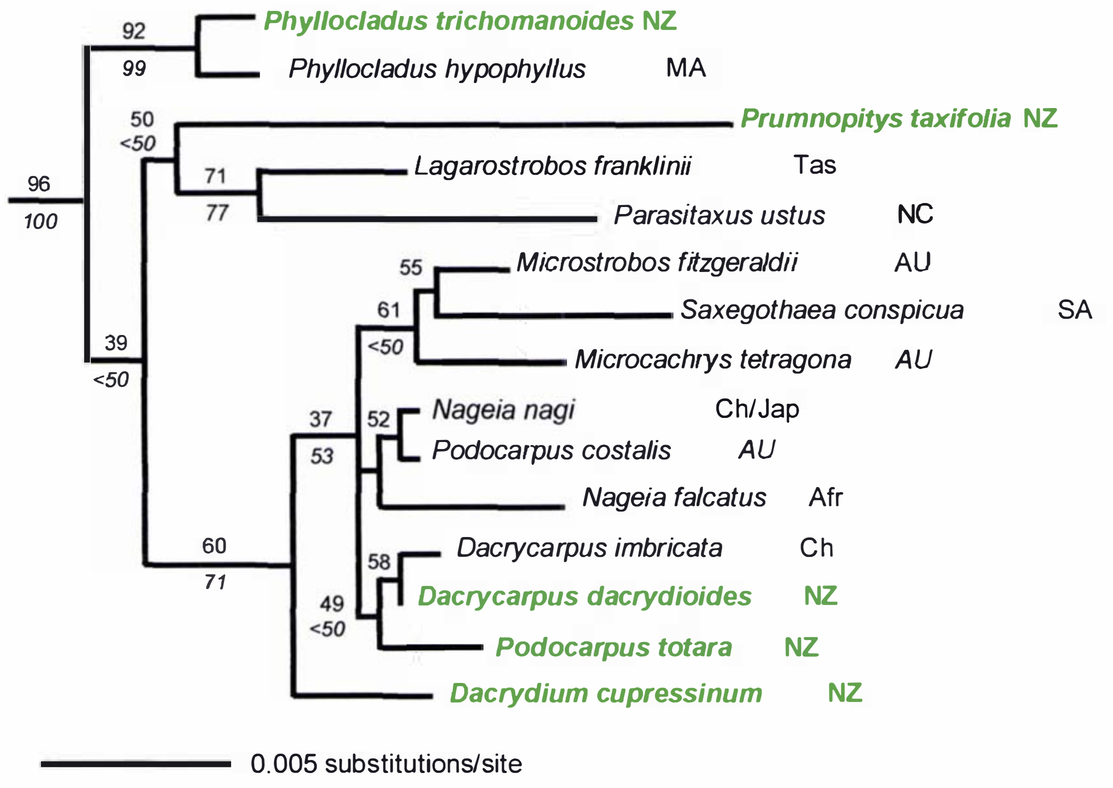


Figure 3.5: Rooted phylogeny for Podocarpaceae based on 18S nuclear rDNA sequence and made with quartet puzzling. Edge lengths were estimated under ML criteria. Araucariaceae were used for outgroup placement. QP-values and bootstrap support (italics) , taken from 100 replicates, are added onto the branches. Substitution model: GTR+I. -Ln likelihood = 3529.45345. ML settings: number of substitution types = 6; nucleotide frequencies: empirical with: A = 0.26080 C = 0.20059 G = 0.24970 T = 0.28890; proportion of sites assumed to be invariable = 0.851412; distribution of rates at variable sites = equal. Abbreviations: **Afr:** Africa; **AU:** Australia; **Ch:** China; **Jap:** Japan; **MA:** Malaysia; **NZ:** New Zealand **Tas:** Tasmania.

3.3.2.3 Molecular Clock Tests

The method of Steel *et al.* (1996) was used to test whether the evolution of the 18S rDNA sequences was clock like. The molecular clock tests were made using taxa one either side of the root position shown in the QP tree of Figure 3.5. Not surprisingly, the podocarp 18S rDNA data showed a lack of a universal molecular clock. Calculations of substitution rate and divergence time estimates were not made for this data set.

3.3.3 Discussion: Are New Zealand Podocarps Gondwanan Relicts?

Although a small selection of Podocarpaceae were included, the analyses contribute to an understanding of phylogenetic relationships between extant New Zealand and overseas species. However, evidence for the evolutionary history of podocarps is limited as only low support values were recovered for most internal edges of the maximum likelihood phylogeny (Figure 3.5). The splitsgraph (Figure 3.4) also showed a rather low fit value and a non bifurcating network structure in which some lineages clearly evolved in a non clock like manner.

Some patterns in the 18S rDNA sequence suggest that *Dacrydium cupressinum* has diverged early from the other investigated species. Thus, it may potentially be representative of an ancient New Zealand podocarp lineage. Phylogenetic analyses of 18S sequence do not exclude the possibility that rimu is a truly Gondwanan element of the New Zealand flora that has had a continuous presence on the archipelago. This inference would be consistent with findings from fossil pollen analyses that recognise this genus from Cretaceous sediments in New Zealand (Mildenhall, 1980). During this era the islands were still part of the Gondwanan supercontinent. As phylogenetic analyses of *rbc*L sequences suggest close relationships between the New Zealand rimu and overseas species of *Dacrydium* (Conran *et al.* 2000), it will be interesting to investigate the divergence time between the extant species of this genus. Such an approach might bring forward evidence for (or against) a vicariant evolution of the genus *Dacrydium*.

More difficult to interpret are the observations under split decomposition, that show the New Zealand species *Dacrycarpus dacrydioides* (kahikatea) at the radiation point of the *Podocarpus* group. Its placement would also suggest it is an early diverged species whose 18S sequence has retained ancestral sequence characters. However, quartet puzzling analyses do not fully support an ancestral status of this taxon. In contrast, the method recovers weak support for a group comprising kahikatea, a Chinese species of the same genus and *Podocarpus totara* (totara) from New Zealand.

In the splitsgraph, totara is positioned on a monotypic branch born in the centre of the radiation. Neither split decomposition nor quartet puzzling recovers support for a direct phylogenetic relationship between the geographically distinct species *Podocarpus totara* and *Podocarpus costalis*. However, this result, which implies that the genus *Podocarpus* is polyphyletic, is not supported by phylogenetic analysis of the *rbc*L gene (Conran *et al.*, 2000). In this study, the authors found great support for the monophyletic origin of all podocarp genera.

18S analyses provide no evidence that the African genus *Nageia* or the South American genus *Saxegothaea* have diverged earlier than the two New Zealand taxa *D. dacrydioides* and *P. totara*. However, more comprehensive molecular studies will be necessary to date the divergence between these extant lineages.

Molecular data shows great genetic differentiation between the genus *Phyllocladus* and other podocarp lineages. In the splitsgraph (Figure 3.3) the genus is situated on a long external edge. This suggests that *Phyllocladus* is a distinct lineage, which has diverged early during the evolution of the Podocarpaceae. This result is in accordance with morphological analyses (Quinn, 1987) and molecular phylogenetic studies using *rbc*L sequences (Conran *et al.* 2000). Although other lineages namely *Prumnopitys* or *Lagarostrobos* show greater genetic differentiation, he tendency in ML analyses for these lineages to join near the root provides some support for the exclusion of genus *Phyllocladus* from the family of Podocarpaceae as suggested, among others, by Molloy and Markham (1999).

Both split decomposition and quartet puzzling suggest that the two species analysed for genus *Phyllocladus* have diverged from a common ancestor rather recently and long after evolution of the lineage. The low genetic distance between the New Zealand and Malaysian species suggests a relatively recent dispersal event. However, as there are no data available for other overseas species of this genus, there are no indications about the direction of dispersal. Moreover, because of the lack of other overseas taxa in the data set, no conclusions can be drawn whether the extant lineage has had an ancient and continuous presence in New Zealand or whether it is part of a more recent radiation. Although *Phyllocladus* has an ancient presence in the New Zealand fossil pollen record since the Upper Eocene, the hypothesis of extinction and recent re-dispersal into New Zealand cannot yet be ruled out.

In the 18S phylogeny the *Prumnopitys taxifolia* lineage derives early from a common ancestor whom it shares with *Lagarostrobos* and *Parasitaxus*. Fossil pollen analyses do not distinguish this species from other podocarps in New Zealand. Because of the lack of overseas *Prumnopitys* species (e.g. the New Caledonian *P. ferruginoides* (Compt.) de Laub. or the South American species *P. andina* (Poepp. ex Endl.) de Laub.) no conclusions can be drawn concerning the evolution of this lineage and its presence in New Zealand. Including these overseas taxa would allow researchers to test the monophyly of *Prumnopitys* as well as infer dispersal routes within the genus.

Further molecular analyses may help to elucidate interspecific relationships for a taxonomic revision of Podocarpaceae. Furthermore, studies using other conservative genome regions (e.g. the *rbc*L gene) may allow researchers to date the divergence between extant podocarp lineages.

3.4 METROSIDEROS BANKS EX GAERTN. 1788 (MYRTACEAE)

3.4.1 Introduction:

The genus *Metrosideros* Banks ex Gaertn. belongs to the large family Myrtaceae which is one of 14 families in the order Myrtales (Cronquist, 1998). The family hosts about 130 genera with over 3000 described species. With a centre of distribution in the Southern Hemisphere, Myrtaceae are wide-spread in temperate, sub-tropical, and tropical climates, and a high number of species occur in warm-temperate Australia and tropical America.

Metrosideros contains about 60 species of trees, shrubs and lianas found in South Africa (one species), the Philippines, New Guinea, north-east Australia, New Caledonia, New Zealand, and some Polynesian Islands. Their aromatic, gland-dotted leaves are always simple and without stipules in opposite arrangement on the branch. Most species are aromatic and have lysigene glands, which bear essential oils. Flowers are 5-merous and aggregated in inflorescences with showy stamen and arranged in terminal cymes or racemes. Petals are frequently reduced and inconspicuous compared to the well-developed androecium of numerous stamens. The gynoecium consists of a single compound pistil of 2-5 carpels, a single style and stigma, and a partly too wholly inferior ovary. Plants are usually hermaphrodite. *Metrosideros* seeds are linear coriaceous capsules.

Based on morphological evidence, Dawson (1976) subdivided the genus *Metrosideros* into three subgenera: subgenus *Metrosideros* with around 23 species, subgenus *Mearnsia* with about 25 species and subgenus *Carpolepis* with three species. In 1984, however, he reconsidered his revision, excluded subgenus *Carpolepis* and raised it to generic rank (Dawson, 1984).

Unfortunately, there is no fossil evidence for the occurrence of *Metrosideros* apart from fossil pollen excavated from Palaeocene sediments in New Zealand (Mildenhall, 1980; Dawson, pers. comm.).

3.4.1.1 New Zealand *Metrosideros*:

New Zealand is well known for its *Metrosideros* species. During December and January when these trees are in bloom, *M. excelsa* Sol. ex Gaertn. (NZ Christmas tree or pohutakawa), *M. umbellata* Cav. (southern rata) and *M. robusta* A. Cunn. (northern rata) provide a most beautiful sight with their bright red flowers. Twelve endemic *Metrosideros* species of shrubs, trees or lianas are currently described in New Zealand and its off shore islands. These are: *M. albiflora* Sol. ex Gaertn., *M. bartlettii* Dawson, *M. carminea* W.R:B. Oliver, *M. colensoi* Hook. f., *M. diffusa* (G. Forst. f) W.R:B. Oliver, *M. excelsa, M. fulgens* Sol. ex Gaertn., *M. perforata* (J. R. Forst. & G. Forst.) A. Rich., *M. parkinsonii* Buchanan., *M. robusta, M. umbellata*, and *M. kermadecensis* W. R. B. Oliv., which is endemic to the Kermadec islands.

As mentioned previously, *Metrosideros* has an ancient presence in the New Zealand fossil pollen record. Pollen of the genus has been recorded from Palaeocene sediments (Mildenhall, 1980).

3.4.2 Phylogenetic Analyses of *Metrosideros* Nuclear rDNA Sequence

Sequence analyses were performed on a DNA alignment consisting of 623 homologous sites of the ITS1, ITS2 and 5.8S regions of nuclear rDNA. The alignment, obtained with ClustalX, comprised all rDNA sequences that were available at the NCBI GenBank for New Zealand taxa as well as a representative number of overseas species. Sequence data for *Cloezia floribunda* Brongn. & Gris was included for outgroup placement. All sequences except for *Metrosideros angustifolia* Dum. Cours. were taken from the study by Wright *et al.* (2000). The rDNA region of *M. angustifolia* was amplified from genomic DNA, which was extracted from herbarium material. DNA extraction protocol, PCR and DNA sequencing procedures were performed as described in Section 2.3. The herbarium voucher number for this specimen is listed in Appendix II.

3.4.2.1 Split Decomposition

To examine the treelike properties of the *Metrosideros* rDNA data, the sequence alignment was investigated using split decomposition under SplitsTree 3.1 (Huson, 1998). The splitsgraph for this data is shown in Figure 3.5. It has a very low fit statistic value of 48.0 indicating that some distances in the data are not well represented by the graph. Excluding the outgroup has no positive effect on the presentation of the data. Relationships among *Metrosideros* species remain identical while the fit value decreases to 44.3 (result not shown).

decomposition. However, edge fitting does not significantly alter the shape of the graph.

One internal split separates two groups, but a low bootstrap value of 36 is assigned to this edge. The low bootstrap value may indicate that only a small number of patterns support this split or that contradictory patterns are present in the sequence data.

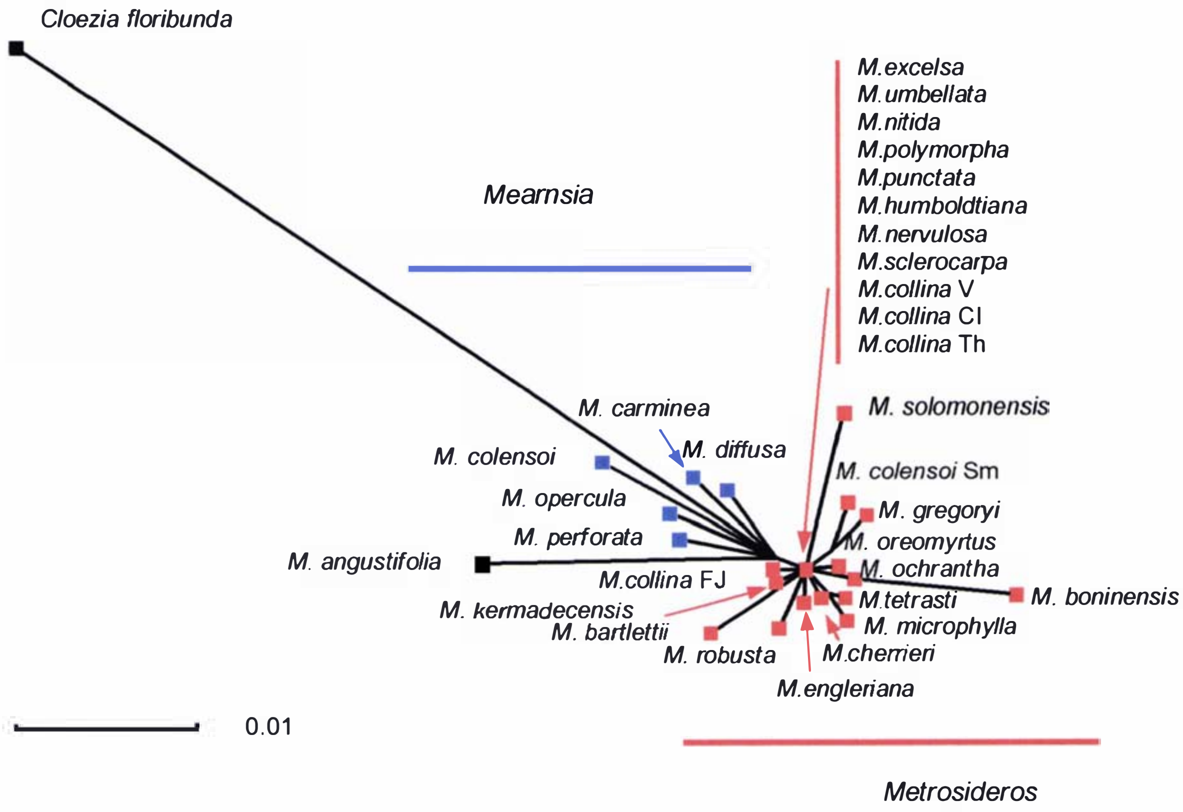


Figure 3.6: Splitsgraph for New Zealand and overseas species of *Metrosideros* based on ITS1, ITS2 and 5.8S of nuclear rDNA. *Cloezia floribunda* (Myrtaceae) is used as outgroup. Scale: 0.01 substitutions per site. Details: Fit = 48.0, characters = 566(of 623), gaps = 48(off), missing = 9(off), constant = 473, nonparsimony = 536, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

The two groups separated by the internal edge distinguish the morphological subgenera recognised by Dawson (1976). Group 1, subgenus *Mearnsia* combines the African species *M. angustifolia*, *M. opercula*, a species from New Caledonia and the four New Zealand taxa: *M. perforata*, *M. colensoi*, *M. diffusa* and *M. carminea*. *Cloezia floribunda,* used as outgroup, joins at the basis of this star like phylogeny.

At the radiation point of the second group, which represents subgenus *Metrosideros*, are two New Zealand species, *M. excelsa* and *M. umbellata* and a large number of taxa from other Pacific islands like Hawaii (e.g. *M. polymorpha*) and New Caledonia (e.g. *M. nitida*), the Lord Howe Islands (*M. sclerocarpa*), Tahiti and Cook Islands (both *M. collina*). The New Zealand species *M. bartlettii*, *M. robusta*, *M. kermadecensis* and a number of taxa from other Pacific landmasses also belong to the second group. Split decompositions does not resolve all relationships within this group of taxa because (i) the rDNA sequence of some species within this unresolved cluster (*M. excelsa*, *M. polymorpha*, *M. nervulosa*, *M. sclerocarpa, M. collina* Vanuatu, *M. collina* Cook Island, *and M. collina* Tahiti) is identical and (ii) some internal edges describing the phylogenetic relationships between closely related taxa would have to be visualised as a network structure which is collapsed by adding more distantly related taxa. As the genetic distances between some species are not represented in the graph, the SplitsTree fit value is decreased.

The splitsgraph shown in Figure 3.7 was made with New Zealand taxa only. The graph has a fit value of 81.4. The root position was tested with *Cloezia floribunda.* and alternatively with *Metrosideros angustifolia.* With both taxa it is identical and joins the phylogeny between both subgroups. In the splitsgraph (Figure 3.7), the root position is indicated by a blue arrow. While the fit values for the splitsgraph increases by choosing *Cloezia* (fit = 88.0) as an outgroup taxon, it decreases slightly by adding *M. angustifolia*: (fit = 81.4). However, as *Cloezia* is more distant to the ingroup taxa, split decomposition recovers some minor support for joining the outgroup with both subgenera *Mearnsia* and *Metrosideros* resulting in a box-like structure at the base of the graph (result not shown).

The relationships between New Zealand *Mearnsia* and *Metrosideros* species remain unchanged by the exclusion of overseas species as shown in Figure 3.7. However in this case, phylogenetic resolution is higher when less taxa are involved as there is a better fit between the pairwise distances that are calculated from the data matrix and distances that are represented in the splitsgraph.

Greater genetic differentiation is found within subgenus *Mearnsia*: between *M. perforata*, *M. colensoi*, *M. diffusa* and *M. carminea.* Under the assumption that all sequences have evolved under a molecular clock, this observation suggests that all members of *Mearnsia* have diverged from a common ancestor at an early stage within the evolution of the genus *Metrosideros.* All species seem approximately equidistant from the ancestral node.

Within subgenus *Metrosideros*, *M. umbellata* is the earliest diverged taxon and is closest to the root. Its position on a relatively short edge indicates that its rDNA may be evolving at a slower than other taxa within the group.

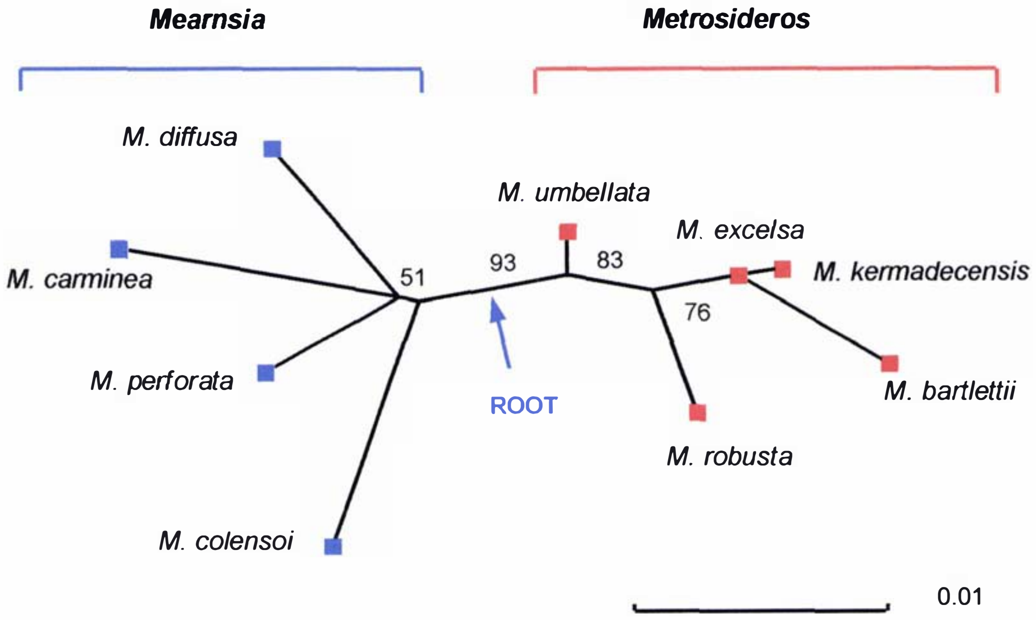


Figure 3.7: Splitsgraph for New Zealand *Metrosideros* species based on ITS1, ITS2 and 5.8S S nuclear rDNA sequence. Scale: 0.01 substitutions per site. Bootstrap values from 100 replicates are assigned to the edges. *M. angustifolia* is used as outgroup and the position of the root is indicated by a blue arrow. Details: Fit = 81.4, characters = 584(of 623) gaps = 37(off) missing = 2(off) constant = 545 non parsimony = 565 dsplits –hamming. For splitsgraph details see Section 2.4.4.4. Scale: 0.01 substitutions per site.

As visualised by the splitsgraph, *M. excelsa* displays an ancestral rDNA sequence type. The taxon appears basal to *M. bartlettii* and *M. kermadecensis.* The low genetic distance between the New Zealand mainland species *M. excelsa* and the Kermadec Islands species *M. kermadecensis* is consistent with a recent dispersal event from New Zealand to these distant oceanic islands.

3.4.2.2 Maximum Likelihood Analyses

Phylogenetic tree reconstruction for the New Zealand *Metrosideros* species was also carried out with quartet puzzling using maximum likelihood criteria. In this, model fitting analyses were performed using a likelihood ratio test with Paup* 4.65 (Swofford, 1999). Model fitting showed that the GTR model allowing a proportion of sites to be invariable provided the best explanation for the evolution of the rDNA data. These settings were used as optimality criteria for quartet puzzling and bootstrap analyses.

M. angustifolia was assumed to represent an outgroup taxon. This taxon forms a basal trichotomy with species from subgenus *Mearnsia* and subgenus *Metrosideros* (Figure 3.8). In root placement studies including *Cloezia*, *M. angustifolia* was found to be the earliest branching lineage within genus *Metrosideros* (result not shown).

Like split decomposition (Section 3.4.2.1), quartet puzzling recovers support for a split separating the two subgenera *Mearnsia* and *Metrosideros.* The method recovers good support for relationships within subgenus *Metrosideros.* QP values, which are assigned to the edges, are between 96 and 100. Phylogenetic relationships within subgenus *Meamsia* receive lower support under ML criteria. However, all internal edges visualised in the quartet puzzling tree are supported by bootstrap values greater 50%. Under altemative model and parameter settings quartet puzzling tends to group *M. carminea* with *M. perforata* rather than with *M. diffusa* while all other phylogenetic relationships remain unchanged (result not shown).

The phylogeny shown in Figure 3.8 also visualises the flower colour of the taxa used in these analyses. In general, species of *Metrosideros* species display red or white flowers. As the phylogeny clearly shows, red and white flowers are present in both subgenera. All species in subgenus *Meamsia* are climbers with small leaves, all except for *M. carminea* have pale white flowers. *M. carminea*, however, has bright red flowers.

In group two all species are trees although *M. umbellata* and *M. robusta* may grow as epiphytes when young. Except for *M. bartlettii,* all species in this group have red flowers. Thus the phylogeny suggests a polyphyletic origin of red flower colour in *Metrosideros*.

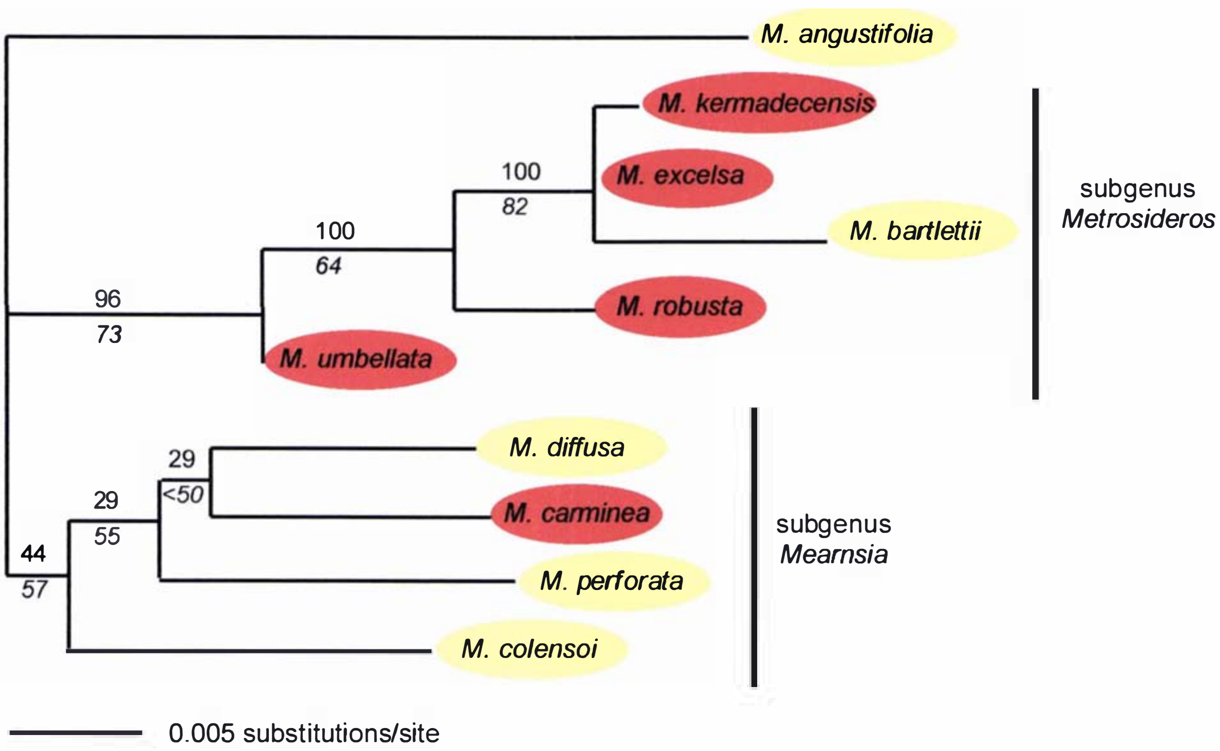


Figure 3.8: Unrooted quartet puzzling tree from *Metrosideros* nuclear rDNA. QP and bootstrap support values (italics) are assigned to the edges. The coloured ovals indicate the flower colour of the species. – ln likelihood = 1184.28411; substitution model: GTR+I; ML settings: number of substitution types = 6; nucleotide frequencies: empirical with A = 0.22388 C = 0.29038 G = 0.27732 T = 0.20842; assumed proportion of invariable sites = 0.845091; distribution of rates at variable sites = equal.

3.4.2.3 Molecular Clock Tests and Substitution Rate Estimates

A molecular clock tests was performed on all New Zealand species using the 'test_clock' option under SplitsTree 3.1 (Huson, 1998; Steel *et al.* 1996). Molecular clock tests indicate that the ITS regions of most species evolves according to molecular clock assumptions. The test showed that the evolution is non-clock-like in *M. umbellata* and *M. excelsa*.

Substitution rate estimates were not performed on this data set, as no fossil evidence is yet available to calibrate with respect to time the divergence between subgenera *Mearnsia* and *Metrosideros*.

3.4.3 Discussion: *Metrosideros* in New Zealand

Phylogenetic analysis of nuclear rDNA loci on New Zealand and overseas species of *Metrosideros* showed a low level of phylogenetic resolution between most species. This makes testing hypotheses difficult. Nevertheless, the subdivision of subgenera recognised by Dawson (1976) was recovered in both split decomposition and quartet puzzling analyses. Under split decomposition the outgroup *Cloezia floribunda* joined with subgenus *Mearnsia*, at the basis of the star like phylogeny (Figure 3.6). The closer alliance of this group with the outgroup is consistent with the greater morphological diversity within *Mearnsia* which suggests a possibly ancestral status of this subgenus (Dawson, pers. comm.).

Maximum likelihood analyses indicated that the evolution of flower colours is polyphyletic for New Zealand taxa. Except for *M. bartlettii* all members of subgenus *Metrosideros* have red flowers while all species from subgenus *Mearnsia* except *M. carminea* have white flowers. On the basis of the rDNA phylogeny recovered for *Metrosideros*, it is hypothesised that the last common ancestor of all New Zealand species had white flowers - like *M. angustifolia* (African Boxwood). Red flowers presumably evolved twice and independently: on the *M. carminea* lineage and on the lineage that gave rise to subgenus *Metrosideros*. The common ancestor of this subgenus most probably had red flowers. In *M. bartlettii* another mutation (or mutations) have restored the ancestral character of white flowers. White or pale mutants are also known from *M. excelsa* and *M. robusta*. This suggests that for *Metrosideros,* changes in flower colour may be a result of simple genetic changes that can occur relatively frequently.

Genetic differentiation between species of subgenus *Mearnsia* (as visualised in Figure 3.6 and displayed below in Tables 3.1 and 3.2) is somewhat higher than between species of subgenus *Metrosideros*. This observation suggests that the radiation from a common ancestor within subgenus *Mearnsia* lies back further in time than the last common ancestor shared by species belonging to subgenus *Metrosideros*. Thus, the radiation within subgenus *Metrosideros* may have been a more recent event.

	M. kermadecensis	***M. excelsa***	***M. bartlettii***	***M. robusta***
M. excelsa	0.0017	-		
M. bartlettii	0.0103	0.0086	-	
M. robusta	0.0120	0.0103	0.0189	-
M. umbellata	0.0137	0.0120	0.0206	0.0120
mean distance:	**0.0120 ± 0.0052**			

Table 3.2: Hamming distances between species of *Metrosideros* subgenus *Metrosideros*.

	M. diffusa	***M. carminea***	***M. perforata***
M. carminea	0.0189	-	
M. perforata	0.0189	0.0206	-
M. colensoi	0.0223	0.0274	0.0206
mean distance	**0.0220 ± 0.0032**		

Table 3.3: Hamming distances between species of *Metrosideros* subgenus *Mearnsia*.

The question whether or not extant species have had a continuous presence in New Zealand since the Palaeocene, when *Metrosideros* pollen is first recognisable in the fossil pollen record, cannot be answered on the basis of these analyses. The low level of phylogenetic resolution between New Zealand and Pacific species might be explained if species are of Late Tertiary and/or Quaternary age and are related by recent dispersal. In this case a hypothesis of extinction and subsequent re-dispersal into New Zealand should not be ruled out. Such an hypothesis might be tested if it were possible to calibrate with respect to time sequence divergence between the Pacific and the African species.

Further molecular analyses should be undertaken particularly in order to explain the phylogenetic relationships between *Metrosideros* and other genera of Myrtaceae. These studies, which could be based on more conserved regions of plastid and nuclear DNA might help to elucidate the evolution of the genus. In this context, additional fossil evidence may be valuable and reveal the origin and past distributions of *Metrosideros*. The availability of more fossil and molecular data would allow researchers to calibrate sequence divergence and make it possible to reconstruct dispersal routes for this genus during the Tertiary and Quaternary.

3.5 MYRSINACEAE R. BROWN 1810

3.5.1 Introduction

The phylogenetic position of the Myrsinaceae with respect to other families of flowering plants has been controversial. The family has generally been placed within the order Primulales (Cronquist, 1998) but also within the order Ericales (APG, 1995). The family comprises about 35 genera and 1000 species which have a distribution ranging from temperate to tropical climates, extending north to Japan, Florida and Hawaii, and South to Australia and New Zealand.

Most species within the family are wooden trees, shrubs or lianas. Leaves are exstipulate and petiolate and often gland-dotted with typically entire lamina margins. Leaf arrangement is alternate or spiral. Flowers are aggregated in axillary or terminal inflorescences, where the terminal inflorescence unit is often racemose. The flowers themselves are often bracteolate, hence small. They are regular, mostly tetramerous or pentamerous, with distinct isomerous calyx and corolla. The fruit is a usually one-seeded berry or a drupe.

No information is available on the early fossil record of this family.

3.5.1.1 New Zealand Myrsinaceae

The Flora of New Zealand (Allan, 1982) distinguishes two New Zealand genera of Myrsinaceae: *Myrsine* L. and the monotypic genus *Elingamita* G.T.S. Baylis which is endemic to the West Island of the Three Kings Islands. *Elingamita johnsonii* G. T. S. Baylis is a small, up to 3 m tall, tree of which only a few individuals are known. It is distinguished from the genus *Myrsine* mainly by its peculiar flower morphology with the corolla shorter than its calyx.

On the New Zealand mainland five *Myrsine* species are currently described: *M. australis* (A. Rich.) Allan, *M. divaricata* A. Cunn., *M. nummularia* Hook., and *M. salicina* Hew. ex Hook. f. In 1998, Heenan suggested to separate *M. argentea* from *M. divaricata* because of differences in morphology and habitat. This taxon is known from one location in the northern South Island. Other *Myrsine* species are endemic to off shore islands. *M. kermadecensis* Cheeseman grows on the Kermadec Sunday Island. *M. oliveri* Allan occurs only on the Great Island of the Three Kings group. *M. chathamica* F. Muell. and *M. coxii* Ckn. are species from the Chatham Islands.

All New Zealand taxa are shrubs or trees of the forest except *M. nummularia*, which is a prostrate shrub growing in montane or subalpine grasslands and herbfields.

At the end of the 19th century all New Zealand species were included in *Myrsine* Linné. After a revision of the genus by Mez (1902) *M. kermadecensis*, *M. australis*, and

M. salicina were placed in *Rapanea* Aubl. The remaining taxa were included in *Suttonia* A. Rich. In 1925, Cheeseman pointed out the artificiality of this treatment and placed all New Zealand species in *Suttonia.* A detailed study of Hosaka (1940) on the status of these genera showed that there was no adequate character to separate *Suttonia* and *Rapanea* and too little morphological support for a generic separation of the type specimen *Myrsine africana* L. As a consequence, Allan (1947) restored the former nomenclature and included all New Zealand taxa in *Myrsine* L. However, Oliver (1951) argued for including all New Zealand taxa in *Rapanea* because he found greater morphological differentiation between the African and the New Zealand *Myrsine* than between the latter and all species placed in *Rapanea*.

Both, *Rapanea* and *Myrsine* species are shrubs or small trees. Flowers are arranged in short axillary clusters and are often formed on the old wood below the leaves. They can be tetramerous or pentamerous, sepals and petals can be free or partly connate. Anthers are attached by the back to the sepals; the ovary is superior. Fruits are one-seeded globular drupes, which are eaten and dispersed by birds. Leaves are often gland-dotted. In New Zealand the genus *Myrsine* includes large and small-leafed forms and with *Myrsine divaricata* one shrub with a divaricating growth habit. Analyses from fossil pollen suggest that *Myrsine* has been present in New Zealand since the Oligocene (Mildenhall, 1980).

3.5.2 Phylogenetic Analyses of *Myrsine* Nuclear rDNA Sequence

Phylogenetic analyses for New Zealand Myrsinaceae in this section were based on nuclear ITS1, ITS2 and 5.8S rDNA sequence data. DNA for *Myrsine* and three Australian species of *Rapanea* (*Rapanea howhittiana* Mez, *R porosa* (F.Muell.) Mez, and *R. variabilis* (R.Br.) Mez) was extracted from dried leaf material and processed as described in Section 2.3.1 Amplification and sequencing of the rDNA region were performed as described in Section 2.3. Sequence data for *Stylogyne sp.* (Myrsinaceae), which was chosen as an outgroup, was taken from Anderberg *et al.* (2000). Sequence data were aligned using the progressive multiple alignment procedure, ClustalX (Thompson *et al.*, 1997). The herbarium voucher numbers for New Zealand and overseas taxa used in this section are also listed in Appendix II.

3.5.2.1 Split Decomposition

Split decomposition was employed to investigate the tree-like properties of the Myrsinaceae rDNA sequence data.

A prominent feature of the splitsgraph shown in Figure 3.9 is the great differences in length of external edges. This phenomenon may account for the relatively low SplitsTree fit value of 67.1.

The chosen outgroup *Stylogyne* joins with the African species *M. africana*. Basal are *Rapanea spp.* and *M. salicina* and the *M. kermadecensis* lineage. Except for *R. porosa*, all taxa, which are close to the root, are genetically diverse and are separated by long external edges. Compared to *Rapanea howhittiana, R. variabilis*, and *R. porosa* is placed on a very short external edge, which suggests that the rDNA of this taxon may evolve at a slower rate.

M. kermadecensis - *M. oliveri* represent a separate evolutionary lineage with *M. kermadecensis* directly descending from an *M. oliveri* – like ancestor. The graph structure suggests a recent dispersal event from The Three Kings group to the Kermadec islands, which are separated by more than 1000 km of Pacific Ocean.

Split decomposition supports an internal edge, which is supported by a bootstrap value of 80% (100 replicates). This split separates species close to the root from a group of taxa, which are genetically less distinct. This group includes *M. argentea, M. australis, and M. divaricata,* from mainland New Zealand and the Poor Knights Islands, respectively, *M. maximowizcii* from Japan, *M. chathamica* and *M. coxii* from the Chatham Islands, as well as *Elingamita johnsonii* from the Three Kings Islands. The extremely low genetic diversity between those taxa indicates a very close phylogenetic relationship within this group and, as will be discussed, implies recent morphological diversification and speciation.

Within this group a split supports the exclusion of *M. coxii* from other taxa in this group. The graph shows that both accessions included for this species are genetically different.

Interestingly, the Japanese species *M. maximowizcii* and the accession of *M. divaricata* originating from the Poor Knights Islands are grouped and separated by a short edge from the remaining taxa.

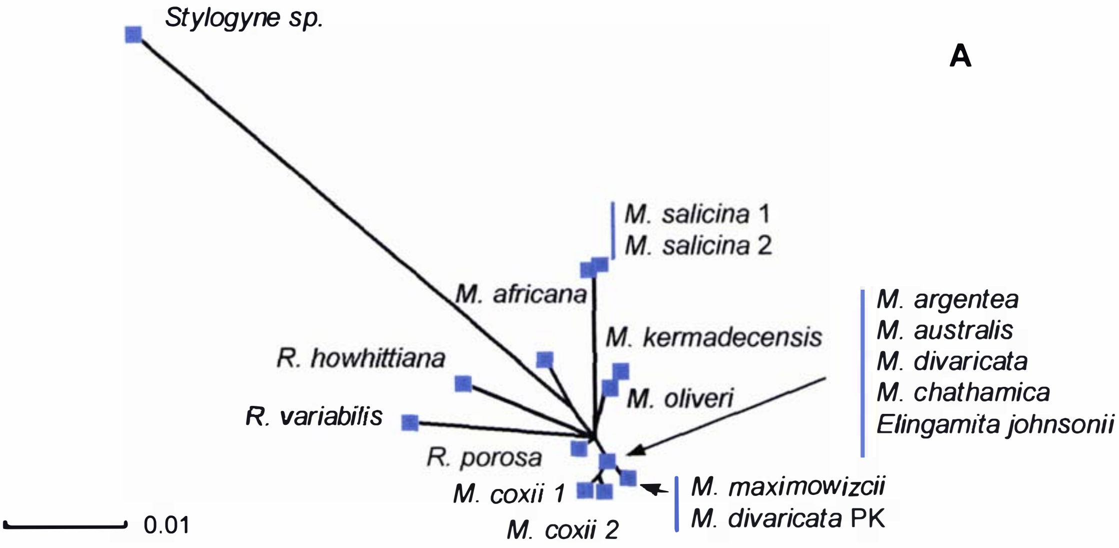


Figure 3.9: Splitsgraph from nuclear ITS and 5.8S rDNA sequence data for Myrsinaceae. *Stylogyne* sp. is used as outgroup. *Stylogyne sp.* is added as outgroup. Scale: 0.1 substitutions per site. Details: Fit = 67.1 characters = 606(of 634) gaps = 23(off) missing = 5(off) constant = 525 non-parsimony= 574 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Herbarium voucher numbers for all accessions are listed in Appendix II.

Split decomposition on a taxa set without the outgroup *Stylogyne sp.* results in a fit value of 66.5 and does not improve the representation of the data. The most problematic taxon seems to be *M. salicina*, which is placed on a comparatively long external edge. The great genetic distance suggests that the rDNA region of this species may evolve at a fast rate. A graph made without this *M. salicina* is shown in Figure 3.10. The graph has a high SplitsTree fit value of 91.0, which implies a better representation of the distances between taxa. The splitsgraph shows that a network structure is the best visualisation for phylogenetic relationships between some species of *Myrsine*. The graph also reveals contradictory patterns that join the outgroup *Stylogyne* with either *M. africana* or *Rapanea* spp.

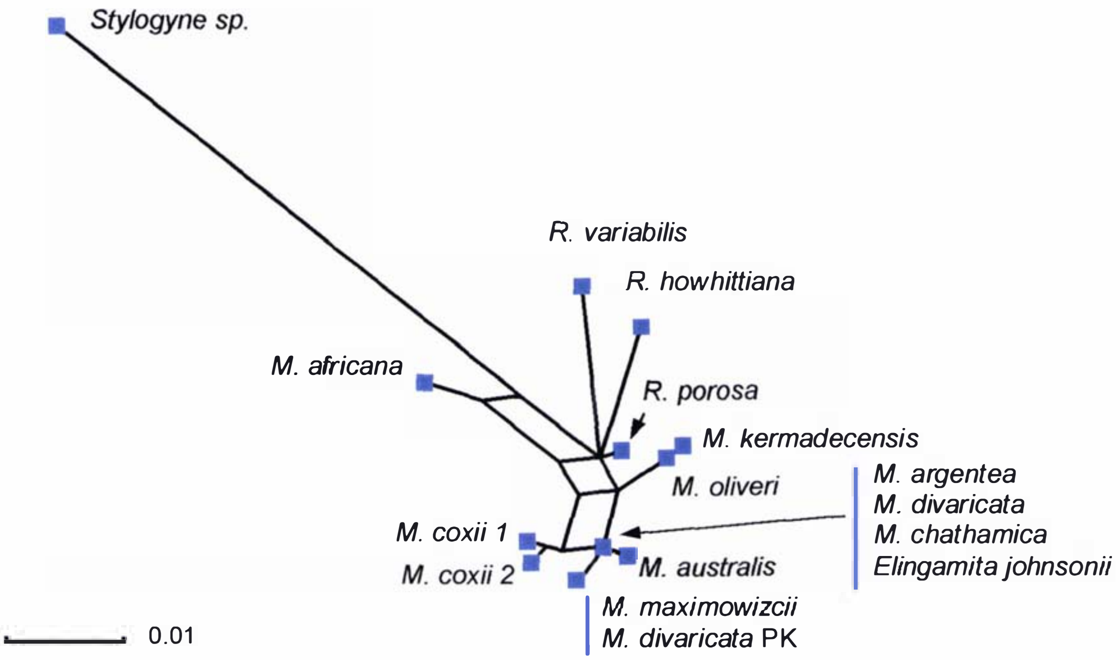


Figure 3.10: Splitsgraph from nuclear ITS and 5.8S rDNA sequence data for Myrsinaceae without *M. salicina*. *Stylogyne sp.* is added as outgroup. Scale: 0.1 substitutions per site Details: Fit=91.0 characters = 606(of 634) gaps = 23(off) missing = 5(off) constant = 532 non-parsimony= 585 -dsplits - hamming. For splitsgraph details see Section 2.4.4.4. Herbarium voucher numbers for all accessions are as listed Appendix II.

3.5.2.2 Maximum Likelihood Analyses

The phylogeny shown in Figure 3.11 was made with quartet puzzling under maximum likelihood criteria. To obtain the best possible parameter settings, model fitting analyses were performed on 11 equally parsimonious trees, which included all species available for *Myrsine*, *Elingamita*, and *Rapanea*. Likelihood ratio tests identified the GTR+I model as the best to describe the evolution of the Myrsinaceae rDNA data.

In the phylogeny, the outgroup taxon *Stylogyne* joins with the African species *Myrsine africana*. *M. africana* is basal to all taxa analysed from the Austral-Pacific region indicating that these taxa are a monophyletic group. This phylogenetic relationship is well supported by a QP value of 84 and gains a bootstrap support of 79% (taken from 100 replicates). Maximum likelihood analyses suggest that, within the Austral-Pacific cluster, the African taxon is closest to the New Zealand species *M. salicina.*

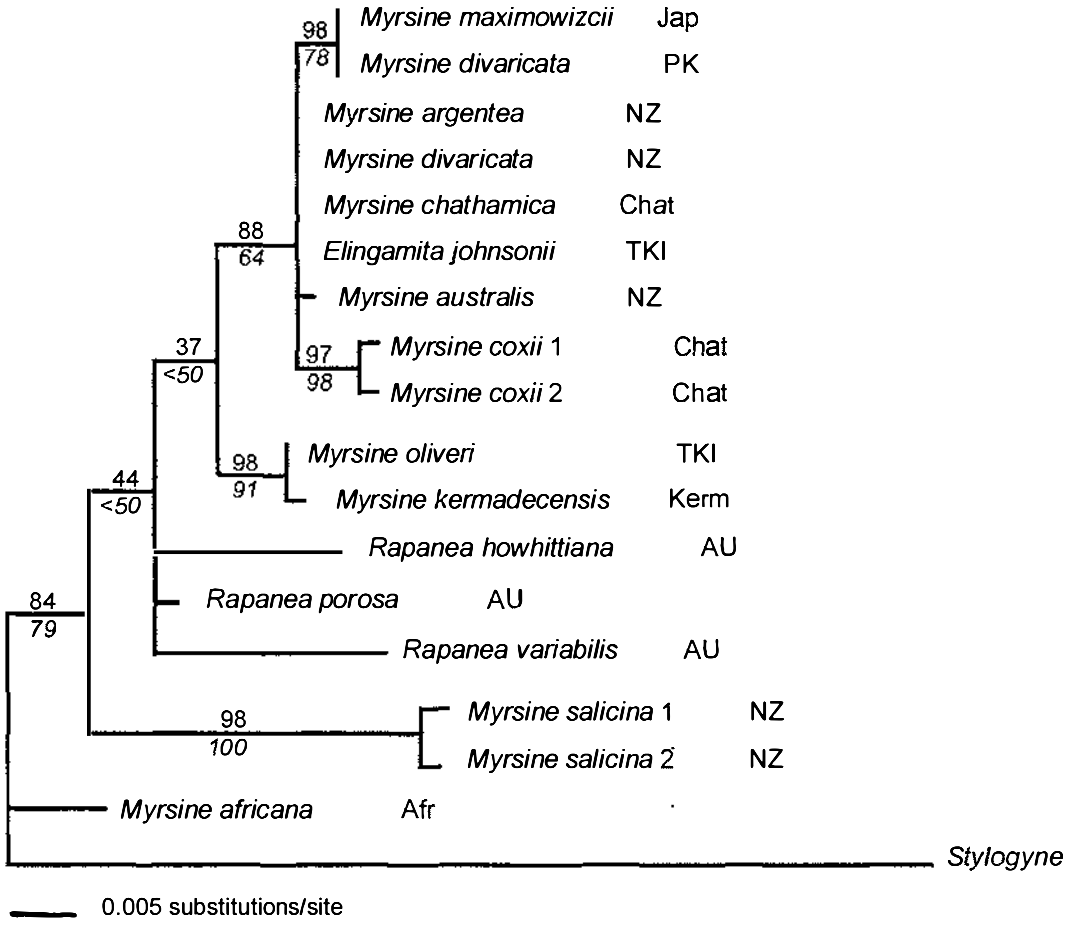


Figure 3.11: Rooted phylogeny for Myrsinaceae based on ITS 1, ITS 2 and 5.8S of nuclear rDNA , made with quartet puzzling and ML parameters under PAUP* (Swofford, 1999). QP values are assigned to the edges, bootstrap values were taken from 100 replicates and values given in italics. Tree is rooted on *Stylogyne* sp. -ln likelihood = 1406.970. Substitution model: GTR; ML settings: assumed nucleotide frequencies: empirical with A = 0.22388 C = 0.29038 G = 0.27732 T = 0.20842; number of substitution types = 6; proportion of sites assumed to be invariable = 0.76098 (observed proportion of constant sites = 0.866337). distribution of rates at variable sites = equal. Abbreviations: **AU**: Australia; **Chat**: Chatham Islands; **Jap**: Japan; **Kerm**: Kermadec Islands; **NZ**: New Zealand; **PK**: Poor Knights Islands; **TKI:** Three Kings Islands.

In the ML phylogeny, the Australian species of *Rapanea* are placed within a New Zealand species cluster. However, this relationship, which suggests a polyphyletic origin of *Myrsine* on the archipelago, gains only low QP reliability values and weak bootstrap support.

Low support is also recovered for joining the *M. kermadecensis-M. oliveri* lineage with a group of taxa from New Zealand, its off-shore islands and Japan. The close relationship between *M. kermadecensis* and *M. oliveri*, however, is well supported. This phylogenetic alliance is also recovered with split decomposition (Figures 3.9 and 3.10) and suggests a recent trans-oceanic dispersal event from the Three Kings to the Kermadec Islands.

A closely related group of taxa includes *M. argentea*, *M. australis*, *M. divaricata*, *Elingamita johnsonii*, two species from the Chatham Islands and *M. maximowizcii* from Japan. As also suggested by split decomposition (Section 3.5.2.1), the *M. divaricata* accession collected on the Poor Knights Islands, which are situated a few kilometres off-shore Northland, New Zealand, is genetically distinct from the accession collected on the New Zealand mainland. This taxon seems more closely related to the Japanese species *M. maximowizcii* than to the mainland *M. divaricata*.

3.5.2.3 Molecular Clock Tests and Substitution Rate Estimates

Molecular clock tests on *Myrsine* were done under the SplitsTree 'test_clock' option (Huson, 1998) and were performed on New Zealand taxa only. The root position was chosen between *M. salicina* and the remaining New Zealand taxa. This root position was validated by maximum likelihood analyses on a data set, which included *Stylogyne* and/or *M. africana* sequences for outgroup placement. Using this root position, molecular clock tests failed for *Elingamita johnsonii*, *M. argentea*, *M. australis*, *M. divaricata* and *M. chathamica*.

To estimate the substitution rate for the *Myrsine* ITS regions, the divergence between the two New Zealand lineage i.e. between *M. salicina* and the remaining New Zealand taxa was calculated under SplitsTree 3.1 'diverge' option. The taxa that failed the MC test as well as ambiguous/heteroplasmic sites and gap positions were excluded from the estimates. A genetic distance of DIST = 0.041 ± 0.008 [substitutions per site] was measured between the two lineages and a substitution rate of $s_r = (0.590 \pm 0.117) \times 10^{-9}$ [substitutions per site and year] was calculated for Myrsinaceae nuclear rDNA regions (ITS 1 and ITS 2). Details and results of the estimate are given in Table 3.4.

number of base pairs in original data set:	ITS 1: 261 bp ITS 2: 208 bp
molecular clock tests:	MC assumption failed for *Elingamita johnsonii*, *M. argentea*, *M. australis*, *M. divaricata* and *M. chathamica*
excluded sites:	ambiguous sites and gap positions: 12 sites invariable sites = 188 bp
distance measure:	Genetic distance between *M. salicina* and *M. coxii*, *M. divaricata* PK, *M. kermadecensis* and *M. oliveri*.
calculated distance:	DIST= 0.041 ± 0.008 [substitutions per site]
fossil date:	First appearance of *Myrsine* in the New Zealand fossil pollen record during the Miocene: 35 million years ago (70 m. y. of parallel evolution)
substitution rate	$s_r = (0.590 \pm 0.117) \times 10^{-9}$ [substitutions per site and year]

Table 3.4: Details and results of substitution rate estimate for the nuclear ITS regions of New Zealand species of *Myrsine*.

To calculate the substitution rate of the *Myrsine* ITS regions, a calibration date was taken from the New Zealand fossil pollen record, which reports the first appearance of *Myrsine* during the Oligocene. The arrival time was chosen at the beginning of this era 35 million years ago. Using this date, a substitution rate of $s_r = (0.590 \pm 0.117) \times 10^{-9}$ [substitutions per site and year] was calculated for New Zealand species of *Myrsine*.

3.5.3 Discussion: Ancient Origins and Recent Speciation?

The phylogenetic analyses of Myrsinaceae based on the ITS and 5.8S regions of nuclear rDNA provided information on the evolutionary history of this family in New Zealand and its relationship to some overseas species.

Root placement with *Stylogyne* sp. identified the African species *M. africana* as the most basal taxon within the group. Analyses under maximum likelihood criteria strongly support a monophyletic Austral-Pacific cluster that includes all species of *Rapanea*, *Myrsine* and *Elingamita* analysed in the present study. This result suggests close biogeographic relationships between Australia and the Pacific region.

Tree building recovered support for grouping the three species analysed for the Australian genus *Rapanea*. DNA sequencing revealed a substantial amount of genetic differentiation between *Rapanea* species. However, from the studies presented here, generic differentiation suggesting two distinct monophyletic groups for Australian *Rapanea* and New Zealand Myrsinaceae is not obvious: tree building methods do not split these taxa away from genus *Myrsine* - the accessions rather appear as part of this genus. The result indicates the need for further investigations and a possible revision of the taxonomic status of *Myrsine* and *Rapanea*.

Maximum likelihood and tree building analyses with quartet puzzling suggest a polyphyletic origin of the New Zealand taxa. Some patterns in the data identify the African species as being more closely related to the New Zealand *M. salicina* than to the Australian *Rapanea* group. However, low ML support values and contradictory patterns of relationship, visualised in the splitsgraph (Figure 3.9), suggest that the position of the root on *M. salicina* is only locally stable. Because of contradictory patterns in the data or different substitution patterns across the alignment, tree building methods may not recover the 'true phylogeny' from the sequence data. A Shimodaira-Hasegawa sites tests (Goldman *et al.*, 2000) on a phylogeny with a basal positioning of *Rapanea* spp., and a constraint monophyletic origin of all New Zealand Myrsinaceae, was not significantly worse. Hence, on the basis of the analyses presented in this section, a polyphyletic origin of the New Zealand taxa cannot be convincingly advocated. Geographic considerations clearly favour a closer relationship between African and Australian species suggesting *M. africana* being more closely related to *Rapanea* spp. than to *M. salicina*. To answer the question of the origin of New Zealand

Myrsine, future sequence analyses need to include a larger group of overseas taxa and compare phylogenetic results from different genome regions.

Nevertheless, the quartet puzzle analyses presented in this chapter place *M. salicina* basal to all other New Zealand taxa. Split decomposition analysis also indicates that this species is genetically very different from other New Zealand species. These results allow the interpretation that *M. salicina* is the only extant representative of a lineage that diverged early during the evolution of Myrsinaceae. The analyses suggest that this lineage is a direct ancestor of the most ancient members of the genus *Myrsine* in New Zealand, which, according to fossil evidence, has been present in New Zealand since the Oligocene (Mildenhall, 1980).

Furthermore, tree building analyses show that *M. oliveri* and *M. kermadecensis* represent a separate lineage, which also may have diverged at an early stage during the evolution of this genus. The genetic distance between this lineage from the root position suggests its divergence shortly after the first appearance of this genus in New Zealand. The close relationship between *M. oliveri* and *M. kermadecensis*, however, clearly suggests a recent trans-oceanic dispersal event from the Three Kings Islands to the Kermadec Islands.

Both split decomposition and maximum likelihood suggest, that the Japanese taxon *M. maximowizcii* and the Chatham Island species *M. chathamica* and *M. coxii* are closely related to the New Zealand mainland species. This suggests recent trans-oceanic dispersal events from mainland New Zealand to Japan and the Chatham Islands. With the Chatham islands being located east of New Zealand, dispersal could have been mediated by the west wind drift which was described as a major mechanism to carry seeds in eastward directions (Mildenhall, 1980; Wright *et al.*, 2000). However, the seeds of *Myrsine* are eaten by birds and are too heavy to travel distances by air. Thus, in the case of *Myrsine,* dispersal across the ocean by migrating birds is much more likely. This scenario would also explain the recent dispersal in northerly direction of a New Zealand ancestor of *M. maximowizcii* to Japan.

Regarding leaf morphology and growth form of all New Zealand *Myrsine* species, *M. salicina* is well distinct from other taxa. The narrow leathery leafs and the tree-like appearance is not found in any other New Zealand taxon. Large ovate leathery leaves combined with a bushy growth form are shared by *M. oliveri*, *M. kermadecensis*, *M. maximowizcii*, *M. chathamica* and *Elingamita johnsonii*. Interestingly, split decomposition suggests that these taxa are in two different phylogenetic groups (Figures 3.9 and 3.10). While *M. oliveri* and *M. kermadecensis* form an early diverged, monophyletic group, *M. chathamica* and *Elingamita johnsonii* have close affinities to species with small leathery leafs like *M. argentea*, *M. divaricata* and *M. coxii*.

The close phylogenetic relationship between Myrsinaceae from New Zealand and its off-shore islands indicates a recent event of species diversification. The low genetic diversity thesis taxa is visualised by split decomposition and quartet puzzling (Figures 3.9 - 3.11) and suggests that these taxa only have speciated during the Quaternary. This conclusion would be consistent with studies on the New Zealand alpine flora that have identified the Quaternary as a major era of speciation (Winkworth *et al.*, 2000; Wagstaff and Garnock-Jones, 1998). The relationships between these closely related taxa was further investigated and is presented in Chapter 5 of this thesis.

3.6 NOTHOFAGACEAE KUPRIANOVA 1962

3.6.1 Introduction

For a long time *Nothofagus* Blume, the southern beech, was placed within the family Fagaceae until sufficient evidence accumulated to support the formation of the monogeneric family Nothofagaceae (Martin and Dowd, 1993). Cronquist (1998) combines the families Fagaceae and Nothofagaceae in the order of Myricales - the order is commonly called Fagales (APG, 1998) and is placed in the subclass Hamameliopsida.

Palynological studies recognise four different pollen types in living *Nothofagus* species (Dettman *et al.*, 1990). Hill and Read (1991) conducted a comparative morphological study of cupules, leaf architecture and leaf epidermis and they concluded that the cupule morphology supports a division of the genus along the pollen types. The authors proposed an infrageneric classification of the genus into four subgenera. These are: *Nothofagus,* with five species in South America, *Fuscaspora* with five species in New Zealand, South America and Tasmania, *Brassospora* with 19 species in New Guinea and New Caledonia and *Menziesospora* with six species in New Zealand, Australia and South America. In 1992, Hill introduced '*Lophozonia*' as a new name for subgenus *Menziesospora*.

Subsequently, molecular phylogenetics on *Nothofagus rbc*L (Martin and Dowd, 1993) and nuclear ITS sequence data (Manos, 1997) were conducted. These studies supported the infrageneric classification based on morphological evidence and substantiated a monophyletic origin of the four different pollen groups. Phylogenies based on these molecular data suggest a hierarchical arrangement of the four monophyletic groups of *Nothofagus*: (*Lophozonia* (*Fuscaspora* (*Brassospora*-*Nothofagus*))) in which *Lophozonia* was suggested as the most ancient group (Manos, 1997).

All *Nothofagus* species are monoecious. The unisexual flowers often arise in the axils of bud scales and leaves. Male flowers have 8-36 stamens and grow in spikes with a perianth of 4-6 imbricate lobes. Female flowers sit on arrested spikes surrounded by an involucre (cupule) which is hardening when the fruit, a 2-3 celled nut, develops. Seeds are propagated by gravity only. Simple, stipulate leaves are alternately arranged. The genus includes deciduous and evergreen trees and shrubs.

The fossil record of *Nothofagus* dates back to the Late Cretaceous, over 70 million years ago. The oldest known fossils are *fusca*-, *menziesii*-, and *brassii*-type pollen excavated from the Late Campanian of the Antarctic Peninsula and from the Maastrichtian of southern South America (Dettmann *et al.*, 1990). The sequence of appearance is not apparent. Older pollen fossils (83 million years) which are not among extant types have been assigned to *Nothofagus*, two extinct pollen types have been described.

3.6.1.1 Hypotheses on the Phylogeography of *Nothofagus*

Traditionally, a close phylogenetic relationship was assumed between the northern beeches (*Fagus* spp.) and the southern beeches (*Nothofagus* spp.). On the basis of this assumptions and motivated by the idea that continental drift occurred too early to affect *Nothofagus*, Darlington (1965) argued for a northern hemisphere origin of the genus. However, the lack of *Nothofagus* fossils in the Northern Hemisphere and the diversity of fossil pollen found in the Southern Hemisphere were interpreted to suggest a southern origin of the genus (Cranwell, 1963; Moore, 1971). When results from molecular studies showed that *Nothofagus* forms an independent evolutionary lineage (Martin and Dowd, 1993) a southern hemisphere origin of the genus gained more and more acceptance.

Today, *Nothofagus* is regarded a pivotal genus for studying evolution and biogeography of the southern hemisphere flora. Southern beeches have been considered to be poor long distance dispersers since their nuts are assumed not viable in sea water and not likely to be eaten and dispersed by birds (Preest, 1963; Burrows and Lord, 1993). Thus, historically, a continuous landmass has been thought critical for range expansion of *Nothofagus*. As fossil evidence showed that all subgenera had evolved before the break up of the former super continent (Dettmann *et al.*, 1990), it has been proposed that the present day distribution of *Nothofagus* is the result of continental drift, following the break-up of Gondwana (Raven & Axelrod, 1972). Linder and Crisp (1995) reasoned that "..since *Nothofagus* was widely distributed in Eastern Gondwana prior to the continental break-up, it appears likely that the relationships of the species between the different Gondwanan fragments should reflect the break-up sequence..". These authors strongly advocated a hypothesis of vicariance and extinction to explain extant phylogeographic patterns.

In contrast, Manos (1997) argued that a vicariant evolution of *Brassospora* and *Nothofagus* beeches is doubtful because fossil evidence indicated widespread extinction of these subgenera from South America, Antarctica, New Zealand and Australia and, respectively, from Antarctica, Tasmania and New Zealand. Pollen studies have also been interpreted to suggest that rare long distance dispersal has occurred across the Tasman Sea from Australia to New Zealand (Mildenhall, 1980; Macphail *et al.*, 1994). Trans-Tasman dispersal has been inferred from the observation that species of all four extant subgenera are recognised in the Australian Cenozoic fossil pollen record which appear only many millions of years later in New Zealand strata. Furthermore, Pole (1994) advocated the possibility of trans-oceanic re-dispersal of *Lophozonia* beeches into New Zealand. He argued that this subgenus might have been extinct on the islands for a certain period because its pollen is absent in some Early Tertiary sediments.

3.6.1.2 New Zealand Nothofagaceae

New Zealand hosts four living members from two *Nothofagus* subgenera. *N. menziesii* (Hook. f.) Oerst. (silver beech) is the only extant member of subgenus *Lophozonia*. *N. fusca* (Hook. f.) Oerst., *N. solandri* (Hook. f.) Oerst., and *N. truncata* (Col.) Ckn. are *Fuscaspora* beeches. It has been suggested that hybridisation is common between these three species (Cockayne and Atkinson, 1926; Wardle, 1988b).

All species are tall trees with evergreen foliage. *Nothofagus* forests cover large areas in the New Zealand South Island, and some mountainous regions in the North Island. They often form dense monotypic stands with little undergrowth. Beeches are absent from all off shore island and show substantial distribution disjunctions on the mainland.

The rare appearance of *Lophozonia* pollen has also been reported from Late Cretaceous and Early Tertiary sediments in New Zealand (Mildenhall, 1980). However, McGlone *et al.* (1996) dispute this observation and propose that *Lophozonia* pollen first appears in Early Eocene sediments. These authors suggest that pollen types that are present in ancient geological strata originate from *Fuscaspora* beeches and from ancestral *Nothofagus* species. Mildenhall (1980), however, proposed that the *Fuscaspora* pollen type first appeared during the Eocene – at the same time with *Brassii*-type pollen. McGlone *et al.* (1996) finds the first appearance of *Brassii* pollen during the Early Oligocene. *Brassospora* beeches are now confined to the tropical Indo-Pacific region. Palynological evidence suggests, that this subgenus has become extinct in New Zealand during the Late Tertiary (Hill, 1992).

Pollen resembling *Nothofagus menziesii* pollen has been found throughout Miocene sediments in Central Otago (Mildenhall and Pocknall, 1989). However, the authors suggested, that this pollen originated from a different species, which occupied more temperate sites.

Macrofossils of *Nothofagus* have been described from Late Cretaceous (McGlone *et al.*, 1996) and from Miocene sediments in New Zealand (Pole, 1993). These fossils depict large-leafed *Nothofagus* specimens that may have been deciduous. All fossils have been assigned to extinct species.

3.6.2 Phylogenetic Analyses of *Nothofagus* Species

Phylogenetic analyses were performed on 14 accession from the two *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. A 622 bp alignment containing the ITS1, ITS2 and 5.8S of nuclear rDNA of *Lophozonia* and *Fuscaspora* was used for analyses on nuclear DNA sequence data. Analyses on plastid DNA sequence data were performed on a 937 bp alignment that contained the *trn*L-*trn*F intergenic spacer region and the *trn*L intron.

Nuclear and plastid sequence data used in this section were determined and analysed for the same individuals. Herbarium vouchers of these are listed in Appendix II. DNA extractions, PCR and sequencing reaction used protocols shown in Section 2.3.1 and 2.3.3 respectively.

To identify a suitable outgroup for both data sets, a blast search was performed using the BLAST option on the NCBI webpage (http://www.NCBI.nlm.nih.gov/). Taxa sequence showing the highest identity scores were chosen and used for analyses. For the nuclear rDNA regions, outgroup placement was studied with *Corylus chinensis* Franch. and *Carpinus hupeana* Hu (Betulaceae). These sequence data, published by Erdogan and Mehlenbacher (2000) and Chen *et al.* (1999) respectively, were obtained from the NCBI GenBank. Outgroup placement on the plastid data set was studied with *Quercus calliprinos* Webb and *Q. suber* Vázquez Pardo and *Q. ilex* Vázquez Pardo (Fagaceae). These sequences were taken from Paffetti *et al.* (unpublished).

3.6.2.1 Split Decomposition on Nuclear rDNA Sequence

The nuclear DNA alignment was analysed under split decomposition to reveal tree-like properties of the data. The splitsgraph shown in Figure 3.12 was made on 14 accession from the two *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. The graph has a high fit value of 90.3, which indicates that the pairwise distances in the distance matrix are well represented by the graph. A low level of contradiction in the data is visualised by a network structure in the splitsgraph. The genus *Nothofagus* is divided by a long internal edge. The split, which separates subgenus *Fuscaspora* from subgenus *Lophozonia*, is well supported by a bootstrap support of 100. Within the subgenera, species from one continent are more closely related to each other than to overseas species.

The root was determined by outgroup analyses with *Corylus chinensis* and *Carpinus hupeana* and joins the phylogeny between both subgenera, positioned closer to *Lophozonia*. A splitsgraph on a data set including the outgroup sequences receives a fit value of 82.6 and does not alter the relationships between ingroup taxa (result not shown).

Although split decomposition shows, that there is a small amount of contradictory patterns in the data, the good phylogenetic resolution displayed in the graph, the high SplitsTree fit value and the high bootstrap support for prominent patterns suggest that the data retains mainly treelike properties.

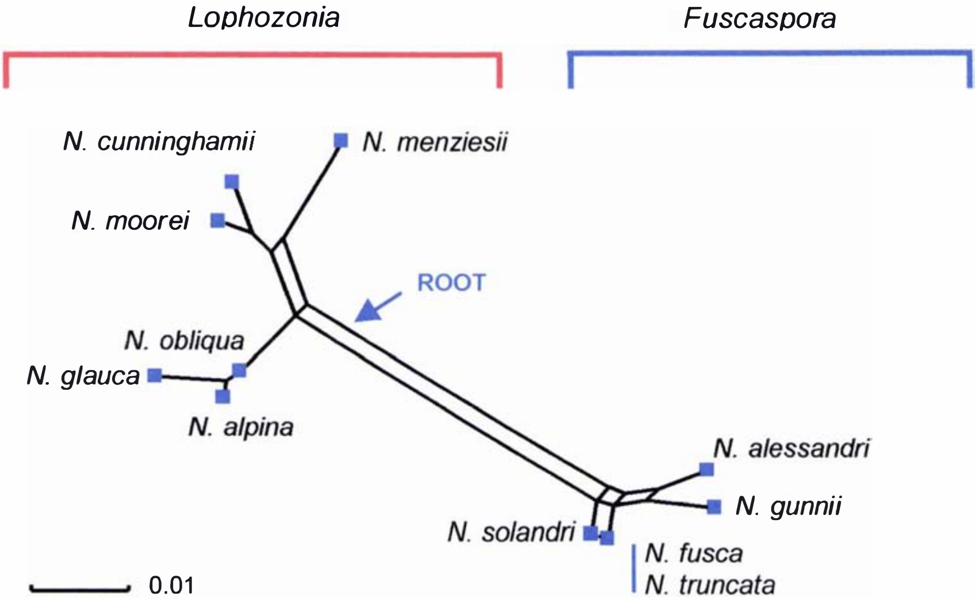


Figure 3.12: Splitsgraph for *Nothofagus* subgroups *Lophozonia* and *Fuscaspora* based on observed distances calculated from nuclear rDNA sequence. The position of the root was determined by outgroup placement with *Corylus chinensis* and *Carpinus hupeana* and is indicated by a blue arrow. Scale = 0.1 substitutions per site. Details: Fit = 90.3 characters = 569(of 621) gaps = 32(off) miss = 20(off) constant = 507 non-parsimony= 518 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

3.6.2.2 Maximum Likelihood Analyses on Nuclear rDNA Sequence

The phylogeny shown in Figure 3.13 was made with quartet puzzling using maximum likelihood criteria. Model fitting analyses identified a GTR invariable sites model as the best description of the evolution of the sequence data. Maximum likelihood parameter settings were estimated by likelihood ratio test comparing six equally parsimonious unweighted trees.

The internal edge, which separates the two subfamilies, is supported by high QP reliability and bootstrap values. Species from one continent are clustered together. Australian *Lophozonia* beeches are more closely related to the New Zealand taxon than to the South American species. In contrast, Australian *Fuscaspora* beeches have greater affiliation to the South American taxon.

As visualised by the length of the external edges, there is greater genetic differentiation between species of *Lophozonia* beeches than between *Fuscaspora* beeches, which implies a more recent radiation of species with the latter group.

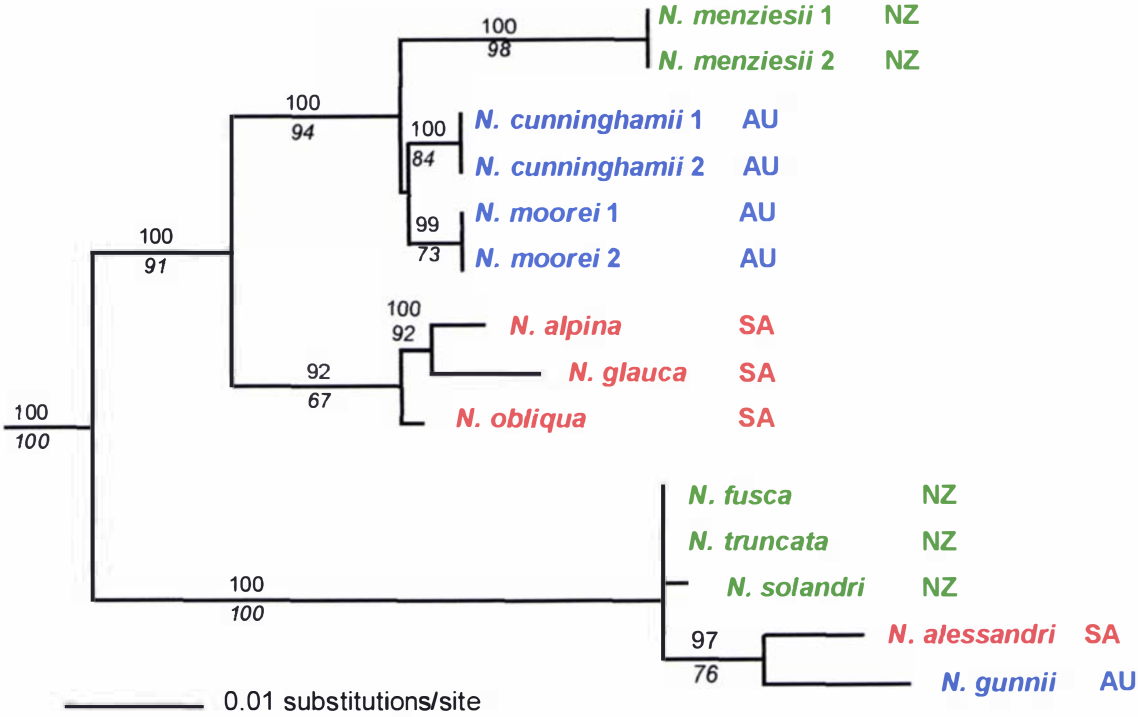


Figure 3.13: Phylogeny for *Nothofagus* subgroups *Lophozonia* and *Fuscaspora* based on nuclear rDNA and made with quartet puzzling and ML parameters under PAUP* (Swofford, 1999). Quartet puzzle and bootstrap values (italics) are assigned to the edges. -ln likelihood = 1708.14555. Parameter model: GTR+I; ML settings: assumed nucleotide frequencies: empirical with A = 0.22388 C = 0.29038 G = 0.27732 T = 0.20842; number of substitution types = 6; proportion of sites assumed to be invariable = 0.128244 (observed proportion of constant sites = 0.744227); distribution of rates at variable sites = equal. Abbreviations: **AU**: Australia, **NZ**: New Zealand, **SA**: South America.

3.6.2.3 Split Decomposition on Plastid Sequence Data

The splitsgraph shown in Figure 3.14 is based on DNA sequence of the *trn*L-*trn*F intergenic spacer region and the *trn*L intron of plastid DNA. The graph reveals treelike properties for these data. The high fit value (100.0) emphasises that the patterns in the sequence data are well represented by the splitsgraph. The split between both *Nothofagus* subgenera examined is supported by a bootstrap value of 100%.

As represented by the length of internal and external edges, the divergence between both *Nothofagus* subgenera is much higher than the divergence between species within one subgenus. No resolution is obtained between the New Zealand species *N. fusca* and *N. truncata* (*Fuscaspora*) and between the South American species *N. alpina*, *N. glauca* and *N. obliqua* (*Lophozonia*) suggesting close phylogenetic relationships between those taxa.

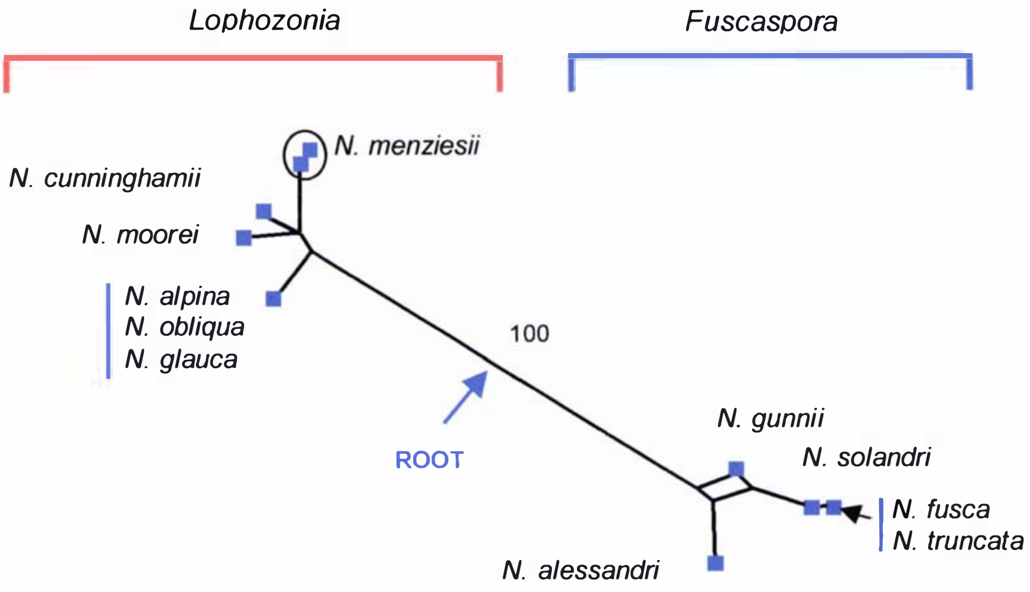


Figure 3.14: Splitsgraph for *Nothofagus* subgroups *Lophozonia* and *Fuscaspora* based on observed distances from plastid sequence data. Bootstrap value shown on the figure was taken from 100 replicates. The position of the root was determined by outgroup placement with *Quercus calliprinos* and is indicated by a blue arrow. Scale = 0.1 substitutions per site. Details: Fit = 100.0, characters = 845(of 937) gaps = 92(off) constant = 799 non-parsimony= 805 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

The only non tree-like aspect of the graph is the phylogenetic relationships between *Fuscaspora* species. Here conflicting signals are visualised as a network structure.

The root, determined by outgroup placement with *Quercus calliprinos*, joins the ingroup on the long edge between both subgenera. Its position is equidistant from both subgenera. Split decomposition recovers some support for joining the outgroup with Fuscaspora beeches visualised as a box-like structure in the graph (result not shown). A graph including the chosen outgroup taxa has a fit value of 95.2.

3.6.2.4 Maximum Likelihood Analyses on Plastid Sequence Data

Bifurcating evolutionary tree analysis models were also used to investigate the same plastid data set. A likelihood ratio test was employed to evaluate the substitution model and ML parameters that best described evolution of the *trn*L data of *Nothofagus*. A GTR invariable sites model was estimated by the tests; these settings were used for quartet puzzling analyses to construct the tree shown in Figure 3.15.

Maximum likelihood analyses on the plastid data set also identify the two *Nothofagus* subgenera as distinct phylogenetic groups. Within both subgenera, species from New Zealand and Australia are clustered; the South American taxa appear basal. This phylogeny is consistent with morphological analyses that suggest a closer relationship between the New Zealand and Australian *Nothofagus* species. All branching patterns within the *trn*L phylogeny are well supported by high QP reliability and bootstrap values.

Using the same plastid data set, tree building under parsimony and maximum likelihood criteria with alternative substitution models and parameter settings recover highly congruent tree shapes. These results, together with those from split decomposition give confidence that the plastid data set is suitable for making phylogenetic inferences.

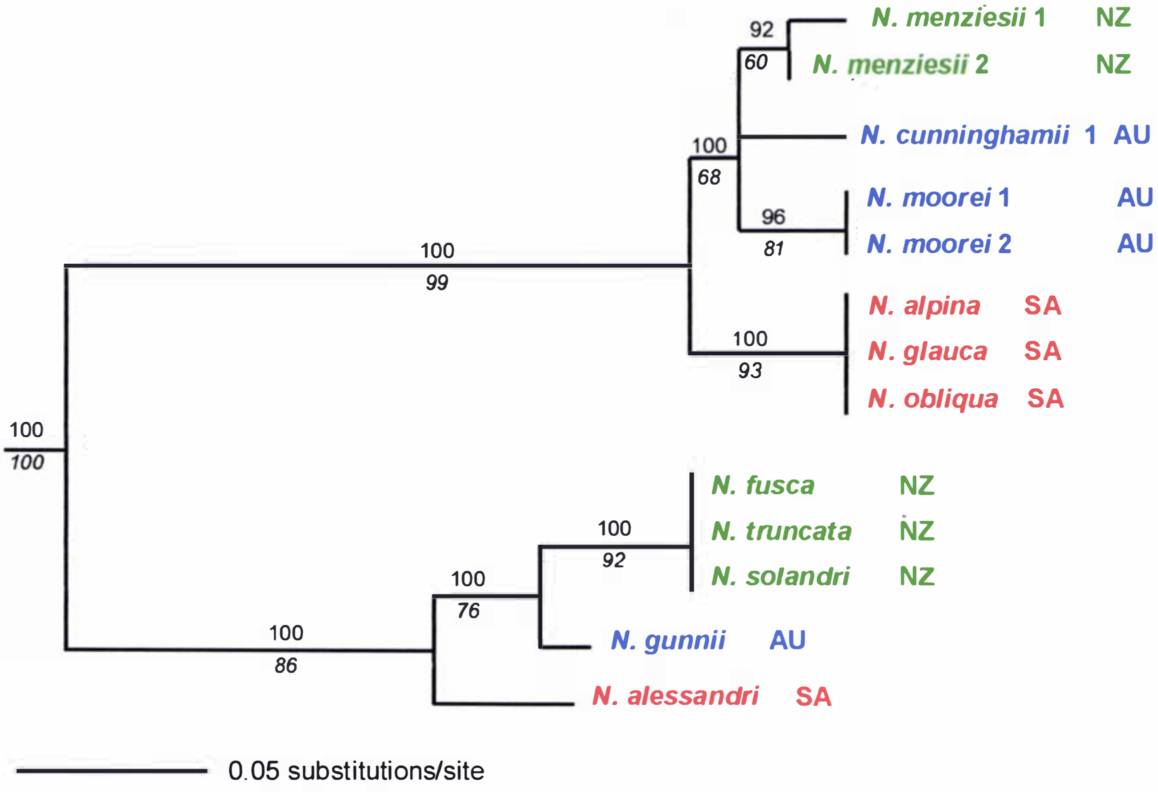


Figure 3.15: Phylogeny for *Nothofagus* subgroups *Lophozonia* and *Fuscaspora* based on *trn*L intron and *trn*L-*trn*F and made using quartet puzzling and ML parameters under PAUP* (Swofford, 1999). Quartet puzzle and bootstrap values (italics) are assigned to the edges. -ln Likelihood = 1593.40216. Parameter model: GTR; ML settings: assumed nucleotide frequencies: empirical with A = 0.22388 C = 0.29038 G = 0.27732 T = 0.20842; number of substitution types = 6; proportion of sites assumed to be invariable = 0.394388; distribution of rates at variable sites = equal. Abbreviations: **AU**: Australia, **NZ**: New Zealand, **SA**: South America.

3.6.2.5 Molecular Clock Tests and Substitution Rate Estimates

Molecular clock tests were performed on both the nuclear and the plastid data set of *Nothofagus.* Analyses were performed on an alignment of the ITS 1 and ITS 2 sequence data as well as on a combined data set from the *trn*L intron and the *trn*L-*trn*F intergenic spacer region. The analyses showed that, in both data sets, sequences of all taxa evolve under molecular clock assumptions. This result allowed sequence substitution rates to be estimated.

To estimate the substitution rate of the DNA regions, the divergence between both subgenera *Lophozonia* and *Fuscaspora* was calculated under the SplitsTree 3.1 'diverge' option. For calculating the divergence between both groups, identical sequences were removed, and gap positions, ambiguous sites and an estimated proportion of invariable sites were excluded from the data matrix. Rate calibration used fossil evidence which suggests a divergence time of ca. 75 million years between the two pollen groups *Lophozonia* and *Fuscaspora* (Dettmann *et al.*, 1990).

Using the divergence between both subgenera and a genetic distance of DIST = 0.083 ± 0.011 [substitutions per site], a substitution rate of s_r = (0.553 ± 0.076) x 10^{-9} [substitutions per site per year] was calculated for the nuclear ITS regions. Details and results of the estimate are given in Table 3.5.

With a genetic distance of DIST = 0.035 ± 0.006 [substitutions per site] a substitution rate of s_r = (0.235 ± 0.039) x 10^{-9} [substitutions per site per year] was estimated for the *trn*L intron and *trn*L-*trn*F intergenic spacer regions. Details and results of the estimate are given in Table 3.6.

A. nuclear data: internal transcribed spacer regions ITS1 and ITS2 of nuclear rDNA

number of base pairs:	ITS1: 231 bp ITS2: 226 bp
molecular clock tests:	MC assumptions were fulfilled for all taxa.
excluded sites:	ambiguous sites and gap positions: 53 sites invariable sites = 180 bp
distance measure:	Genetic distance between *Lophozonia* and *Fuscaspora*
calculated distance:	DIST= 0.083 ± 0.011 [substitutions per site]
fossil date:	Divergence between Fuscaspora and Lophozonia in the Cretaceous: 75 million years (150 million years parallel evolution)
substitution rate	s_r = (0.553 ± 0.076) x 10^{-9} [substitutions per site per year]

Table 3.5: Details and results of substitution rate estimate for the nuclear ITS regions of *Nothofagus*.

B. plastid data: *trn*L-*trn*F intergenic spacer and *trn*L intron:

number of base pairs:	*trn*L-*trn*F intergenic spacer:476 bp *trn*L intron: 440
molecular clock tests:	MC assumptions were fulfilled for all taxa.
excluded sites:	gap positions: 91 sites invariable sites = 52 bp
distance measure:	Genetic distance between *Lophozonia* and *Fuscaspora*
calculated distance:	DIST= 0.035 ± 0.006 [substitutions per site]
fossil date:	Divergence between Fuscaspora and Lophozonia in the Cretaceous: 75 million years (150 million years parallel evolution)
substitution rate	s_r = (0.235 ± 0.039) x 10^{-9} [substitutions per site per year]

Table 3.6: Details and results of substitution rate estimate for the *trn*L intron and *trn*L-*trn*F intergenic spacer regions of *Nothofagus*.

3.6.2.6 Dating the Divergence Times for New Zealand *Nothofagus* Species

Using the substitution rates estimated for nuclear and plastid sequence data in the previous section, divergence times were estimated between Australian and New Zealand *Nothofagus* species. The results are summarised in Table 3.7 and Figure 3.16.

species	molecular marker	genetic distance	divergence time
N. menziesii - *N. cunninghamii*/*N. moorei*	nuclear ITS	0.0293 ± 0.0077	**39-17 m. y.**
	*trn*L intron/*trn*L-*trn*F spacer	0.0071 ± 0.0025	**25-8 m. y.**
N. fusca - *N. gunnii*	nuclear ITS	0.0203 ± 0.0067	**28-11 m. y.**
	*trn*L intron/*trn*L-*trn*F spacer	0.0047 ± 0.0024	**18-4 m. y.**

Table 3.7: Divergence time estimates between Australian and New Zealand *Lophozonia* and *Fuscaspora* beeches derived from two molecular markers.

Results for both data sets are consistent in suggesting that New Zealand *Nothofagus* species have been separated for less than the period of time elapsed since New Zealand and Australia were joined as part of Gondwana. Assuming that the Australian species are ancestral, the divergence time estimates suggest that extant lineages in New Zealand post-date the Miocene. Both data sets suggest that, in New Zealand, *Fuscaspora* beeches are of more recent origin than *Lophozonia* beeches. The results are also consistent with the hypothesis of two independent trans-Tasman dispersal events.

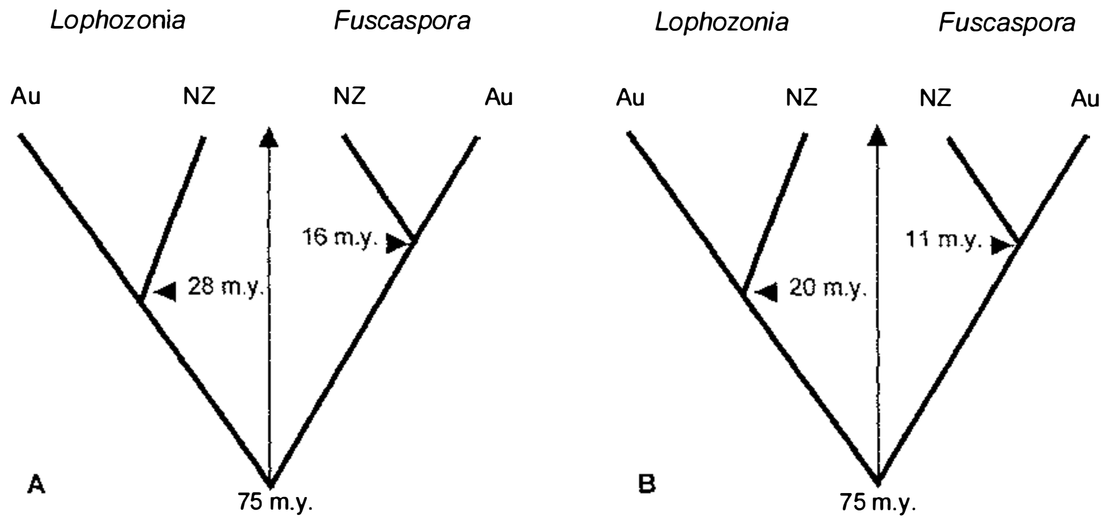


Figure 3.16: Scheme for the divergence between Australian and New Zealand *Lohpozonia* and *Fuscaspora* beeches. The divergence between the two pollen groups is assumed to be 75 million years (m.y.); mean values were used to label the additional nodes and taken from the estimates listed in Table 3.7. **A:** estimates based on nuclear ITS sequence data and a substitution rate of $s_r = (0.553 \pm 0.076) \times 10^{-9}$ substitutions per site per year. **B:** estimates based on sequence data from plastid *trn*L intron and *trn*L-*trn*F spacer regions and a substitution rate of $s_r = (0.235 \pm 0.039) \times 10^{-9}$ substitutions per site and year.

3.6.3 Discussion: Is *Nothofagus* Dispersed or Vicariant?

Results from DNA analyses on two the subgenera *Lophozonia* and *Fuscaspora* presented in this section recover support for the infrageneric classification of *Nothofagus* as suggested by Hill and Read (1991). This result confirms other molecular studies published by Martin and Dowd (1993), Manos (1997) and Setoguchi *et al.* (1997). These authors agree in placing subgenus *Lophozonia* basal in the evolutionary hierarchy. Results from split decomposition and quartet puzzling shown here indicate that genetic distances within *Lophozonia* are noticeably greater than within Fuscaspora. This result may also be an indication for greater antiquity of *Lophozonia* beeches and as such are consistent with the basal positioning of this subgenus.

Internal structures of the DNA phylogenies published to date vary only slightly, indicating good correspondence between the different DNA markers used. A prominent feature of phylogenetic trees for *Lophozonia* and *Fuscaspora* beeches is that South American taxa are placed basal to Australian and New Zealand species. This result is most consistent with the hypothesis that extant species have derived from a South American ancestor. Furthermore, most phylogenies suggest affiliation between Australian and New Zealand species with the latter being derived from an Australian ancestor. Palynology provides evidence that some *Nothofagus* species appear first in Australia and later in New Zealand (Mildenhall, 1980; Macphail, 1997). The evolutionary chronology is also recovered by analyses of plastid sequences presented in this chapter.

Results from rDNA sequence analyses, on the other hand, may suggest closer phylogenetic relationship between New Zealand and South American *Fuscaspora* beeches (Figure 3.12 and Table 3.8 below). Here, genetic differentiation - the mean hamming distance - between New Zealand and Australian beeches is higher than between New Zealand and South American beeches. As the majority of data sets analysed for *Nothofagus* do not correspond with this result, it is anomalous and unexplained.

subgenus	distance* from Australian species	distance* from South American species
Lophozonia	0.02540	0.04157
Fuscaspora	0.02079	0.01617

Table 3.8: *Mean hamming distance between New Zealand and overseas species of *Nothofagus*.

Divergence time estimates between *Lophozonia* and *Fuscaspora* beech presented in this section were based on substitution rate estimates from nuclear ITS and *trn*L intron/*trn*L - *trn*F spacer regions. Although values from nuclear ITS were higher than values derived from the plastid data set, estimates were congruent and within the range of the standard deviation. Both markers agreed that *Fuscaspora* beeches had been in New Zealand for a shorter period of time than had *Lophozonia* beeches. The upper time limit for the presence of *Lophozonia* in New Zealand has been estimated to be 39 million years. This result is more consistent with palynological evidence published by McGlone *et al.* (1996) who suggested that *Lophozonia* beeches have been present in New Zealand only since the Early Eocene. On the basis of molecular results the presence of *Lophozonia* beeches during the Cretaceous and Early Tertiary as suggested by Mildenhall (1980) can only be explained by extinction and redispersal of this subgenus in New Zealand.

Molecular data suggests that *Fuscaspora* beeches are unlikely to have been in New Zealand for more than 28 million years. This date correlates with palynological evidence presented by Mildenhall (1980) who suggested that *Fuscaspora* beeches have been present in New Zealand since the Eocene. Molecular and palynological studies reject the hypothesis of a relictual Gondwanan distribution of all New Zealand southern beeches but suggest trans-Tasman dispersal events of *Nothofagus* from Australia to New Zealand during the late Tertiary. This result is also consistent with area cladogram analysis performed on members of all four *Nothofagus* subgenera (Swenson *et al.*, 2001).

As studies by Preest (1963) clearly showed, *Nothofagus* seeds are not viable in seawater and very unlikely to be eaten and dispersed by birds. Thus, hypotheses of post Gondwanan dispersal, leaves open the question for possibilities of trans-oceanic dispersal of *Nothofagus*.

Analysing nuclear ITS and *rbc*L sequence Manos (1997) also noticed extremely low genetic differentiation between geographically distant species. He suggested that the taxa of interest might have extremely low rates of genetic evolution. Indeed, when compared to estimates made by other authors, substitution rates derived for *Lophozonia* and *Fuscaspora* nuclear ITS and plastid *trn*L intron/*trn*L - *trn*F spacer regions (Section 3.6.2.5) are below average (Wolfe *et al.*, 1987; Savard *et al.*, 1993). However, as MC tests suggested a clock like evolution of plastid and nuclear DNA regions, divergence time estimates are still expected to be valid in spite of a slow overall rate.

3.7 PROTEACEAE JUSS., NOM. CONS. 1789

3.7.1 Introduction

The Proteaceae are a large family within the flowering plants systematically organised within the order of Proteales (Cronquist, 1998). Described are about 1050 species in 75 genera.

Johnson and Briggs (1975) published a classification of the family, which was updated using a molecular systematic approach (Hoot and Douglas, 1998). The following subfamilies were defined: Persoonioideae, Bellendenioideae, Eidotheoideae, Proteoideae, Sphalmioideae, Carnavonioideae, and Grevilleoideae. Some well-known genera are *Protea* L. (Proteoideae), *Grevillea* R. Br. ex Knight, nom. cons., *Banksia* L. f., nom. cons. and *Macadamia* F. Muell. (Grevilleoideae).

The family has its distribution predominantly in the temperate and tropical climates of the Southern Hemisphere. Species prefer warm temperatures and mainly require a long dry season. Biogeographic and phylogenetic analyses suggest a Gondwanan origin of the family (Johnson and Briggs, 1975; Hoot and Douglas, 1998).

The Proteaceae include herbaceous as well as woody plant species. Leaves are usually simple, exstipulate, and mostly alternate. The flowers are often in dense cone-like inflorescences or heads with involucral bracts. The perianth is uniseriate and usually tetramerous. The androecium contains nearly always four basifixed stamens, each one directly opposite or attached to a tepal lobe. The gynoecium is monomerous and consists of a single, simple carpel with a long style and a single stigma.

Fossil grains of proteaceous type have been extracted from Tertiary deposits in Australia (Cookson, 1950). Fossil fruiting cones, found in Tertiary strata in Victoria, were assigned to taxa that reveal great similarity to the extent genus *Banksia,* which is a prominent component of the modern Australian flora (Cookson and Duigan, 1950.

The validity of Proteaceous fossils reported from northern hemisphere countries has been strongly questioned (Burbridge, 1960) and a southern hemisphere origin of the family is widely accepted.

3.7.1.1 New Zealand Proteaceae

New Zealand has two species of Proteaceae: *Knightia excelsa* R. Brown an endemic species from subfamily Grevilleoideae and *Toronia toru* A. Cunn. I.A.S. Johnson & B.G. Briggs the sole member of a New Zealand endemic genus which belongs to subfamily Persoonioideae.

Until recently, *Toronia toru* (toru), was included in the Australian genus *Persoonia* (Allan, 1982). The species is an erect small branching tree, which can reach 12 m in height. Its leaves are long and narrow (linear-lanceolate) and up to 20 cm long and 15 mm wide. They are arranged alternately or in more or less erect whorls along the branchlets. The fragrant flowers are of brownish-yellow colour and 10 to 18 mm in diameter. They arise in racemes with six to sixteen individual flowers. The fruits are 12 to 18 mm long tomentose drupes, which have a reddish colour when ripe. Toru occurs in lowland and montane shrublands in the New Zealand North Island roughly between 35° and 38° S latitude. Pollen of the genus has been recorded as present in New Zealand only since the Lower Quaternary (Mildenhall, 1980).

Knightia excelsa (rewarewa, NZ honeysuckle) is a tall tree, which can be up to 30 m high and can have a trunk of 1 m or more in diameter. Adult leaves are linear-lanceolate, acutely serrate and of thick, leathery structure. The flowers have spirally curled perianth segments of red-brown colour, the protruding style is about 3 cm long. Individual flowers are bird-pollinated and arranged in dense racemes, which are about 10 cm long. The tomentose fruits are up to 4 cm long and take one year to open, splitting into two halves and releasing winged seeds. Rewarewa grows from lowland to montane forests in the North Island and in the South Island in the Marlborough Sounds. The genus *Knightia* has had a continuous pollen record in New Zealand since the Eocene (Mildenhall, 1980).

3.7.2 Phylogenetic Analyses of Proteaceae *atp*B Sequence

A DNA sequence alignment of the plastid coded ATP synthase beta subunit gene (*atp*B) was used for phylogenetic analyses on Proteaceae. These data were taken from the study by Hoot and Douglas (1998). The outgroup species used by these authors (*Platanus occidentalis* L. (Platanaceae) and *Elaeagnus* sp. (Elaeagnaceae)) were also used in this analysis. Sequences for these taxa were from Savolainen *et al.* (2000). Analyses in this section were not performed on all available Proteaceae *atp*B data but on a representative taxa set. Accessions were chosen to include members of all phylogenetic groupings recovered by Hoot and Douglas (1998).

An alignment of 1359 homologous sites of the *atp*B gene was investigated with split decomposition under SplitsTree 3.1 (Huson, 1998) and quartet puzzling under PAUP 4.65 (Swofford, 1999).

3.7.2.1 Split Decomposition

The splitsgraph of Proteaceae, shown in Figure 3.17, is based on 1359 homologous sites of the *atp*B gene. Although it was calculated without outgroup sequences, it has a relatively low fit value of 67.3 - indicating that a splits graph cannot completely represent the distances in the pairwise distance matrix. The lowered fit statistic is explained by the presence of contradictory relationships at the deeper phylogenetic levels. The absence of internal tree-like structure between more anciently diverged taxa is represented as an unresolved polytomy in the splits graph. This apparent weakness of the data is expected to limit analyses with outgroups. Attempts to root the splitsgraph with *Platanus occidentalis* and *Eleaegnus sp.* suggested that the root joined in the middle of the polytomy.

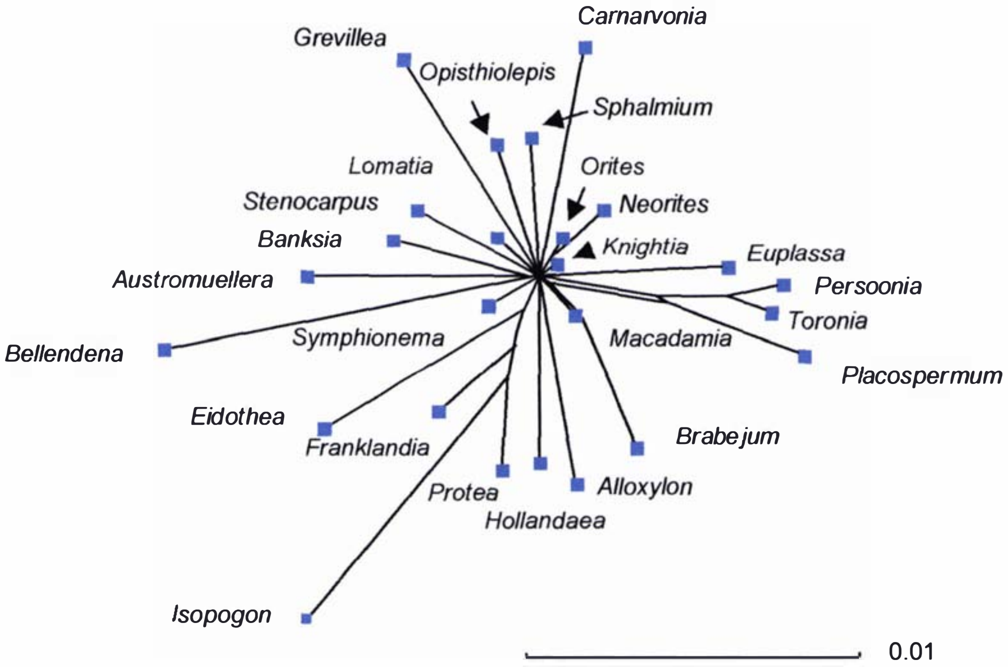


Figure 3.17: Splitsgraph for Proteaceae based on observed distances calculated from *atp*B sequence. Scale = 0.1 substitutions per site. Details: Fit = 66.1 characters = 1340 (of1359) missing = 19 (off) constant = 1147 non-parsimony = 1279 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

Despite the lack of phylogenetic resolution at deeper levels, some clear patterns of relationship do occur at higher levels of the phylogeny, between more closely related taxa. Analysis with split decomposition indicates that *Toronia* and *Persoonia* have only recently diverged from a common ancestor. The basal positioning of the Australian genus *Placospermum* suggests a recent dispersal origin of *Toronia* from Australia. Results from pollen analysis would be consistent with this conclusion as they report the first appearance of *Toronia* pollen in New Zealand sediments during the Quaternary (Mildenhall, 1980)

In contrast, the New Zealand species *Knightia* is a deep branching taxon indicating the early divergence of this taxon from a Proteaceous ancestor. However, in spite of its inferred antiquity, *Knightia* is situated on a branch, which is very short compared to other basal taxa. This suggests that the *atp*B gene region of *Knightia* may not be evolving under a molecular clock. Examinations using the relative rates test of Steel *et al.* (1996) implemented in SplitsTree 3.1 (Huson, 1998) confirm this suggestion.

3.7.2.2 Maximum Likelihood Analyses

The lack of deep branching resolution in Proteaceae *atp*B sequences was also observed when using other methods of tree building. The phylogeny shown in Figure 3.18 was made under the HKY model. This model and the chosen parameter settings were determined by likelihood ratio tests using nested models of sequence evolution. Model fitting analyses were performed on 8 equally parsimonious trees.

Low bootstrap and low quartet puzzle values were recovered for the deep branching taxa. These results suggest that the reason for the low resolution lies in the data itself and is not explained by potential insufficiencies of the tree building methods used.

As also observed under split decomposition, Neighbor Joining/bootstrap and quartet puzzling analyses suggest the recent divergence between the New Zealand genus *Toronia* and its Australian relative (*Persoonia*). The basal positioning of the Australian genus *Placospermum* combined with palynological evidence suggests recent trans-Tasman dispersal between Australia and New Zealand.

The placement of *Knightia* as an early diverged species in the genus in the *atp*B phylogeny is consistent with and corroborates inferences from pollen analysis suggesting an ancient presence in New Zealand since the Eocene (Mildenhall, 1980).

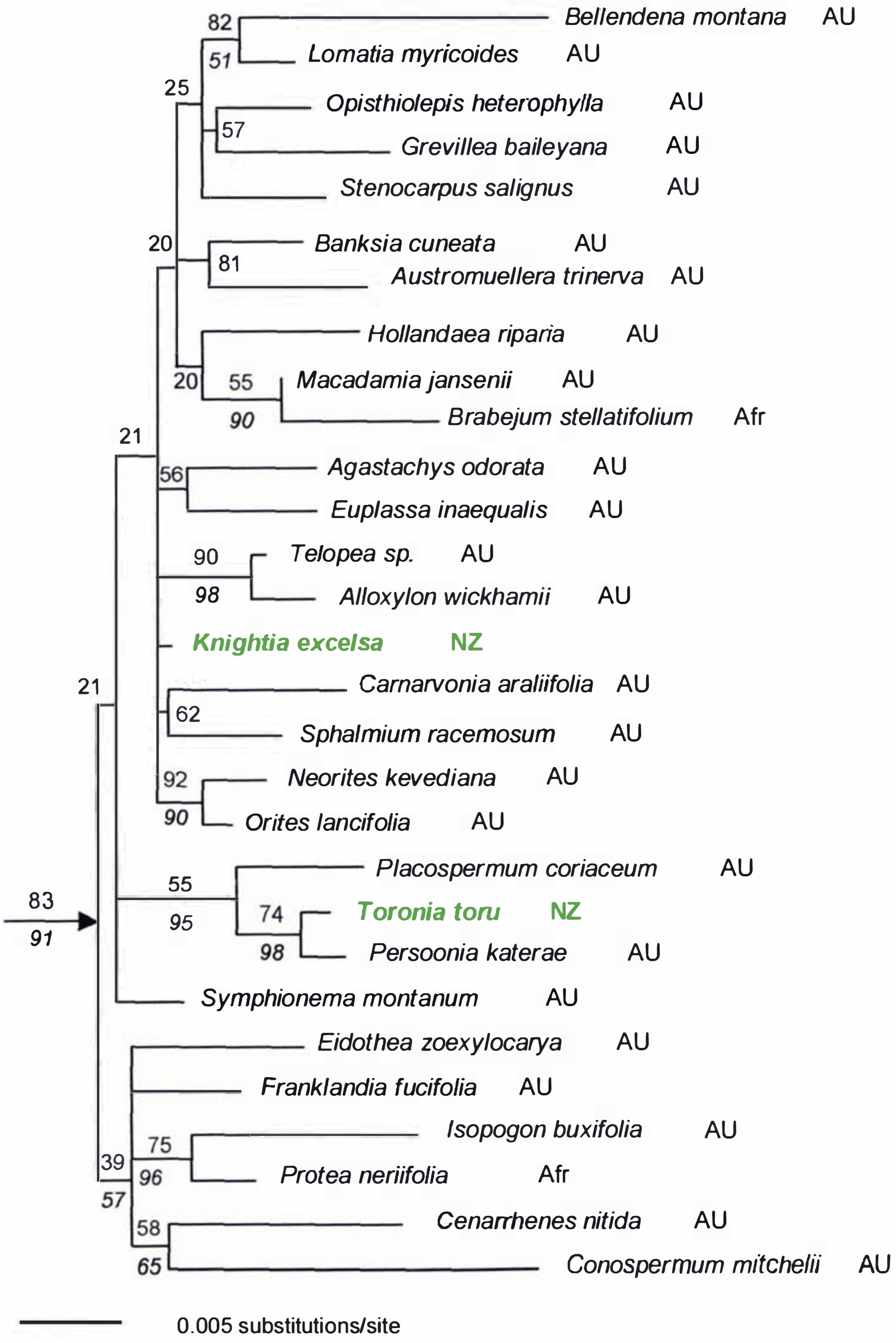


Figure 3.18: Rooted phylogeny for Proteaceae based on plastid coded *atp*B gene made with quartet puzzling and ML parameters under PAUP* (Swofford, 1999). The arrow indicates the position of the root as inferred by outgroup placement with *Platanus* and *Elaeagnus*. Quartet puzzle and bootstrap values (italics) are assigned to the edges. -ln Likelihood = 4123.48467. Parameter model: HKY. ML-settings: assumed nucleotide frequencies: empirical with A = 0.22388 C = 0.29038 G = 0.27732 T = 0.20842; number of substitution types = 2; proportion of sites assumed to be invariable = 0.732807 (observed proportion of constant sites = 0.818731); transition/transversion ratio = 0.92881 (kappa = 5.94440). distribution of rates at variable sites = equal. Abbreviations: **AU**: Australia, **NZ**: New Zealand, **Afr**: Africa.

3.7.2.3 Molecular Clock Tests and Substitution Rate Estimates

Molecular clock tests performed under the SplitsTree 'test_clock' option (Huson, 1998) suggest that in some Proteaceae the evolution of the *atp*B gene does not follow molecular clock assumptions. The lack of phylogenetic resolution at deeper levels in the tree and the non clock-like rate of *atp*B evolution in some taxa including *Knightia* make it impossible to use these data to quantitatively test the question of whether the presence of *Knightia* in New Zealand predates the Gondwanan break-up.

3.7.3 Discussion: *Atp*B and the Trans-Tasman Connections of New Zealand's Proteaceae

The reanalysis of the *atp*B data of Hoot and Douglas (1998) presented in this section, highlights a number of features of the data not raised by the earlier analyses.

The data corroborates inferences from palynological observations that find evidence for the ancient presence of *Knightia* pollen in New Zealand and the relatively young presence of *Toronia* pollen. Tree building analyses based on *atp*B sequence data show *Knightia* branching deeply within the phylogeny. In contrast, *Toronia* appears as a close relative to the Australian taxon *Persoonia katerae*. The strong relationship between these two taxa and the Australian species *Placospermum coriaceum* implies that *Toronia* and *Persoonia* have relative recently diverged from a common Australian ancestor. The molecular data suggests a trans-Tasman dispersal event during the Pliocene-Pleistocene that gave rise to the extant New Zealand species.

However, only limited phylogenetic information was gained by the analyses of *atp*B sequence data as they do not contribute towards a detailed understanding of phylogenetic relationships among Australian and New Zealand Proteaceae. The reason for this is the nature of substitution patterns in the *atp*B gene. At higher taxonomic levels the marker provides relatively treelike data, however at deeper levels the marker provides poor resolution, suggesting site saturation. It is possible, that some sites in the *atp*B gene in the taxa examined accumulate changes quickly while only few sites accumulate changes slowly. This situation may result in a good indication of relationships at a shallow phylogenetic level but poor resolution at a deeper level.

Hoot and Douglas (1998) also used the *atp*B-*rbc*L spacer region for phylogenetic analyses on the same set of accessions. This plastid-coded region is very variable and hence suitable only for studying relationships between closely related taxa. In this thesis work, preliminary analyses of this region showed great variation in sequence length between Proteaceae. This made it difficult to be confident whether homologous positions were correctly aligned. Hence, these data were not considered potentially useful for further investigating basal relationships within the Proteaceae phylogeny. To elucidate phylogenetic relationships between more distantly related taxa additional sequence data for genes showing conservative substitution patterns (e.g. nuclear ribosomal gene sequences) are needed.

3.8 WINTERACEAE LINDL., 1830

3.8.1 Introduction

Traditionally Winteraceae were placed in the order of Magnoliales (Cronquist, 1998). More recently the group was excluded and the order Winterales was established by Reveal (1993) accommodating Winteraceae as its only family. The family Winteraceae combines nine extant genera with about 100 described species.

For decades botanists have considered Winteraceae to be one of the most archaic representative of the hypothetical ancestral prototype of the early flowering plants. A cladistic study of morphological characters suggested an alignment of the Winteraceae with the Illiciaceae and Schisandraceae (Donoghue & Doyle, 1989). Recent phylogenetic studies based on molecular data suggested that Winteraceae are an early diverged angiosperm family (Soltis *et al.*, 1999; Ul *et al.*, 1999). However, no close relationship was recovered between Winteraceae and Illiciaceae or Schisandraceae. The study rather supported affiliation of Winterales with Piperales and placed both groups within subclass Magnoliidae.

All species of Winteraceae have their distribution in the tropical and subtropical regions of the Southern Hemisphere. They are found in Mexico, South America, South East Asia, New Zealand and Madagascar (but not in Africa). *Drimys* J.R. Forster & G. Forster extents from South America into Mexico. *Tasmannia* DC. occurs in Australia, New Guinea, and on the Philippines. *Bubbia* Tiegh. is found in Australia, New Guinea, and New Caledonia. New Caledonia is also home to species from *Exospermum* and *Zygogynum*. Baill. New Zealand hosts the endemic genus *Pseudowintera* Dandy. The genus *Takhtajania* Baranova & Leroy is known only from a single collection made in Madagascar in 1909 (Leroy, 1980).

The Winteraceae include trees and shrubs; all bearing essential oils. The simple leaves are usually entire and without stipules, they are petiolate and usually gland-dotted. Leaf arrangement is alternate or spiral. Flowers are small and mostly actinomorphic, they can be perfect or unisexual. Flowers are hypogynous with a corolla organised in two whorls, carpels of the gynoecium are free. Inflorescence is a solitary flower or in cymes. Winteraceae have small and oily berry-like or follicular fruits with endosperm.

Pollen of the family was detected in ancient sediments as early as the Lower Cretaceous. Presumably the earliest winteraceous fossils have been reported from Israel and estimated to be 105 to 110 million years old (Walker *et al.*, 1983).

3.8.1.1 New Zealand Winteraceae:

Pseudowintera Dandy is endemic and the only genus of Winteraceae in New Zealand. Allan (1982) distinguishes three species. All species are evergreen shrubs with a distribution in lowland and montane forest communities. Leaves are entire with glands obvious from the surface. The flowers arise as axillary fascicles of two to ten along the stem. They are about 1.5 cm in diameter, faintly aromatic and of greenish-white colour. The ripe fruits are baccate and carry two to six fruits.

In New Zealand Winteraceae pollen have been reported from Upper Cretaceous sediments (Mildenhall, 1980). However, the possibility could not be excluded that the pollen originated from a genus, which is now extinct in New Zealand.

3.8.2 Phylogenetic Analyses of Winteraceae Nuclear rDNA Sequence

An alignment of 666 bp from nuclear ITS1, ITS2 and 5.8S sequence data from the Winteraceae was constructed using ClustalX (Thompson *et al.*1998) and used for phylogenetic analyses. Winteraceae rDNA sequence data was taken from Suh *et al.* (1993). Outgroup analyses were attempted using Canellaceae sequence, which was taken from the same authors. However, as genetic differences between Winteraceae and Canellaceae were high in both ITS regions, sequences were problematic to align. For this reason no outgroup was added to the alignment.

3.8.2.1 Split Decomposition

The splitsgraph built from Winteraceae rDNA sequence shown in Figure 3.19 is a very tree like graph. The fit value of 79.5 and bootstrap values between 80 and 90% for some internal edges indicate good support and relatively few contradictory patterns in the sequence data.

The Madagascan species *Takhtajania perrieri* is the most distant taxon. The SplitsTree fit value increases slightly when this taxon is not included in the analyses (result not shown).

Winteraceae from different geographic regions are shown as monophyletic groups, which are separated by relatively long internal edges. Interestingly, within these groups, the New Caledonian species *Exospermum stipidatum*, *Zygogynum balansae* and *Z. bicolor* appear to be polyphyletic.

The graph suggests recent divergence between the two species included for the Australian genus *Tasmannia*. A small amount of genetic differentiation is recovered between accessions of *T. inspida*. Recent speciation is also suggested for the two New Zealand species of *Pseudowintera*.

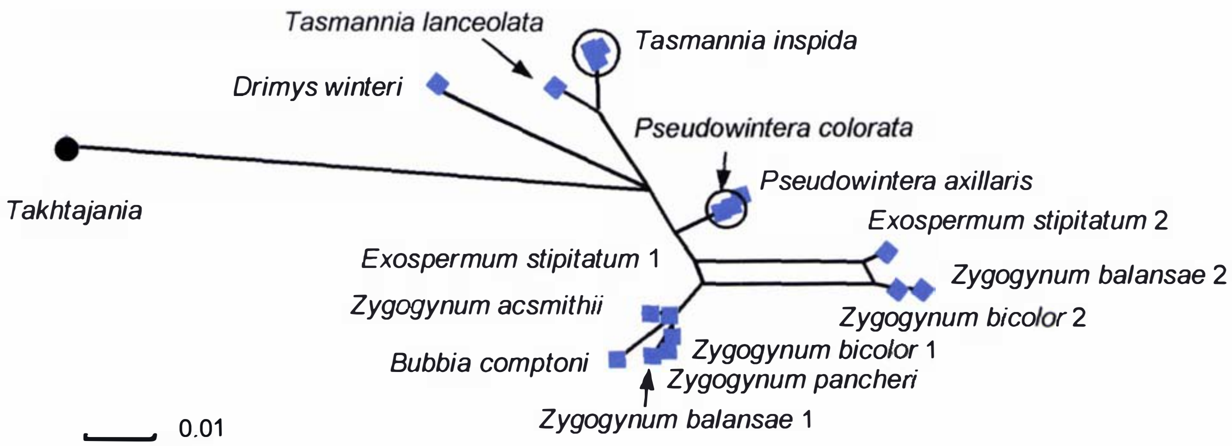


Figure 3.19: Splitsgraph for Winteraceae based on observed distances from nuclear rDNA sequence data. Bootstrap values shown on the figure are taken from 100 replicates. Scale = 0.1 substitutions per site. Details: Fit = 79.5, characters = 604(of 666), gaps = 62(off) constant = 469 non-parsimony = 531, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

3.8.2.2 Maximum Likelihood Analyses

The nuclear data set was also analysed with quartet puzzling under maximum likelihood criteria. The model and parameter settings were determined by model fitting analyses on one most parsimonious tree. The ML phylogeny is shown in Figure 3.20.

Consistent with the tree-like splitsgraph a high degree of support was recovered for the monophyly of taxa from different geographic regions, which is indicated by high QP values. Based on the angiosperm phylogeny published by Qiu *et al.* (1999), *Takhtajania perrieri* was given outgroup status and placed basal in the phylogeny. The Australian genus *Tasmannia* appears as the earliest diverged genus in the phylogeny. The South American genus *Drimys* is sister taxon of the more derived Indo-Pacific group containing a New Zealand lineage (two species of *Pseudowintera*) and a New Caledonian lineage (*Exospermum*, *Bubbia* and *Zygogynum*).

Sequence divergence between the New Zealand taxa is very low, suggesting a very close relationship between both *Pseudowintera* species. Split decomposition as well as quartet puzzling analyses indicate that *Pseudowintera axillaris* may be directly derived from a *P. colorata* – like ancestor.

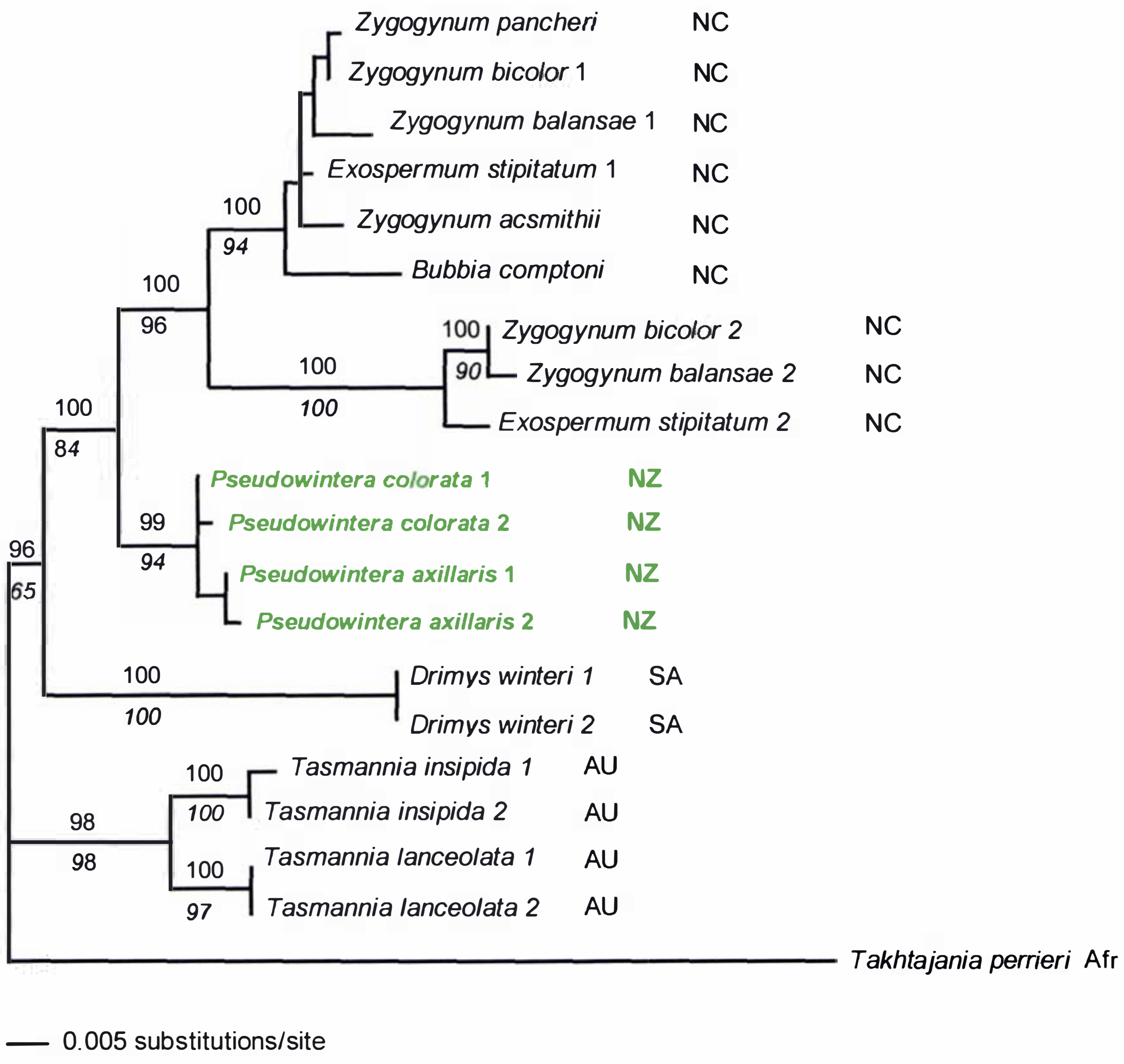


Figure 3.20: Unrooted quartet puzzling (1000 steps) tree for Winteraceae based on nuclear rDNA (604 bp of 666 bp) QP values added onto the edges. Maximum Likelihood used as optimality criterion. Edge length estimated under ML. *Takhtajania* was given outgroup status. QP and bootstrap support-values (italics) taken from 100 replicates are added onto the branches. -ln likelihood = 1787.54938. Substitution model: HKY. ML settings: number of substitution types = 2; nucleotide frequencies: empirical with: A = 0.26080 C = 0.20059 G = 0.24970 T = 0.28890; transition/transversion ratio = 1.7886 (kappa = 3.62259); proportion of sites assumed to be invariable = 0.483868 (observed proportion of constant sites = 0.774834); distribution of rates at variable sites = equal. Abbreviations: **Afr:** Africa (Madagascar); **AU:** Australia; **NC:** New Caledonia; **NZ:** New Zealand; **SA:** South America.

3.8.2.3 Molecular Clock Tests and Substitution Rate Estimates

Molecular clock tests performed under the SplitsTree 'test_clock' option (Huson, 1998) showed that the ITS regions of the New Zealand *Pseudowintera* lineage did not evolve according to molecular clock assumptions. Despite this observation, a substitution rate was calculated using the divergence between the New Zealand species of *Pseudowintera* and the Australian genus *Tasmannia* and the break up of

Gondwana to calibrate the molecular clock. This rate was expected to be lower than average.

The genetic distance between *Pseudowintera* and *Tasmannia* was calculated under the SplitsTree 'diverge' option. With the distance of DIST= 0.0516 ± 0.0084 [substitutions per site] and a divergence of 130 million years between the two lineages a substitution rate of $s_r = (0.397 \pm 0.064) \times 10^{-9}$ [substitutions per site per year] was estimated for the nuclear ITS regions of Winteraceae. This rate was similar to that used by Suh *et al.* (1993) to argue for a vicariant distribution of Winteraceae in the Southern Hemisphere.

number of base pairs:	its 1: 253 bp its 2: 270 bp
molecular clock tests:	MC assumption failed for both New Zealand species of Pseudowintera, which suggests a slower substitution rate for these taxa.
excluded sites:	ambiguous and gap positions: 14 sites invariable sites = 0
distance measure:	Genetic distance between *Pseudowintera* and *Tasmannia*
calculated distance:	DIST= 0.0516 ± 0.0084 [substitutions per site]
fossil date:	Separation of New Zealand from Australia 65 million years ago. (130 m.y. parallel evolution between *Pseudowintera* and *Tasmannia*)
substitution rate	$\underline{s_r = (0.397 \pm 0.064) \times 10^{-9}}$ [substitutions per site per year]

Table 3.9: Details and results of substitution rate estimate for the nuclear ITS regions of Winteraceae.

3.8.3 Discussion: Pseudowintera

Phylogenetic analyses on Winteraceae based on nuclear rDNA sequence showed great genetic differentiation between geographically distant genera. Closer phylogenetic relationships were only found within taxa of similar geographic origin. As demonstrated in both phylogenies made with split decomposition and quartet puzzling, the most distant taxon is the Madagascan species *Takhtajania perrieri* (Figures 3.19 and 3.20). Molecular studies on angiosperm phylogeny published by Qiu *et al.* (1999) recovered support for a basal position of this taxon within the family. Based on these findings, phylogenetic analyses presented in this chapter used *Takhtajania* as an outgroup. Quartet puzzling suggested the Australian genus *Tasmannia* to be the closest relative of the distant Madagascan species. The South American genus *Drimys* diverges next in the phylogeny and is a sister group to a Pacific clade. Analyses suggested recent speciation between members of this group. Evidence was also recovered for a recent divergence between both extant species of *Pseudowintera* in New Zealand.

Molecular clock tests on Winteraceae showed that the ITS regions of *Pseudowintera* do not evolve under MC assumptions. In spite of this observation *Pseudowintera* was included in substitution rate estimates for Winteraceae. Suh *et al*. (1993) estimated substitution rates for both ITS regions using the average divergence between the *Bubbia-Zygogynum* complex and *Pseudowintera.* To calibrate a molecular clock for ITS sequences in the Winteraceae, these authors also used the separation of New Zealand from Gondwana. Their resulting substitution rates ranged between 0.3 and 0.6 x 10^{-9} [substitutions per site per year]. This rate is similar to the substitution rate of $s_r = (0.397 \pm 0.064) \times 10^{-9}$ [substitutions per site per year] estimated in this section, which used the divergence between New Zealand and Australian taxa as a calibration point.

The first Winteraceous pollen to occur in New Zealand appears in Upper Cretaceous sediments. However, these are more similar to pollen of the South American genus *Drimys* than they are to extant *Pseudowintera* pollen (Mildenhall, 1989). During the Upper Cretaceous, Gondwana was still a continuous landmass - Antarctica connected Australia and New Zealand was connected with South America.

If the *Pseudowintera* lineage has had a continuous presence in New Zealand since the break-up of Gondwana, then it might be possible to speculate that the genus evolved from a *Drimys*-like ancestor after the break-up of the super-continent, subsequently evolving its unique pollen type.

3.9 DISCUSSION

The studies presented in this chapter investigated available sequence data and sought to evaluate its potential for testing hypotheses on origins of some elements within the New Zealand forest flora. Results were presented from analyses of seven plant groups, which have living representatives in the New Zealand flora.

3.9.1 Data and Analyses

3.9.1.1 Sampling and General Sequence Substitution Properties

At present the modest availability of taxa for some groups weakens or limits the phylogenetic inferences that can be drawn. This is particularly the case for the Podocarpaceae, Myrsinaceae and genus *Metrosideros*. For these three plant groups, analyses mainly concentrated on elucidating relationships among New Zealand taxa. Including more overseas taxa in future studies is expected to be helpful in improving current understanding of southern hemisphere relationships and the origins of some taxa in New Zealand.

The properties of the sequence data used to make phylogenetic inferences in the present chapter were investigated under split decomposition, LRT tests for determining the best models of sequence evolution as well as molecular clock tests. Good phylogenetic resolution was observed for data sets available for Araucariaceae, Myrsinaceae, Nothofagaceae and Winteraceae. The rDNA sequence data for *Metrosideros* offered good resolution between all New Zealand taxa.

Less useful information was gained in analyses of the podocarp 18S region. Split decomposition and molecular clock tests showed, that the rDNA regions of different taxa evolve at different rates. SplitsTree visualised these differences as long and short external edges. Such extreme differences in sequence evolution between taxa are likely to cause problems for tree building methods (Lockhart and Cameron, 2001).

A similar problem may have reduced the phylogenetic resolution in the Proteaceae *atp*B data, analysis of which produced graphs with irregular edge lengths. The splitsgraph also visualised a star like relationship between basal taxa, suggesting that the *atp*B gene may not be optimal for studying early events in the evolution of the Proteaceae. The reason for this appears to be that the sites accepting mutations are too fast evolving, and this produces a saturation effect, i.e. the observed number of base substitutions in the *atp*B gene is significantly lower than the actual number of substitutions (see Section 2.4.2).

3.9.1.2 Rates of Sequence Evolution

The sequence analyses presented in this chapter suggested that a slow rate of sequence evolution is a common feature of many new Zealand forest plants. This phenomenon is predominantly observed in taxa from lineages with an ancient pollen record in New Zealand. A particularly slow rate of sequence evolution was inferred in *Agathis australis*, *Dacrycarpus dacrydioides*, *Metrosideros umbellata, Knightia excelsa* and *Pseudowintera* spp. Except for *M. umbellata*, all taxa are sole representatives of a New Zealand lineage.

In the reconstructed phylogenies for their respective plant groups, *A. australis*, *D. dacrydioides*, and *M. umbellata* are situated at a basal tree position. This is also true for *K. excelsa* although the phylogeny for Proteaceae remains largely unresolved at its base. These observations corroborate the proposed antiquity of these species. Rates of base substitution were calculated for the Araucariaceae *rbc*L gene, for non-coding *trn*L regions of *Nothofagus* and for nuclear ITS of *Myrsine*, *Nothofagus*, and Winteraceae. The values are listed in Tables 3.10 and 3.11.

The rates estimated for regions of the plastid genome (I) the Araucariaceae *rbc*L gene (0.208 ± 0.017) x 10^{-9} [substitutions per site per year] and (ii) the *Nothofagus trn*L intron and *trn*L-*trn*F intergenic spacer (0.235 ± 0.039) x 10^{-9} [substitutions per site per year] were highly congruent (see Table 3.10). In general, non-coding regions have been found to be faster evolving than coding regions as there no functional constraints on the evolution of the DNA region (Page and Holmes, 1998). However, the rate estimated for the non-coding *trn*L regions (*trn*L intron and *trn*L-*trn*F intergenic spacer) of *Nothofagus* was not significantly lower than the rate estimated for Araucariaceae *rbc*L gene. A possible explanation is that estimates on coding regions, previously made by some authors, often do not take into account the sites in the sequence that are invariable. If invariable sites are included, sequence substitution rates in coding regions are expected to be underestimated.

Savard *et al.* (1993) used the divergence between *Alnus* spp. and *Betula* spp. to estimate the rate of *rbc*L sequence evolution for this group with $s_r = 0.04 \times 10^{-9}$ [substitutions per site per year]. This rate is considerably lower than the values obtained for Araucariaceae. However, it has been found that rates of sequence evolution may differ greatly between species (Wolfe *et al.*, 1987). This discrepancy may also be explained by the use of wrong calibration dates in either case. The possibility that the fossil dates used for calibrating substitution rates are inaccurate cannot be excluded.

	substitutions per site per year
Araucariaceae *rbc*L region:	$s_r = (0.208 \pm 0.017) \times 10^{-9}$
***Nothofagus trn*L regions:**	$s_r = (0.235 \pm 0.039) \times 10^{-9}$
Betulaceae *rbc*L	$s_r = 0.04 \times 10^{-9}$

Table 3.10: Substitution rate for plastid regions for Araucariaceae and *Nothofagus*.

The substitution rates for nuclear ITS regions derived from data sets of three New Zealand taxa were estimated to be between 0.4 and 0.6×10^{-9} [substitutions per site per year] (see Table 3.11). With $s_r = (0.397 \pm 0.064) \times 10^{-9}$ the lowest rate among New Zealand taxa was calculated for Winteraceae. This observation is consistent with results from molecular clock tests and split decomposition, which suggested a comparatively low rate of sequence evolution for this taxon. The substitution rate for Winteraceae was estimated in spite of results from MC tests using the divergence between the New Zealand species of *Pseudowintera* and the Australian genus *Tasmannia* and the break up of Gondwana to calibrate the molecular clock.

In this chapter, the substitution rates estimated for nuclear rDNA regions were approximately twice as high as the rates estimated for the two plastid DNA regions of *Nothofagus* and Araucariaceae. These differences in rates of sequence evolution are consistent with findings made by other authors: plastid DNA regions have been found to evolve at a slower rate than nuclear DNA regions (e.g. Wolfe *et al.*, 1987).

Sang *et al.* (1994) proposed an average rate of nucleotide substitution of 0.78 % per million years for the herbaceous genus *Dendroseris*. This rate equals $s_r = 7.8 \times 10^{-9}$ [substitutions per site per year]. This value is ten times higher than the rates estimated in this chapter. This great difference may be explained by the shorter generation times in herbaceous plants compared to woody species such as those analysed here from the New Zealand forest. A substitution rate of $s_r = 1.1 \times 10^{-9}$ [substitutions per site per year] was estimated for birches and alders (Savard *et al.*, 1993), and such a rate is more comparable with those estimated here.

	substitutions per site per year
Myrsine:	$s_r = (0.5899 \pm 0.1165) \times 10^{-9}$
Nothofagus:	$sr = (0.5531 \pm 0.0760) \times 10^{-9}$
Winteraceae:	$s_r = (0.397 \pm 0.064) \times 10^{-9}$
Betulaceae	$s_r = 1.1 \times 10^{-9}$

Table 3.11: Substitution rates for nuclear ITS regions of three New Zealand plant groups.

3.9.2 Biogeographic History of New Zealand Forest Plants

Some general conclusions can be drawn concerning the biogeography and evolution of New Zealand forests plants. As might be expected from its relatively complex geological history, the origins of its flora are also complex. Some lineages are truly ancient, with a direct link to Gondwana. However, others are more recent and have arrived via transoceanic dispersal. In this respect, New Zealand has both the character of an ancient land and also an oceanic island.

The fossil pollen record gives evidence that several extant forest plant groups were present in New Zealand before the break-up of Gondwana. Five of these taxa were included in the analyses presented in this chapter: *Agathis*, *Dacrydium*, *Metrosideros, Nothofagus* and Winteraceae (Mildenhall, 1980; Macphail, 1997).

The molecular data clearly suggested an ancient origin of *Agathis australis* in New Zealand. In the phylogeny, the taxon represented the only member of an ancient New Zealand lineage. With findings from fossil data analysed by Daniel (1989) this result is most consistent with a Gondwanan origin of the New Zealand kauri. Divergence time estimates confirmed the observation that the New Zealand kauri may have diverged from other *Agathis* species 70 million years ago.

In the phylogeny of Podocarpaceae based on 18S rDNA sequence, *Dacrydium cupressinum*, which is the only species in the genus *Dacrydium*, occurred basal to other podocarp genera from other southern hemisphere lands. This result corroborates evidence from fossil pollen and suggests, that besides *Agathis australis, Dacrydium cupressinum* has derived in New Zealand from Gondwanan ancestors. This result is consistent with fossil records that suggest a presence of this genus on New Zealand land since the Late Cretaceous (Mildenhall, 1980; Pole, 1994).

New Zealand hosts members of two *Metrosideros* subgenera. Analyses of rDNA sequence data indicated greater genetic diversity between members of subgenus

Mearnsia that suggests that this group radiated in New Zealand prior to subgenus *Metrosideros*. On the basis of the molecular analyses presented in this chapter, no conclusions could be drawn on the age of the *Metrosideros* lineages in New Zealand. Further studies need to include a larger number of overseas taxa especially from Australia to date the divergence between species of these neighbouring landmasses. The availability of more fossil evidence would also help to elucidate the evolution and historical biogeography of this genus.

Palynological studies have provided evidence, that *Myrsine* has been present in New Zealand since the Oligocene (Mildenhall, 1980). The plants are generally bird-dispersed and thus, long-distance dispersal is likely to have been mediated by migrating birds. The genus has a wide distribution from the temperate to tropical belts predominantly in the Southern Hemisphere. Phylogenetic analyses presented in this chapter concentrated on New Zealand species and included only a very limited number of overseas species. Maximum likelihood analyses of rDNA sequence data suggested a possibly polyphyletic origin of New Zealand Myrsinaceae. In the quartet puzzling tree *Myrsine salicina* appeared basal to a Austral-Pacific group including species of *Myrsine* from New Zealand, the Kermadec Islands and Japan, *Rapanea* from Australia and *Elingamita*, an endemic taxon from the Three Kings Islands.

Substitution rate estimates made in the present study for *Myrsine* ITS sequences were based on the first appearance of the genus in New Zealand during the Oligocene. Calibration of sequence divergences with this date, suggested an ITS substitution rate very similar to the rate estimated for the same genome region of New Zealand *Nothofagus* species. These findings corroborate fossil evidence and suggest a continuous presence of *Myrsine* in New Zealand since the Oligocene. However, molecular data also suggests trans-oceanic re-dispersal of the genus subsequent to the Oligocene and this implies a polyphyletic origin of extant taxa in New Zealand.

Results from sequence analyses on *Nothofagus* suggested that extant species in both subgenera *Lophozonia* and *Fuscaspora* from Australia, New Zealand, and South America diverged after the break-up of the Gondwanan continent. Comparing the macrofossil flora from Australia and New Zealand, Mildenhall 1980 and Macphail (1997b) found evidence that the same extinct species of *Nothofagus* appear first in Australia and subsequently in New Zealand. Considering this evidence, it has been suggested, that the evolution of *Nothofagus* species in New Zealand has not only been a result of in situ evolution but also of long-distance dispersal (Mildenhall 1980; Hill, 1992). Pole (1994) also noticed the absence of *Nothofagus* pollen in many Early Tertiary localities and suggested a temporary but complete regional loss of *Nothofagus* from New Zealand. Hence, analyses of molecular data are consistent with evidence from fossil records and the interpretation that the extant distribution of species is a result of trans-oceanic dispersal.

Knightia is regarded as another ancient member of the New Zealand forests as pollen of this genus is recognised since the Eocene (Mildenhall, 1980, Macphail, 1997). However, in Australia the same pollen type is known from the Late Cretaceous. This evidence suggests trans-Tasman dispersal of this genus after the break-up of Gondwana. Although the *atp*B data analysed in this chapter lacks the power of resolving deeper phylogenetic relationships between Proteaceae (Section 3.7), it can be concluded that *Knightia* belongs to one of the more ancient lineages within the phylogeny.

The second New Zealand member of Proteaceae, *Toronia toru* has a pollen record dating back only to the Quaternary (Mildenhall, 1980; Macphail, 1997). Morphological and molecular analyses have suggested *a* close relationship between the Australian taxon *Persoonia katerae* and *T. toru* (Pole, 1994). The phylogenies presented in Figures 3.17 and 3.18 of this chapter also reveal a close relationship between *Persoonia* and *Toronia*. Furthermore, the phylogeny suggests, that both species share a common ancestor with the Australian taxon *Placospermum coriaceum*. Molecular results are consistent with the recent dispersal of a Proteaceous ancestor of *Toronia* from Australia to New Zealand.

New Zealand Winteraceae have been regarded as potential 'Gondwanan relicts' (Mildenhall, 1980; Maphail, 1997). Winteraceous pollen has been identified from Early Cretaceous sediments in New Zealand. Molecular analyses by Suh *et al*. (1993) and analyses presented in this chapter are consistent in suggesting that *Pseudowintera* represents an ancient Winteraceous genus that may have evolved in New Zealand from Gondwanan ancestors after the break-up of the continent.

CHAPTER 4

Molecular Markers for Studying Closely Related Taxa

4.1 INTRODUCTION

Genetic information has been used as a measure for evolutionary processes on all taxonomic levels. Molecular phylogenetics applies comparative analysis of protein or DNA sequences for reconstructing relationships between different species and genera or even between different phyla. Molecular phylogeography takes this concept further and emphasises historical and phylogenetic perspectives on intraspecific population structure (Avise, 1998).

Below species level, genetic variation is determined to discriminate between subgroups or individuals of one species. Measuring the amount of genetic diversity between populations can be useful information for conservation biology and horticulture. Traditionally, studies on population level have focused on the relative frequencies of different forms of genes (alleles) within populations. When two or more alleles of a gene coexist, the population exhibits polymorphism for this gene locus (genetic polymorphism). While classical geneticists have used phenotypic assessment of allele variation, molecular biology distinguishes differences in allele genotype. The evaluation of genetic polymorphism forms the basis of molecular population genetics.

Several molecular marker technologies are currently available to reveal measurable variation in genomes. They are used to discriminate between closely related taxa on the basis of their genetic disposition. Approaches that target specific regions of a genome can be distinguished from 'DNA fingerprinting methods'. While the latter typically screens the whole genome for variation, the former analyses subsets of the whole genome. Both approaches provide a means to visualise unique DNA patterns in different individuals.

Allozymes: Enzymes that display allelic variation at the same locus are called allozyme markers. Their relative migration speed during gel electrophoresis is caused by the charge differences of exchanged amino acids. However, many enzymes are invariant within populations or even between species and thus often provide low resolution between taxa of interest.

RFLP: Restriction endonuclease based methods such as Restriction Fragment Length Polymorphism (RFLP: Avise, 1994) provide a means for locating polymorphic genome regions. This DNA fingerprinting method reveals variation of DNA sequences in fragment length after restriction of whole genomic DNA or a proportion of it.

RAPD: Random Amplified Polymorphic DNA markers (RAPD, Welsh and McClelland, 1990) are generated by amplification of random DNA segments with short primers (usually about ten nucleotides long) of arbitrary nucleotide sequence. Amplification products, which typically display a multiple band pattern, are screened by gel electrophoresis. Optimally, the presence or absence of a DNA fragment is determined by the presence or absence of a compatible priming site. However, non-specific priming may lead to up to 60% artefactual polymorphism (Mueller and Wolfenbarger, 1999).

SSR: Microsatellites or Simple Sequence Repeats (SSR, e.g. Tautz, 1989) consist of two to six nucleotides that are organised as tandem arrays in a genome. The number of these repeats can vary greatly between individuals of one species. By developing PCR primers for the flanking regions of the microsatellite region the repeat unit can be amplified and allele variation is represented in the size of PCR products carrying different numbers of microsatellite repeats. The development of a sufficient number of microsatellite primers for screening closely related taxa has been found time expensive but information content and robustness of this screening method are particularly high (Mueller and Wolfenbarger, 1999). One PCR approach that makes use of SSR to identify polymorphic genome regions is termed ISSR (Inter Simple Sequence Repeats). It has been useful in numerous applications of molecular ecology and evolution (e.g. Wolfe *et al.*, 1998a, 1998b).

PCR based methods like RAPDs and SSR have received much attention in assessment of genetic diversity. Both approaches are attractive because they may require mere nanogram quantities of DNA.

4.1.1 Amplified Fragment Length Polymorphism

Zabeau and Vos (1993) have developed a method combining the DNA restriction and amplification process. The technique, known as Amplified Fragment Length Polymorphism (AFLP), is a DNA fingerprinting method by which many different gene loci are addressed in a single assay. It is a very sensitive approach, which produces a high density of DNA markers, and makes it possible to locate genetic polymorphism below species level without prior knowledge of the genome investigated (Mueller and Wolfenbarger, 1999).

The AFLP method (see Section 2.3.6) is based on the fragmentation of genomic DNA by restriction with endonucleases. Double stranded oligonucleotides (linkers) are then ligated to the 3-' and 5'- restriction site ends of the different sized DNA pieces. The known linker sequences serve as priming sites for a first round PCR. A second round selective PCR follows the less specific preamplification step. In the second round selective primers with two to four overhanging bases are used to amplify subsets of the genome. By changing the GC-content of the overhanging bases to over or under 50%

GC-rich or AT-rich DNA parts of a genome can be targeted. The banding pattern of each selective primer pair is different. The group of amplified fragments can be separated using gel electrophoresis and banding patterns can be visualised using DNA staining techniques (e.g. DNA silver staining, see Section 2.3.2.2).

The AFLP procedure detects polymorphisms that are caused either by single nucleotide changes or by deletions and insertions. Nucleotide changes as well as deletions or insertions effect the presence or absence of a particular DNA band in an AFLP profile by (i) changing a restriction site (ii) mutating the nucleotides adjacent to a restriction site which are used for selective primer annealing or (iii) changing the length of a DNA fragment.

AFLP is suitable for investigating genomic DNA of various origins and complexities (Lin and Kuo, 1995). It has been successfully implemented in the fingerprinting of micro-organisms (e.g. Gancheva *et al.,* 1999), fungi (e.g. Lee *et al.,* 2000), insects (e.g. Cervera *et al.,* 2000), ferns (Perrie *et al.*, 2000), higher plants (e.g. Beismann *et al.*, 1997; Sharbel *et al.*, 2000) and animals (e.g. de Knijff *et al.,* 2001).

Numerous different scientific questions have been approached using this technique: Genetic variation and diversity between individuals and populations have been analysed (Han *et al.*, 2000; Gaudeul *et al.,* 2000), population dynamics and outcrossing rates have been investigated (Cervera *et al.*, 2000; Krauss, 2000a, 2000b). Phylogenetic relationships, paternity and hybridisation have been studied (Ganter and de Barros Lopes, 2000; Gerber *et al.*, 2000), taxa have been discriminated and unknown samples have been identified (Nair *et al.,* 2000; Beisman *et al.*, 1997). AFLP has also been used for biogeographical questions such as postglacial colonisation of habitats (Sharbel *et al.,* 2000) and the approach has been emerging as the method of choice for quantitative trait analysis and genetic linkage mapping (Hackett *et al.,* 2000; Klein *et al.,* 2000).

Although the AFLP technique may be a powerful tool for DNA fingerprinting analysis of closely related taxa, profiles may share few common fragments when genome sequence homology is less than 90%. AFLP has been found less appropriate for comparing genomes that are evolving rapidly such as those of some microbes (Janssen *et al.,* 1996). Also the interpretation of these data for phylogeny reconstruction is problematical. Much of the difficulty associated with the use of DNA fingerprint profiles in tree building analyses is due to uncertainty in homology assessment. That is, between closely related groups of individuals it may be possible to identify homologous fragments but with more distantly related taxa this is extremely difficult (Mueller and Wolfenbarger, 1999).

AFLP is an extremely reliable method but it requires DNA of reasonable quality (Jones *et al.,* 1998). Experience with endemic New Zealand plant taxa suggests that the effectiveness of AFLP is reduced when DNA quality is poor (McLenachan *et al.,* 2000). Interfering plant metabolites diminish the efficiency of restriction enzymes and the partial digest of genomic DNA results in the presence of extra bands that may be misinterpreted as polymorphisms (Lin and Kuo, 1995).

4.1.2 Aims and Expectations

While AFLP fingerprint analysis may be of limited use for reconstructing some phylogenies (Winkworth, 2000) this technique provides a very effective method for identifying variable regions of plant genomes (Lou & Boutry, 1995; Lockhart & McLenachan, 1997). The sensitivity and reliability of AFLP can be utilised to locate polymorphic genome regions from which sequence specific PCR markers – sometimes called "SCAR" or "STS" markers (e.g. Melotto *et al.*, 1996; Shan *et al.*, 1999; Schupp *et al.*, 1999) – can be derived.

Molecular markers of this type have several advantages when studying natural plant populations. Although DNA fingerprinting may provide fine scale resolution of genetic variation, screening large numbers of accessions is labour intensive. In contrast, AFLP derived markers may allow researchers to screen a large number of accessions far more rapidly. Also, the need for good quality DNA may limit the use of herbarium tissue and some material collected from natural populations – an important source of material for rare or otherwise inaccessible taxa (McLenachan *et al.,* 2000).

The aim of the work presented in this chapter was to use AFLP to locate polymorphic genome regions for *Myrsine* and *Nothofagus*. From these regions sequence specific DNA markers were to be derived that show high variability amongst closely related individuals.

4.2 AFLP DERIVED DNA MARKERS FOR *MYRSINE*

As standard markers (rDNA, *ndh*F) were not sensitive enough to resolve phylogenetic relationships between closely related taxa of *Myrsine* (see Section 3.4) faster evolving genome regions needed to be identified. The use of direct sequencing loci was perceived as an advantage over DNA fingerprinting methods as herbarium material was to be included in these studies and the DNA quality of this material was expected to be poor.

For preliminary AFLP analysis on *Myrsine,* AFLP profiles were run with AT-rich and GC-rich primers to target the corresponding portions of the genomes. AFLP profiles were obtained from genomic DNA of *M. australis*, and several samples of *M. divaricata* with diverse geographic origins. The AFLP profiles from all primer combinations examined showed genetic variation between individuals. Polymorphic bands were excised and processed from two profiles using the primer combinations *Mse*-CAG/*Eco*-ATA (Figure 4.1) and *Mse*-CAG/*Eco*-ACG (Figure 4.2).

a) *Mse*-CAG/*Eco*-ATA

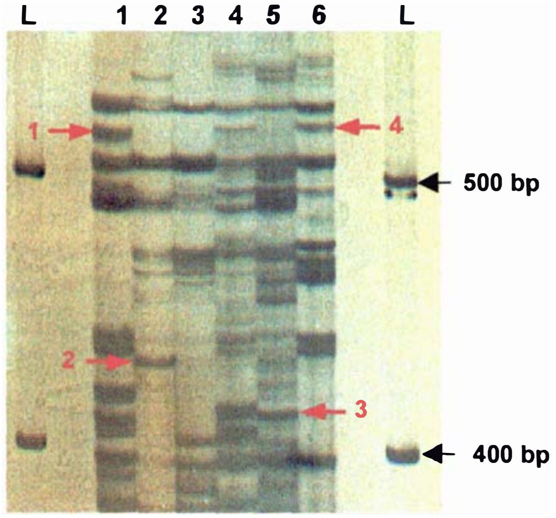


Figure 4.1: AFLP profile from *M. australis* and *M. divaricata*. Primer combination: M-CAG/ E-ATA. Lanes L: 100 bp ladder; **1:** *M. australis*, origin: Stewart Island; **2-6**: *M. divaricata*, **2:** Auckland Islands **3:** Stewart Island: Maori Beach **4:** Peel Forest **5:** Reporoa Bog **6:** Poor Knights Islands: sample 1. Arrows and numbers indicate excised DNA bands 1-4.

fragment	1	2	3	4
reamplified	+	+	+	-
cloned	+	-	+	-
sequenced	+	-	+	-
name	**Ma-af1**	-	**Md-af2**	-

Table 4.1: Relative success at each step in attempt to isolate four markers extracted from AFLP profiles obtained using the selective primer pair M-CAG/ E-ATA as shown in Figure 4.1. + indicates step successful. - indicates the step was unsuccessful.

The relative success at each step in attempt to isolate markers extracted from AFLP profiles is summarised in Tables 4.1 and 4.2. DNA sequence was obtained for AFLP fragments Ma-af1, Md-af3, Md-af5, Md-af7, Md-af8, Md-af9, Md-af11, Md-af12, Md-af13, Md-af15 and Md-af17 and are given in Appendix IV. With fragments 3, 9, and 17 possibly homologous gene regions were excised. This suggestion was confirmed by sequence analysis, which indicated that the cloned sequences differed at only two positions.

b) *Mse*-CAG/*Eco*-ACG

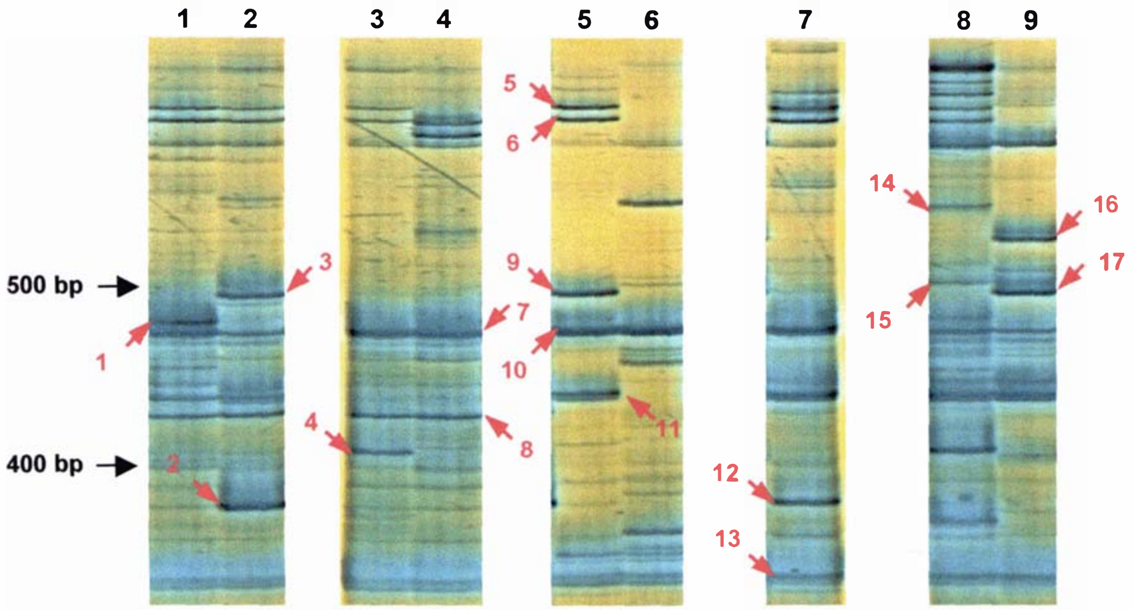


Figure 4.2: AFLP profile from *M. divaricata*. Primer combination: M-CAG/ E-ACG Lanes **1:** Manapouri; **2:** Mt. Oxford; **3:** Auckland Islands; **4:** Invercargill; **5:** Cobb Valley; **6:** *M. argentea*, Mt. Burnett: sample 1; **7:** Whakapapa; **8:** Tararua ranges; **9:** Poor Knights Islands: sample PK1. Arrows and numbers indicate excised DNA bands 1-17.

fragment	1	2	3	4	5	6	7	8	9
reamplified	-	-	+	+	+	+	+	+	+
cloned	-	-	+	-	+	-	+	+	+
sequenced	-	-	+	-	+	-	+	+	+
name	-	-	**Md-af3**	-	**Md-af5**	-	**Md-af7**	**Md-af8**	**Md-af9**

fragment	10	11	12	13	14	15	16	17
reamplified	+	+	+	+	+	+	+	+
cloned	-	+	+	+	-	+	+	+
sequenced	-	+	+	+	-	+	-	+
name	-	**Md-af11**	**Md-af12**	**Md-af13**	-	**Md-af15**	-	**Md-af17**

Table 4.2: Relative success at each step in attempt to isolate seventeen markers extracted from AFLP profiles. The potential molecular markers were obtained using the selective primer pair M-CAG/ E-ATA as shown in Figure 4.2. + indicates step successful. - indicates the step was unsuccessful.

DNA sequences obtained from the AFLP fragments were checked against the NCBI GenBank to identify possible matches. The identity score recovered by the algorithm were too insignificant to assume any homology between an existing DNA sequence and the *Myrsine* AFLP fragments. Translation of the DNA sequences to amino acid sequence showed that none of the fragments contained a continuous open reading frame.

Primers were designed from cloned fragments Ma-af1, Md-af3, Md-af9, and Md-af11 and used to amplify the four regions from the genomic DNA of *Myrsine* spp. DNA sequences of the primers are given in Table 4.3. Preliminary analyses were performed on *M. australis* (1 accession), *M. chathamica* (1 accession), and *M. divaricata* (2 accessions). PCR products were initially checked by gel electrophoresis on agarose and polyacrylamide. PCR protocols were optimised for amplification of the new molecular markers from *Myrsine spp.* The conditions used are described in Section 2.3.3.

fragment	primer name	primer sequences	preliminary observation	comments
Ma-af1	Ma-af1 AF	GTTCAGAAATAGGTGTC	a single amplification product observed for *Myrsine* spp. on 1% agarose	further investigated
	Ma-af1 AR	CAATGCCTTAGCCAATTTC		
Md-af3	Md-af3 AF	CTCGTGGTATGGGAAAGACTGG	very weak amplification product observed for *Myrsine* spp. on 1% agarose	not further investigated
	Md-af3 AR	GCTCCAGGCCGTGAACCTTCC G		
Md-af9	Md-af9 AF	GCTTTGGGCAAAGTTATACG	a single amplification product observed for some species on 1% agarose	further investigated
	Md-af9 AR	GCACATGTTTTACCATTACC		
Md-af11	Md-af11 AF	TACGATACTTGGTATGCG	two discrete bands observed for all *Myrsine* spp. on 1% agarose	not further investigated
	Md-af11 AR	CACTAATTATTGCATTCATG		

Table 4.3: Potential new molecular markers for *Myrsine*: primer sequences and preliminary observations on PCR amplification products.

4.2.1 Characterisation of Ma-af1

The AFLP derived marker Ma-af1 was successfully amplified from genomic DNA of *M. argentea*, *M. australis*, *M. divaricata*, *M. chathamica*, *M. coxii*, *M. kermadecensis*, *M. maximowizcii*, *M. oliveri*, *M. salicina,* as well as of *Rapanea perosa*, *R. variabilis*, *R. howhittiana*, and *Elingamita johnsonii*. No product was obtained for *Myrsine africana*. The length of the amplified products ranged between 470 and 450 bp. Preliminary analyses by gel electrophoresis on a 5% polyacrylamide matrix revealed that a single band was obtained for all samples save *Myrsine australis*. Single band PCR products for *Myrsine* and *Rapanea* were sequenced using a direct sequencing protocol.

Investigation of *M. australis* population samples showed that most genomes of individuals carried two different sized Ma-af1 copies of 468 and 454 bp length. Two out of eight accessions of *M. australis* displayed a single band with 468 bp. The different sized products may be attributed to allelic variation in *M. australis*. Amplification products for ten New Zealand taxa are displayed in Figure 4.3.

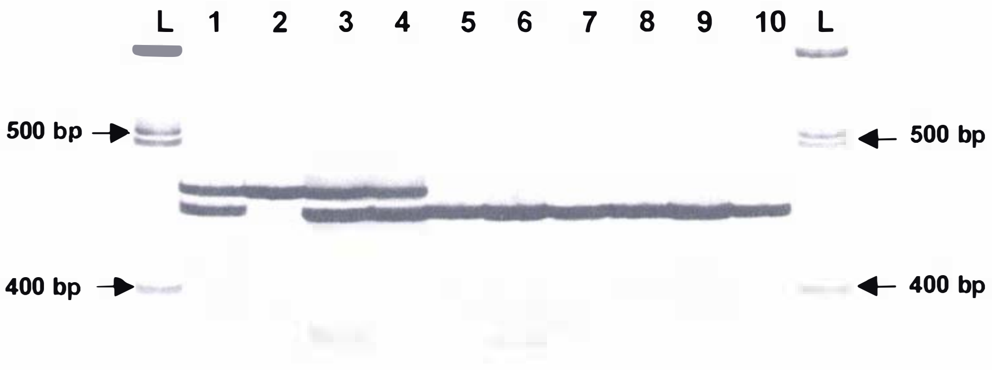


Figure 4.3: Ma-af1 amplification products for some accessions of *Myrsine spp*. **L**: 100 bp ladder; lanes **1:** *M. australis*, Lake Rotopounamu, sample ALR2; **2:** *M. australis* Lake Rotopounamu, sample ALR1; **3:** *M. australis,* Tararua Ranges; **4:** *M. australis*, Kaikoura; **5**: *M. argentea* sample Myarg_3; **6**: *M. divaricata*, Stewart Island, sample MyD_2; **7**: *M. chathamica* sample MyC_2 **8**: *M. coxii* sample Mycox_1; **9**:*M. kermadecensis* MyK_ 2; **10**: *M. salicina;* Lake Rotopounamu, SLR 1. Herbarium voucher number for the samples are listed in Appendix II.

Sequence analysis revealed the presence of a direct sequence repeat in the larger Ma-af1 copy of *M. australis* (Figure 4.4). In all other investigated samples the Ma-af1 region carried no repetitive sequence motif.

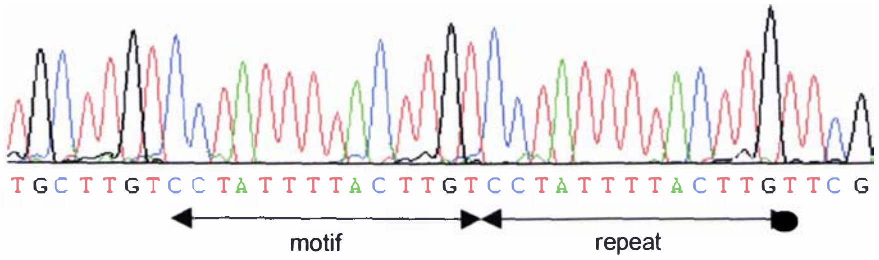


Figure 4.4: Repetitive DNA sequence motif of 14 bp length from *M. australis* Ma-af1 fragment between position 158 bp and 186 bp.

4.2.2 Characterisation of Md-af9

Amplification of Md-af9 from genomic DNA and with primers Md-af9 AF and Md-af9 AR (Table 4.3) was successful for all accessions available for *M. chathamica*, *M. coxii*, *M. kermadecensis*, *M. maximowizcii*, and *Elingamita johnsonii*. Using the same PCR protocol the fragment could also be amplified in some accessions of *M. australis* and *M. divaricata*. The size of the amplified products ranged between ca. 360 bp and 310 bp. PCR products from these accessions were successfully used for direct sequencing (Section 2.3.8).

Using a two-step PCR protocol (nested PCR, Section 2.3.3.2) and internal sequence specific primers Section (2.1.10.1), additional products were obtained for accessions of *Myrsine salicina*, *M. argentea*, *M. australis*, *M. divaricata*, *Rapanea variabilis*, and *R. howhittiana*. This gene region could not be amplified from genomic DNA of *Myrsine africana*.

Gel electrophoresis of amplified products in a 5% polyacrylamide matrix showed that multiple fragments were amplified for some accessions of *Myrsine* spp. and for all *Rapanea* species. Since a subset of these fragments could be amplified using a single primer these products were found to be a result of mispriming events. Although attempts were made to increase stringency of the PCR protocol and additional primers were designed to optimise the amplification process, it was not possible to obtain a single sized fragment from genomic DNA of these taxa.

Multiple sequence alignment of Md-af9 revealed the existence of two Md-af9 types, which differed in size by 47 bp. The smaller type was amplified and sequenced for individuals of *M. australis* and some individuals of *M. divaricata*. In all accessions of these taxa only one Md-af9 type was found. Accessions of *M. salicina* and *Rapanea* spp. were suspected to carry both Md-af9 types in their genomes.

4.2.3 Genetic Distances between Taxa

To assess genetic variability shown by the new molecular markers for *Myrsine* pairwise hamming distances were compared between Ma-af1, Md-af9 and combined nuclear rDNA sequence. The comparison was made for ten New Zealand taxa for Ma-af1 and with seven taxa for Md-af9. The results are shown in Table 4.4.

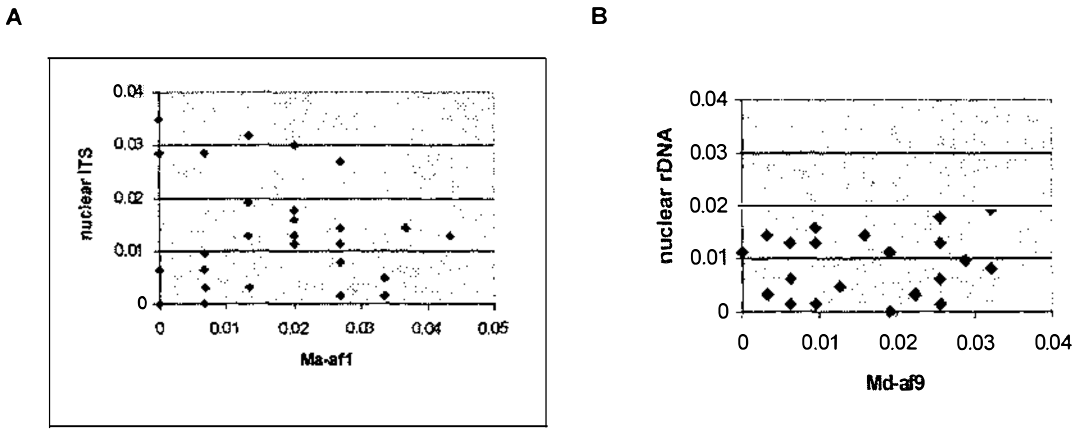
Interestingly, in many cases the values for the three markers differ greatly for the same pair of taxa. This may indicate that in different taxa the regions evolve at different rates or that the data have evolved in a non-treelike manner.

		MyArg	MyA	MyC	MyCox	MyD-PK	MyD	MyK	MyO	MyS
MyA	Ma-af1	0.02676								
	rDNA	0.00159								
	Ma-af9	-								
MyC	Ma-af1	0.00669	0.03344							
	rDNA	0	0.00159							
	Ma-af9	-	0.02508							
MyCox	Ma-af1	0	0.02676	0.00669						
	rDNA	0.00635	0.00794	0.00635						
	Ma-af9	-	0.03135	0.00627						
MyD-PK	Ma-af1	0.00669	0.03344	0.01338	0.00669					
	rDNA	0.00317	0.00476	0.00317	0.00952					
	Ma-af9	-	0.01254	0.02194	0.02821					
MyD	Ma-af1	0	0.02676	0.00669	0	0.00669				
	rDNA	0	0.00159	0	0.00635	0.00317				
	Ma-af9	-	-	-	-	-				
MyK	Ma-af1	0.01338	0.03679	0.02007	0.01338	0.02007	0.01338			
	rDNA	0.0127	0.01429	0.0127	0.01905	0.01587	0.0127			
	Ma-af9	-	0.01567	0.02508	0.03135	0.0094	-			
MyO	Ma-af1	0.02007	0.04348	0.02676	0.02007	0.02676	0.02007	0.02676		
	rDNA	0.01111	0.0127	0.01111	0.01746	0.01429	0.01111	0.00159		
	Ma-af9	-	0.0094	0.01881	0.02508	0.00313	-	0.00627		
MyS	Ma-af1	0	0.02676	0.00669	0	0.00669	0	0.01338	0.02007	
	rDNA	0.02857	0.02698	0.02857	0.03492	0.02857	0.02857	0.03175	0.03016	
	Ma-af9	-	0	0	0	0	-	0	0	
ElJ	Ma-af1	0	0.02676	0.00669	0	0.00669	0	0.01338	0.02007	0
	rDNA	0	0.00159	0	0.00635	0.00317	0	0.0127	0.01111	0.02857
	Ma-af9	-	0.0094	0.01881	0.02508	0.00313	-	0.00627	0	-

Table 4.4: Comparison of pairwise hamming distances for New Zealand species of *Myrsine* between new molecular markers Ma-af1 and Md-af9 and nuclear rDNA. Distances were obtained with Splitstree 3.1. Abbreviation: **MyArg:** *M. argentea* (Myarg_3); **MyA:** *M. australis* (MyA_9); **MyD:** *M. divaricata* (MyD_2), **MyD-PK:** *M. divaricata* (MyD_PK1), **MyC:** *M. chathamica* (MyC_1); **MyCox:** *M. coxii* (Mycox_1); **MyK:** *M. kermadecensis* (MyK_2); **MyO:** *M. oliveri* (MyO_2); **MyS:** *M. salicina* (MyS_4); **ElJ:** *Elingamita johnsonii*. Herbarium vouchers are listed in Appendix II.

To determine how well the information contents (i.e. the genetic distance a region describes between two taxa) of the new markers and nuclear rDNA correlate, the pairwise distances were plotted against each other in a co-ordinate system (see Figure 4.5). A linear function will be described when the values of both data matrices correlate perfectly (correlation coefficient = 1).

In the case of the AFLP derived markers low correlation coefficients indicate a poor match between path lengths calculated for nuclear rDNA and Ma-af1 and Md-af9 regions.



Correlation coefficient = -0.0906

Correlation coefficient = 0.117

Figure 4.5: Correlation between molecular markers. **A**: Ma-af1 and nuclear rDNA. **B**: Md-af9 and nuclear rDNA.

4.2.4 Split Decomposition and Spectral Analyses

Split decomposition and spectral analyses were used to visualise the correlation between splits and phylogenetic information in rDNA and the new molecular markers for *Myrsine*. The splitsgraphs are shown in Figure 4.6. As the Md-af9 region does not amplify for *M. argentea*, *M. divaricata* and *M. salicina* a graph is shown for a reduced set of taxa consisting of *M. australis*, *M. divaricata* from the Poor Knights Islands, *M. chathamica*, *M. coxii*, *M. oliveri*, *M. kermadecensis*, and *Elingamita johnsonii*. Split spectra ('distance Hadamard spectra', Hendy and Penny, 1993) which are visualised in Figure 4.7 compare the support of different splits recovered from all three data sets using seven taxa.

Splitsgraphs and distance spectra show that only some partitions are supported by all three data sets. The support gained for splits varies greatly between the molecular markers. The graph shows that evolutionary processes vary greatly between the different genome regions investigated. The great differences between edge length as visualised by the splitsgraph suggests that the sequences are not evolving according to molecular clock assumptions.

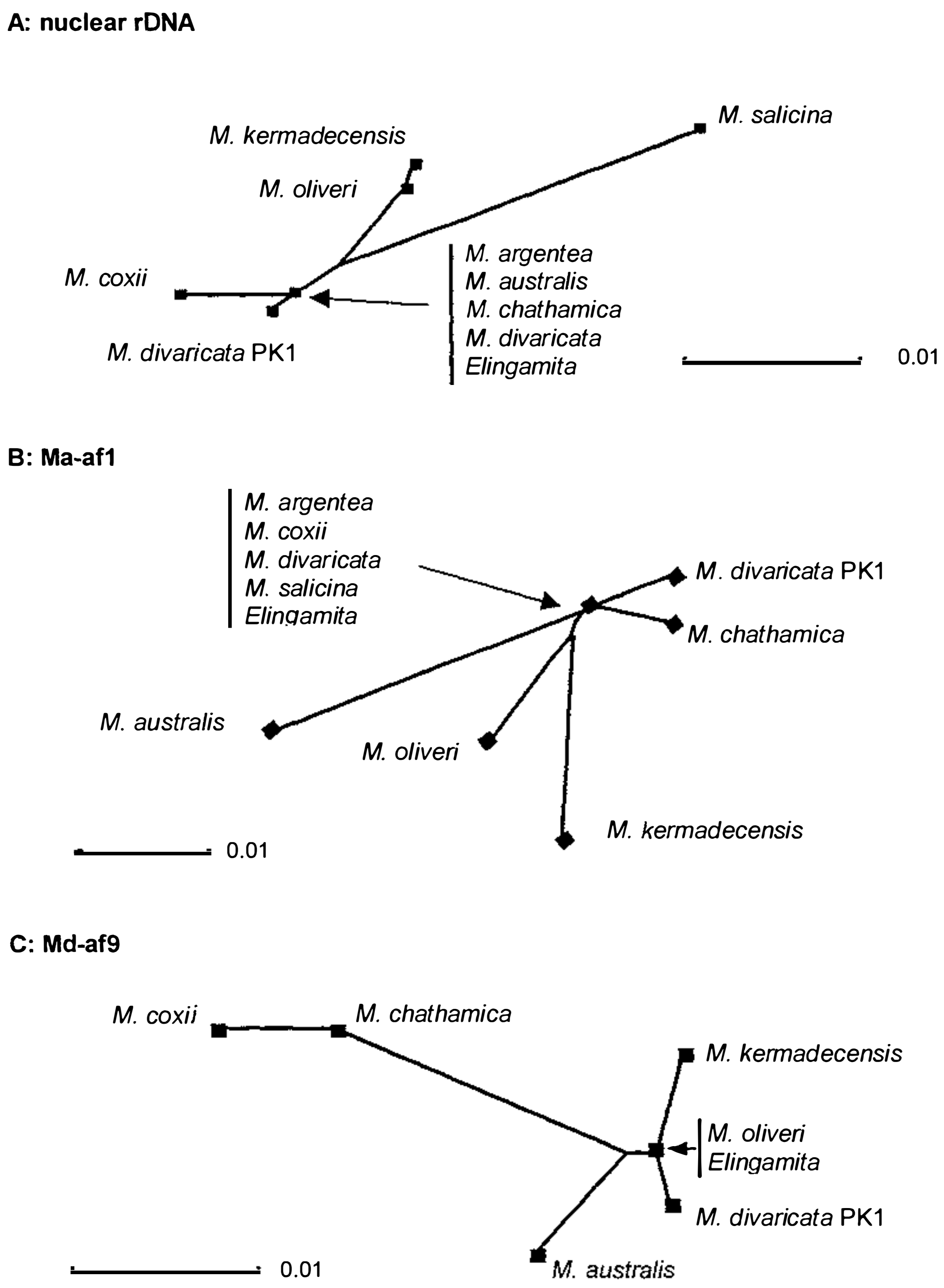


Figure 4.6: Splitsgraphs for New Zealand species of *Myrsine* and *Elingamita* **A:** based on nuclear rDNA. Details: Fit = 86.7, taxa = 10, characters = 628(of 649), gaps = 18(off), missing = 3(off), constant = 600, nonparsimony = 619, -dsplits -hamming. **B:** based on DNA sequence of the AFLP derived marker Ma-af1. Details: Fit = 100.0, taxa = 10, characters = 299(of 331), gaps = 14(off), missing = 18(off), constant = 279, nonparsimony = 298 -dsplits -hamming. **C:** based on DNA sequence of the AFLP derived marker Md-af9. Details:Fit = 100.0 taxa = 7, nchar = 319(of 329,) gaps = 1(off), miss = 9(off), const = 306, nonparsi = 313, -dsplits -hamming. Samples: *M. argentea*: Myarg_2; *M. australis*: MyA_LR1; *M. divaricata*: MyD_Stl1; *M. divaricata*: MyD_PK1; *M. chathamica*: MyC_1; *M. coxii*: Mycox_1; *M. kermadecensis*: MyK_2; *M. oliveri*: MyO_2; *M. salicina:* MyS_AT; *Elingamita johnsonii*. Herbarium vouchers are listed in Appendix II.

The Hadamard spectra show that all three data sets separate *M. australis*, *M. divaricata* PK1 and *M. kermadecensis* from the remaining taxa. While the Ma-af1 region separates *M. chathamica* and *M. oliveri*, nuclear rDNA and the Md-af9 region find no support for this split. These two regions support genetic distinctness of *M. coxii*. Nuclear rDNA and Ma-af1 support a phylogenetic relationship between *M. kermadecensis* and *M. oliveri*.

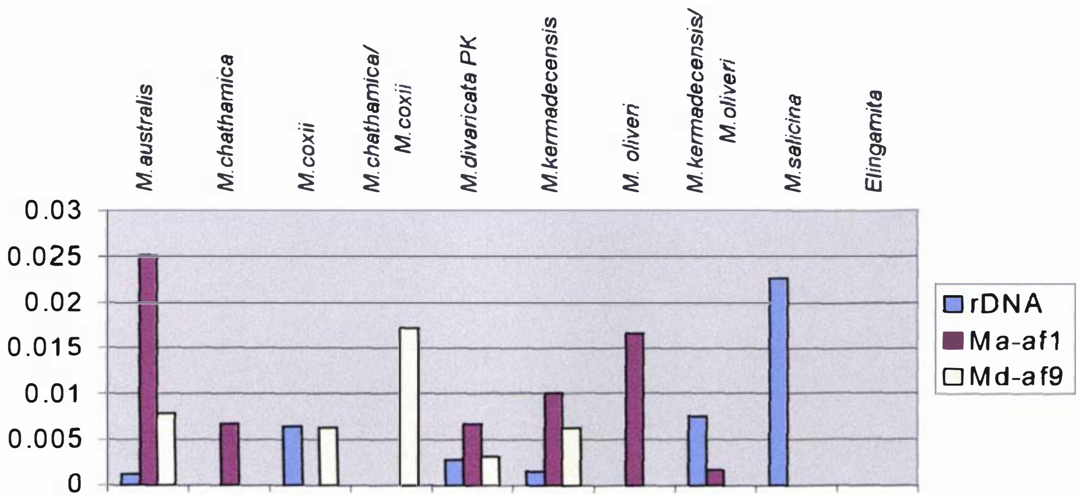


Figure 4.7: Distance Hadamard spectra of nuclear rDNA and the new molecular marker Ma-af1 and Md-af9 for New Zealand species of *Myrsine*. Entries represent patterns of support for separating the group of taxa indicated above from the remaining data set. No contradiction was observed for any of the splits. Samples: *M. australis*: MyA_LR1; *M. divaricata*: MyD_PK1; *M. chathamica*: MyC_1; *M. coxii*: Mycox_1; *M. kermadecensis*: MyK_2; *M. oliveri*: MyO_2; *M. salicina*: MyS_AT; *Elingamita johnsonii*. Herbarium vouchers are listed in Appendix II.

4.2.5 Molecular Clock Like Properties of the Data

As no outgroup sequences could be added to the Ma-af1 and Md-af9 data sets, the placement of the root in Ma-af1 and Md-af9 phylogenies was assumed to be the same as that inferred from outgroup analyses using the nuclear rDNA sequence. For molecular clock tests on Ma-af1, the root was placed on the edge leading to *Myrsine salicina* as indicated in Section 3.4.2. For the Md-af9 region, which was not characterised for *Myrsine salicina*, the root was placed on the edge leading to M. kermadecensis/M. oliveri and the remaining taxa. The tests were performed using the SplitsTree 3.1'test_clock' option as described by Huson (1998). The tests showed that both regions do not evolve according to molecular clock assumptions.

4.2.6 Summary: The AFLP Derived Markers Ma-af1 and Md-af9

Phylogenetic properties of two polymorphic gene regions isolated from AFLP profiles were investigated for *Myrsine*. The AFLP derived molecular markers Ma-af1 and Md-af9 were successfully amplified from genomic DNA of *Myrsine* spp. GenBank blast searches showed that both regions had no similarity to known DNA sequences. This can be regarded as a strong indication that both regions are not encoded by the plastid genome, but rather by the mitochondrial or nuclear genomes, which are more variable in gene composition. The presumable allele variation in the Ma-af1 region of *M. australis* supports this assumption and suggests a nuclear origin of this AFLP derived molecular marker.

PCR primers derived for both new molecular markers were highly specific and did not amplify in a large number of taxa. Both gene regions Ma-af1 and Md-af9 showed genetic variability at the sequence level between taxa. This result demonstrates that AFLP is a useful tool for detecting variable regions ('hot spot regions') in genomes.

Genetic distances between taxa obtained from both new molecular markers did not correspond with distances from nuclear rDNA regions. This could be demonstrated using the correlation coefficients between two regions. For both markers the coefficients were extremely low: 0.0934 for Ma-af1 and nuclear rDNA and 0.1189 for Md-af9 and nuclear rDNA (Figure 4.5).

The molecular clock test described by Steel *et al.* (1996) showed that both new molecular markers do not evolve according to a molecular clock.

4.2.6.1 The Ma-af1 Region

The Ma-af1 was isolated from an AFLP profile of *M. australis* to be used as a potential new molecular marker in closely related species of *Myrsine*. PCR primers were derived from the isolated fragment and amplified a homologous product in *Myrsine* spp., *Rapanea* and *Elingamita*. The fragment did not amplify in *M. africana*.

Some accessions investigated for *M. australis* carried two different sized copies, which could be separated by electrophoresis in a polyacrylamide gel matrix. This result can be interpreted as allele polymorphism on homologous parental chromosomes. DNA sequencing revealed a 14 bp insertion - a direct repeat unit - in the larger Ma-af1 copy. A single band was amplified in all other investigated species of Myrsinaceae.

Split decomposition and the comparison of pairwise hamming distances show that this region is particularly fast evolving for *M. australis*. In the splitsgraph (Figure 4.6B), the great genetic distance between *M. australis* and other New Zealand species of *Myrsine* has been visualised by a long external edge. The corresponding Hadamard spectrum showed the high support for this split. However, the DNA region did not reveal great phylogenetic resolution between other species of Myrsinaceae. This limits the use of Ma-af1 for phylogenetic purposes on interspecific and intergeneric levels. However, basepair differences in Ma-af1 sequence and allele variation allow the identification of individuals of *M. australis*; the Ma-af1 allele frequencies could be used to investigate relationships between populations of *M. australis.*

4.2.6.2 The Md-af9 Region

The Md-af9 region was derived from a polymorphic band identified in AFLP profiles of *M. divaricata*. The DNA marker was present in AFLP patterns of a subset of accessions from different geographic locations. The corresponding fragment was successfully processed and sequenced for two different individuals of *M. divaricata.* Results from DNA sequencing indicated that the DNA markers from two independent AFLP fingerprints were homologous as the DNA sequence differed by only one basepair.

PCR primers were derived from the isolated fragments to amplify the homologous regions from genomic DNA of other Myrsinaceae. Although several different primer pairs and amplification protocols were tested, homologous fragments could only be amplified in a small number of taxa. Amplification and DNA sequencing protocols were successful for *M. coxii*, *M. chathamica*, *M. divaricata* from Poor Knights Islands, *M. maximowizcii*, *M. kermadecensis*, *M. oliveri*, *Elingamita johnsonii,* and for some accessions of *M. australis*. A homologous Md-af9 amplification product could not be obtained from genomic DNA of *M. africana, M. argentea, M. divaricata* from the New Zealand mainland, *M. salicina*, or species of *Rapanea.*

The Md-af9 region shows great genetic variability between some New Zealand species of *Myrsine*. Analyses with split decomposition demonstrate treelike properties of the data, which provides good phylogenetic resolution between taxa. The high genetic variability of the region and the fact that the marker could not be amplified in a large number of taxa demonstrate the high specificity of this DNA marker. The Md-af9 region is suitable for phylogenetic analyses on a relatively small group of species, which may indicate a phylogenetic relationship between those taxa.

4.3 AFLP DERIVED DNA MARKERS FOR *NOTHOFAGUS*

Preliminary rDNA sequencing (Chapter 6.2.1) indicated that to investigate the nature of the geographic disjunction between populations of *Nothofagus menziesii* in New Zealand, molecular markers were required that showed faster rates of evolution than rDNA sequences

To locate polymorphic genome regions in *N. menziesii*, AFLP profiles were obtained for six duplicate accessions from different geographic origins. AT-rich selective primers were chosen to target non-coding and more variable proportions of the genomes.

The AFLP profiles from all primer combinations examined showed genetic variation between individuals. Polymorphic bands were excised and processed from two profiles using the primer combinations *Mse*-CAG/*Eco*-ATA (Figure 4.8).

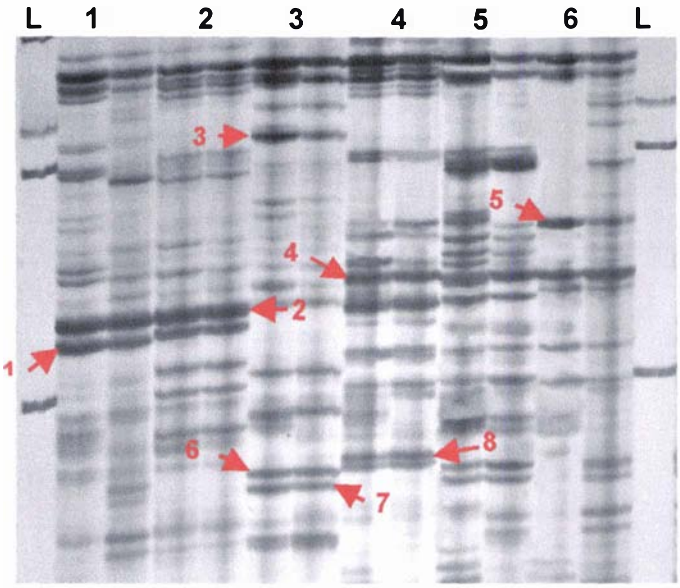


Figure 4.8: AFLP profile for *Nothofagus menziesii* from selective primer pair *Mse*-CAG-/*Eco*-ATA. **L:** 100 bp ladder **1:** Longwood: S15-1; **2:** Haast: S9-1; **3:** Abel Tasman: S1-1; **4:** Bryant Range S3-1; **5:**Tararua East: N10-10; **6:** Urewera: N5-6. Herbarium vouchers are listed in Appendix III. Arrows and numbers indicate excised DNA bands 1-8 with 8 being a double band.

AFLP band	1	2	3	4	5	6	7	8
reamplified	+	+	-	+	+	+	+	+
cloned	+	+	-	+	-	+	+	+
sequenced	-	+	-	+		-	+	+
fragment name	-	**Nm-af2**	-	**Nm-af4**	-	-	**Nm-af7**	**Nm-af8**

Table 4.5: AFLP fragments extracted from *Nothofagus* AFLP profiles made with selective primer pair *Mse*-CAG/*Eco*-ATA. **+** indicates step successful. - indicates the step was unsuccessful.

Eight fragments were excised from *Nothofagus menziesii* *Mse*I-*Eco*RI profiles and processed (Figure 4.8). DNA sequences was obtained for fragments Nm-af2, Nm-af4, Nm-af7 and Nm-af8 (Table 4.5). For AFLP band no. 8 a double fragment was excised and processed, but only one product was cloned and sequenced.

Sequences for fragments Nm-af2 and Nm-af4 were identical save an additional repetitive sequence motif that was present in Nm-af4. This DNA region was of particular interest for further investigations. PCR primers were derived for Nm-af2 and Nm-af7 and are listed in Table 4.6.

The DNA sequences of the products as given in Appendix IV were checked against the NCBI GenBank. The identity score recovered by the algorithm were too insignificant to assume any homology between an existing DNA sequence and the *Nothofagus* AFLP fragments. Furthermore, translation of the DNA sequences into amino acid sequence showed that none of the fragments contained a continuous open reading frame.

fragment	primer name	primer sequences	preliminary observation	comments
Nm-af2	Nm-af2 AF	CTGGTATTCTGTAAAGAAAATG	a single amplification product observed for all *Nothofagus* spp. on 5% polyacrylamide	further investigated
	Nm-af2 AR	CATGTTCGATGTACCATGACTTC		
Nm-af7	Nm-af7 AF	TTCAATAAGCCACAAAGGCA	two discrete bands observed on 5% polyacrylamide for all *Nothofagus* spp.	not further investigated
	Nm-af7 AR	CAGTGTATTAGTTGCGTGG		

Table 4.6: Potential new molecular markers for *Nothofagus*: primer sequences and preliminary observations on amplification products.

Amplification products were obtained with both primer pairs derived from polymorphic bands Nm-af2 and Nm-af7. A doublet product was obtained from genomic DNA of *N. menziesii* with primers Nm-af7 AF and Nm-af7 AR. The sizes of the amplification products were identical in all samples investigated from distant geographic locations. This DNA region was not further investigated.

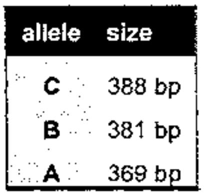
Polymerase chain reaction using primers Nm-af2 AF and Nm-af2 AR and *N. menziesii* genomic DNA produced single or double DNA bands. As the banding patterns seemed to indicate the presence of allele polymorphism for population samples of silver beech, the potential of the Nm-af2 region as a new molecular marker was further investigated.

4.3.1 Characterisation of Nm-af2

Amplification of the Nm-af2 region with primers Nm-af2 AF and Nm-af2 AR (Table 4.6) by PCR from genomic DNA was successful for accessions available from *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. The size of the amplification products ranged between 320 bp and 370 bp. PCR protocols used for amplification of Nm-af2 are described in Section 2.3.3. Sequence analysis of the Nm-af2 fragment of *Nothofagus* revealed the presence of repetitive sequence motifs in accessions of both subgenera.

The PCR products were initially analysed by gel electrophoresis in 5% polyacrylamide. Amplification products of Nm-af2 from *Fuscaspora* beeches displayed a single DNA band for the Australian species *N. gunnei* and South American species *N. alessandri*. Homologous products could also be obtained for some accessions of *N. solandri* and *N. truncata*. Non-specific amplification products were obtained with primers Nm-af2 AF and Nm-af2 AR from genomic DNA of *N. fusca*. These products were not suitable for analyses by direct sequencing protocols.

Single, one-sized products were obtained for all *Lophozonia* beeches save *N. menziesii*. The Nm-af2 primers amplified either one single sized product or two different sized products from genomic DNA of New Zealand silver beech. This result is interpreted as allele polymorphism on homologous nuclear chromosomes. Three different sized alleles were found which could be distinguished by size (Table 4.7). Five allele combinations were detected in all silver beech accessions analysed (Figure 4.9). All samples carried no more than two different sized Nm-af2 regions which is an indication that a single copy of this region was inherited from each haploid parental gamete. In 240 samples of *N. menziesii* the allele C was never found to be homozygous - it was only detected in combination with either allele A or B.



allele	size
C	388 bp
B	381 bp
A	369 bp

Table 4.7: Sizes of three Nm.af2 alleles of *N. menziesii*.

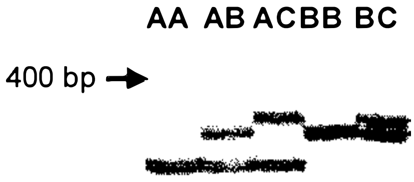
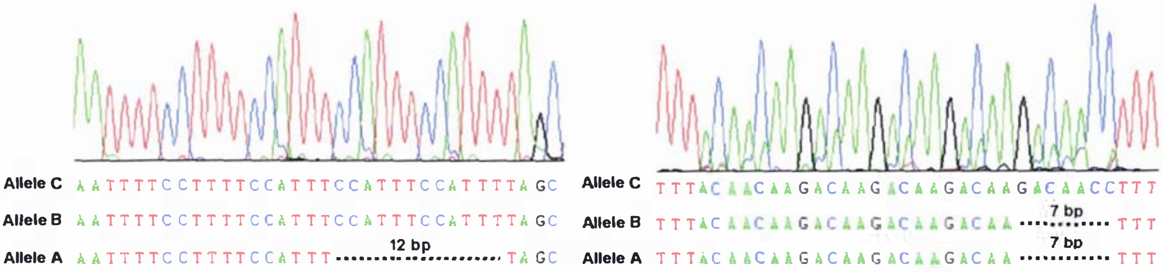


Figure 4.9: Nm-af2 allele combinations found in N. *menziesii*.

Sequence analyses showed that the three alleles for *N. menziesii* differed in number of minisatellite repeats at two sites. The sequences of the repeat units are shown in Figure 4.10.



site 1

site 2

Figure 4.10: Repetitive sequence motifs from *N. menziesii* Nm-af2 alleles A, B and C. Site 1: position 214 bp - 242 bp; site 2: position 306 bp - 335 bp.

Comparison of the repetitive sequence motifs using the BLAST algorithm available at the NCBI GenBank (http://www.ncbi.nlm.nih.gov/) showed that the sequence motif at site 1 (tttccttttccatttccatttccattt, Figure 4.10), is common in many other eukaryotes. Identical sequence repeats were found e.g. in several genome regions of *Drosophila melanogaster*, on chromosome two of *Plasmodium falciparum*, on chromosome five of *Arabidopsis thaliana*, on chromosome six of the human genome, or in the lactate dehydrogenase gene sequence of rice *and Toxoplasma gondii*. The sequence motif at site 2 (acaagacaagacaagacaa, Figure 4.10) was repeatedly recognised by the BLAST algorithm in the genome of *Drosophila melanogaster*. However, the probability scores recovered by the software programme for both queries suggest that the identities may not indicate homology but are a result of coincidence.

4.3.2 Genetic Distances between Taxa

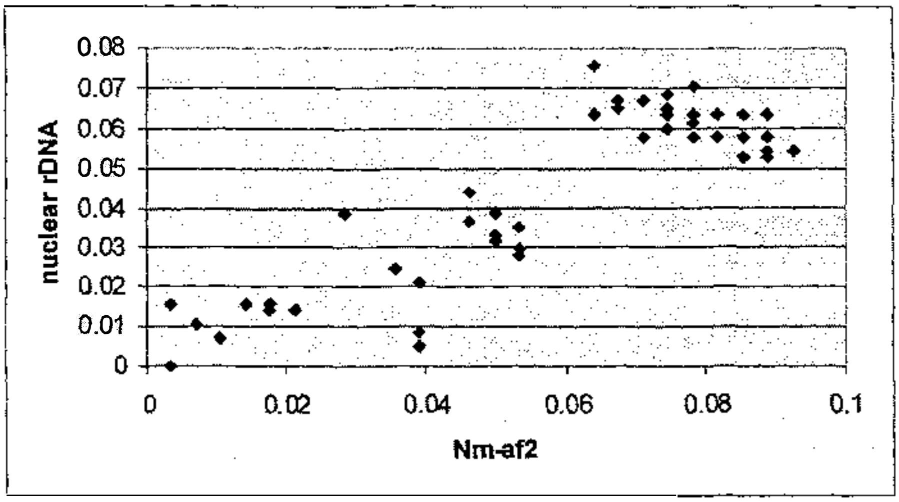
Pairwise hamming distances were calculated from Nm-af2 and nuclear rDNA (ITS1, ITS2 and 5.8S regions) and compared between accessions of *Lophozonia* and *Fuscaspora* species. For *N. menziesii* sequence data for allele B was used in the comparison. The results are shown in Table 4.8.

Pairwise distances between taxa for Nm-af2 and nuclear rDNA are remarkably similar. Only three out of 45 pairs differ by more than 50%. In over 85% of the comparisons the new molecular marker was more variable than the nuclear rDNA regions.

		N.menz	N.cun	N.moor	N.alp	N.obl	N.glau	N.trun	N.sol	N.ales
N.cun	Nm-af2	**0.03559**								
	rDNA	0.0246								
N.moor	Nm-af2	**0.03915**	**0.01068**							
	rDNA	0.02109	0.00703							
N.alp	Nm-af2	0.02847	**0.04982**	**0.05338**						
	rDNA	**0.03866**	0.03339	0.02988						
N.obl	Nm-af2	**0.04626**	**0.04982**	**0.05338**	**0.03915**					
	rDNA	0.03691	0.03163	0.02812	0.00527					
N.glau	Nm-af2	**0.04626**	**0.04982**	**0.05338**	**0.03915**	0.00712				
	rDNA	0.04394	0.03866	0.03515	0.00879	0.01054				
N.trun	Nm-af2	**0.07829**	**0.08541**	**0.08897**	**0.07829**	0.08541	**0.08541**			
	rDNA	0.06327	0.058	0.05448	0.058	**0.05272**	0.06327			
N.sol	Nm-af2	**0.08185**	**0.08897**	**0.09253**	**0.08185**	**0.08897**	**0.08897**	**0.00356**		
	rDNA	0.06327	0.058	0.05448	0.058	0.05272	0.06327	0		
N.ales	Nm-af2	**0.06762**	**0.07473**	**0.07829**	**0.06762**	**0.07473**	**0.07829**	**0.01779**	**0.02135**	
	rDNA	0.06678	0.06503	0.06151	0.06503	0.05975	0.0703	0.01406	0.01406	
N.gun	Nm-af2	0.06406	**0.07117**	**0.07473**	**0.06406**	**0.07117**	**0.07473**	0.01423	**0.01779**	0.00356
	rDNA	**0.07557**	0.06678	0.06327	0.06327	0.058	0.06854	**0.01582**	0.01582	**0.01582**

Table 4.8: Pairwise hamming distances for *Nothofagus* spp. for the new molecular marker Nm-af2 and the standard molecular marker nuclear rDNA, made with Splitstree 3.1. The higher value in each partition is highlighted. Abbreviations: ***N.menz***: *N. menziesii*; ***N.cun***: *N. cunninghamii*; ***N.moor***: *N. moorei*; ***N.alp***: *N. alpina*; ***N.obl***: *N. obliqua*; ***N.glau***: *N. glauca*; ***N.trun***: *N. truncata*; ***N.sol***: *N. solandri*; ***N.ales***: *N. alessandri*, ***N.gun***: *N. gunnii*.

The correlation of genetic distance information between the new molecular marker Nm-af2 and nuclear rDNA has been visualised in Figure 4.11. In the graph, pairwise distances are plotted against each other in a co-ordinate system. The correlation coefficient of 0.8747 supports the high congruence between both marker regions. A linear regression was calculated from the values and is shown in Figure 4.11.



Correlation coefficient = 0.878

Figure 4.11: Correlation between molecular markers Nm-af2 and nuclear rDNA.

4.3.3 Split Decomposition and Spectral Analyses

Split decomposition and spectral analyses were used to visualise the splits and phylogenetic information in nuclear rDNA regions and Nm-af2. The splitsgraphs are shown in Figure 4.12. Distance Hadamard spectra for the same data are visualised in Figure 4.13.

The splitsgraphs show that many of the tree splits suggested by the rDNA data also occur in the Nm-af2 data.

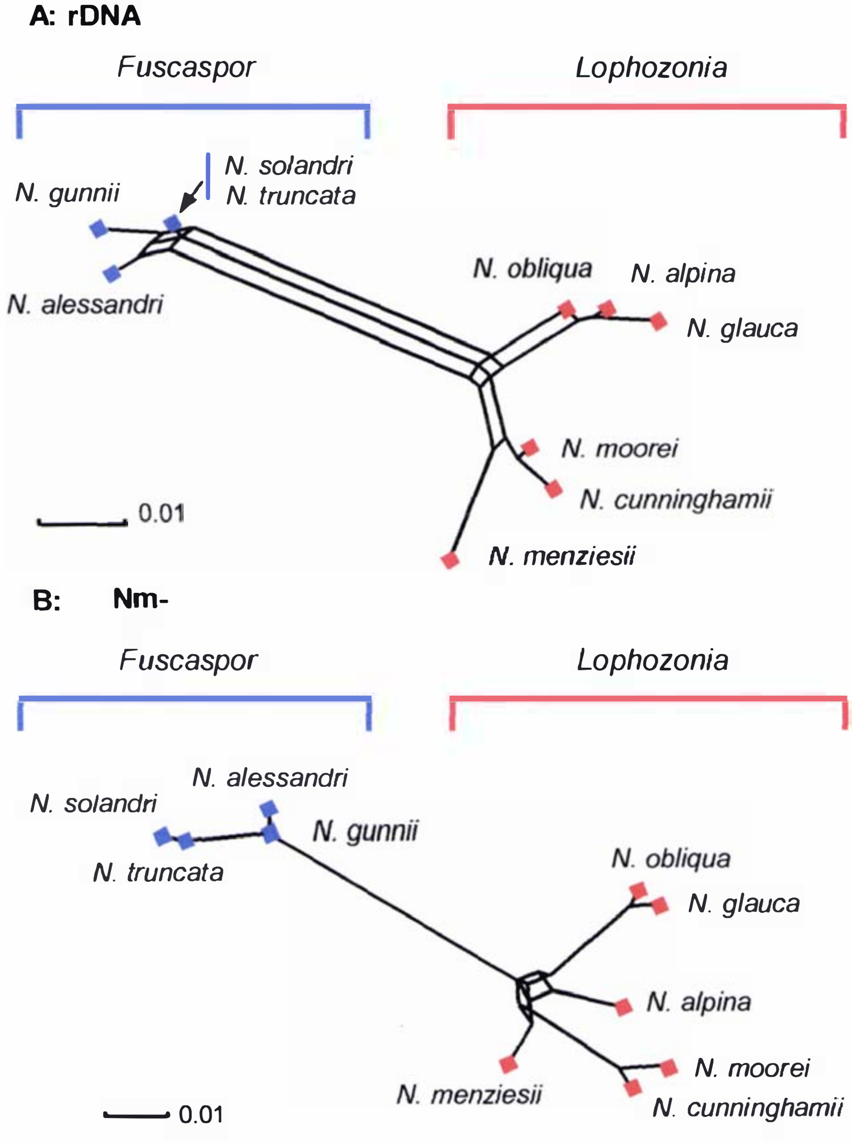


Figure 4.12: Splitsgraphs for *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. **A:** based on nuclear rDNA regions of nuclear rDNA. Details: Fit = 94.2, characters = 569(of 622) gaps = 33(off) missing = 20(off) constant = 509 nonparsimony = 527 -dsplits -hamming. **B:** based on DNA sequence of AFLP derived marker Nm-af2. Details: Fit = 91.7 characters = 281 (of 333) gaps = 49 (off) missing = 3(off) constant = 241 nonparsimony = 252, dsplits -hamming.

The Hadamard spectra visualise the support for splits in the two data sets. A high degree of correlation is suggested for the split dividing the two *Nothofagus* subgenera *Fuscaspora* and *Lophozonia* into genetically distinct groups. The splits separating New Zealand taxa from all others is also advocated by both data sets. In both data sets no support is recovered for a group combining Australian and New Zealand *Fuscaspora* (*N. solandri*, *N. truncata* and *N. gunnii*) beeches. In *Lophozonia* beeches the corresponding taxa (*N. menziesii*, *N. moorei* and *N. cunninghamii*) are grouped by the rDNA data set but not by the Nm-af2 data. However, both data sets recover support for a split separating both Australian *Lophozonia* beeches from the remaining taxa. A group combining the three South American Lophozonia beeches is also only supported by the rDNA regions but not by the new molecular marker.

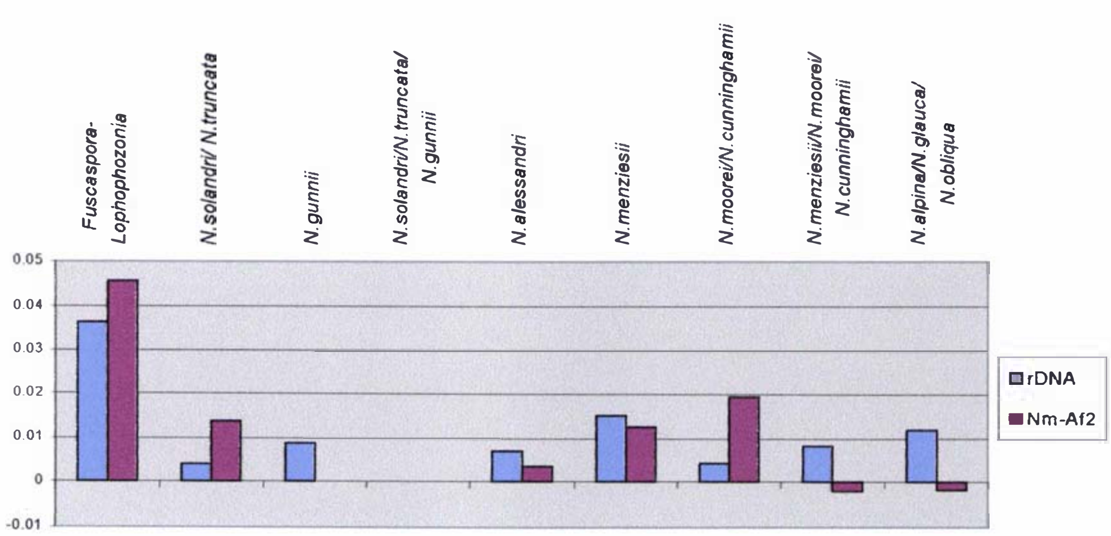


Figure 4.13: Distance Hadamard spectra of nuclear rDNA and the new molecular marker Nm-af2 for *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. Entries represent patterns of support minus patterns of contradiction for separating the group of taxa indicated above from the remaining data set.

4.3.4 Molecular Clock-Like Properties of the Data

Molecular clock tests were performed on the Nm-af2 data set for *Lophozonia* and *Fuscaspora* using the SplitsTree 3.1 'test_clock option' (Huson, 1998). The position of the root was inferred by outgroup analyses using the nuclear rDNA region, where the root joins on the long edge between both subgenera *Lophozonia* and *Fuscaspora*.

Molecular clock tests showed that the region was evolving in accordance with a molecular clock. This finding allowed substitution rates to be estimated for this marker.

A substitution rate was then estimated for Nm-af2 using the genetic distance between *Lophozonia* and *Fuscaspora* beeches. The distance was estimated under the SplitsTree 'diverge option' for the same taxa included in the splitsgraph shown in Figure 4.12. For calculating the distance between both groups, invariable sites were estimated under the Jukes Cantor model and excluded from analyses. Using these parameters, a substitution rate of $s_r = (0.461 \pm 0.051) \times 10^{-9}$ [substitutions per site per year] was calculated for the Nm-af2 regions of *Lophozonia* and *Fuscaspora* beeches. Details and results of the estimate are given in Table 4.9.

Number of base pairs:	279bp
Molecular clock tests:	MC assumption was fulfilled for all taxa.
Excluded sites:	invariable sites: 82 ambiguous sites and gap positions: 51
Distance between *Fuscaspora* and *Lophozonia*:	0.06915 ± 0.01060 [substitutions per site]
Fossil date:	Divergence between *Fuscaspora* and *Lophozonia* in the Cretaceous: 75 m.y. (150 m.y. parallel evolution)
Substitution rate	$s_r = (0.461 \pm 0.051) \times 10^{-9}$ [substitutions per site per year]

Table 4.9: Substitution rate estimates for the new molecular marker Nm-af2 for *Nothofagus*.

4.3.5 Summary: the AFLP Derived Marker Nm-af2

The new molecular marker Nm-af2 is one out of four polymorphic bands that were successfully excised from an AFLP profile and sequenced. With AFLP fragment Nm-af4, a homologous region from a second individual had been cloned and sequenced.

The region contains a number of repetitive sequence motifs ranging from two to six base pairs. Polyacrylamide gel electrophoresis of PCR products indicated the presence of allele polymorphism in population samples of *Nothofagus menziesii*. In individuals of silver beech three different sized alleles, called A, B, and C, were detected, which were present in five allele combinations. DNA sequencing revealed that the alleles differ in number of minisatellite repeats at two sites. The BLAST algorithm, which was performed to compare the AFLP fragments to data available at the NCBI GenBank, recognised these repetitive sequence motifs in other eukaryotic genomes including plants, insects and humans.

In 240 individuals of *N. menziesii* allele C never appeared homozygous. With 15%, the highest frequency of allele C was found in a central North Island population (Section 6.2.3). From this value a homozygosity of 2.25% is expected. Given that 49 individuals were sampled from this area, the probability of having included a homozygous individual is only 68%. Hence, the lack of individuals, which are homozygous for allele C, could be a result of sampling error - if more samples were included in the studies, homozygous individuals would be found.

The lack of allele C homozygotes may also indicate that a loss of fitness is connected with allele C. However, the nature of the loss of fitness remains unclear. As no continuous open reading frame was detected in either of the alleles it has to be concluded that the Nm-af2 region is not coding for a protein or oligopeptide. Hence, the possibility of a frame shift in translation caused by the seven-basepair insertion in allele C can be ruled out.

The phylogenetic properties of these genome regions were analysed and compared with results from the nuclear rDNA regions. The overall phylogenetic signals obtained from the new molecular marker under SplitsTree 3.1 (Huson, 1998) are highly compatible with results from rDNA and plastid markers presented in Sections 3.6.2.1 and 3.6.2.2. The AFLP derived molecular marker revealed greater sequence variability between species of *Nothofagus* than the rDNA regions. The results indicate a high suitability of the Nm-af2 region for phylogenetic analyses in *Nothofagus*. The region may help to elucidate relationships between closely related taxa of *Lophozonia* and *Fuscaspora*. However, as the region could not be amplified from all accessions of *Nothofagus*, i.e. no universal primer sequences could be designed, the use of this marker for more distantly related taxa may also be limited.

4.4 DISCUSSION

For studying closely related taxa of *Myrsine* and *Nothofagus* polymorphic bands were excised from AFLP profiles and converted into sequence specific DNA markers. This was done as an alternative to encoding information in AFLP profiles for intraspecific studies. For phylogenetic analyses, homology assessment of gene derivatives is crucial. This can be accomplished more easily (e.g. by sequence comparison) when dealing with specific marker regions, than from profiles where homology assessment can be problematic. Furthermore, techniques like PCR, gel electrophoresis and sequencing, which are used for analysing SCAR markers, are extremely robust. Requirements on template (i.e. the amount and quality of DNA) are particularly low: This allows the inclusion of samples for which only herbarium material or dried plant tissue is available and only limited amounts of good quality DNA can be obtained (McLenachan *et al.*, 2000).

4.4.1 How Efficient is the Approach Used to Derive Sequence Specific Markers?

The approach to generate sequence specific markers from AFLP profiles which was used in this study was previously implemented by several authors (Lockhart and McLenachan, 1997; Shan *et al.*, 1999). However, Shan *et al.* questioned the efficiency of this approach. They presumed that fragments amplified from genomic DNA might not display genetic differentiation because polymorphisms in AFLP profiles may be related to mutations at the restriction sites.

However, for *Myrsine* and *Nothofagus*, the efficiency of retaining genome regions, which showed variation at the sequence level was extremely high. 50% of fragments that were amplified from genomic DNA showed allele polymorphism and were regarded as potentially useful for phylogenetic studies (Table 4.10).

genus	no. of selective AFLP profiles viewed	no. of bands cut from wet gels	no. of fragments cloned and sequenced	no. of fragments attempted to amplify from genomic DNA	no. of fragments used for phylogenetic study
Myrsine	8	21	21	4	2
Nothofagus	6	9	4	2	1

Table 4.10: Overview over total efforts needed for obtaining three molecular markers for New Zealand forest genera *Myrsine* and *Nothofagus*.

4.4.2 Are the new Molecular Markers Identical to the Isolated AFLP Bands?

Experiments on the cloning of DNA bands extracted from AFLP bands by Shan *et al.* (1999) have shown that more than one fragment is typically amplified in the process. To increase the likelihood of cloning the fragment located in the AFLP profile the authors analysed 50 colonies from a single transformation and performed restriction analyses on the cloned inserts. Fragments that displayed a restriction pattern shared by a majority of the clones were further investigated.

For the *Myrsine* Md-af9 and Md-af15 regions two presumably homologous bands were isolated from the AFLP profile (Section 4.2.1). Sequence analysis confirmed the homology of these gene regions. The independent results provide evidence for the identity of both, the new molecular marker Md-af9 and the polymorphic AFLP fragment.

Two homologous gene fragments were also independently cloned and sequenced for *Nothofagus*. Here DNA sequences of the two gene regions Nm-af2 and Nm-af4 differed only by a 6bp direct repeat. This was a clear indication for the presence of gene polymorphism at this particular locus and explained the banding patterns observed in the AFLP profiles.

For marker Ma-af1 no evidence was brought forward to confirm the identity of this region and the AFLP polymorphism observed. However, the direct repeat unit that was observed in this gene region of *M. australis* confirmed the polymorphic nature of the Ma-af1 region.

4.4.3 What Types of Markers Were isolated from AFLP Profiles?

Different types of markers have been obtained by conversion of AFLP polymorphic bands into sequence specific markers. Variability of the gene regions has been reported to occur with respect to the size, sequence and number of alleles amplified by the primer pairs (McLenechan *et al.*, 2000). A common feature of several polymorphic genome regions that have been located from diverse genera using AFLP is that the amplified fragments contain small direct repeats. This observation was reported by Lou and Boutry (1995) for a hypervariable fragment isolated from *Hevea* using RFLP and by McLenachan *et al.* (2000) for *Rhopalostylis*.

In this study on *Myrsine* and *Nothofagus* two out of three new molecular markers (Ma-af1 and Nm-af2) contained small direct repeats in one or more taxa. The absence of marker Md-af9 in the AFLP profile of several taxa was coupled with a 46 bp deletion.

For *Myrsine,* experience with the AFLP based approach showed that the isolated polymorphic regions were variable for a small number of taxa only. The limited range of the specificity of the AFLP derived molecular markers may find its reason in the nature of the AFLP technique. The strength of this method lies in the comparison of taxa with more than 90% genetic similarity (Janssen *et al.,* 1996) because it allows the identification of changes in the genome which may be based on a single mutation event.

Results presented in the current thesis show, that using AFLP to derive sequence specific markers may be of limited value when trying to investigate a wider range of species. The high specificity of the primer sequences may not permit the establishment of a phylogenetic framework for a group of more distantly related taxa. Studies on the New Zealand plant genera *Myosotis*, *Rhopalostylis* and *Phormium* which also used AFLP to identify fast evolving gene regions for phylogeographic studies confirm this observation (McLenachan *et al.,* 2000; Winkworth, 2000).

4.4.4 What is the Potential of AFLP Derived Markers?

All new molecular markers obtained in this study were investigated as direct sequencing loci for some taxa of interest. DNA alignment for Ma-af1, Md-af9 and Nm-af2 allowed differentiation of taxa on the basis of base pair differences. In general, direct sequencing loci that reveal variation between taxa are potentially useful markers for the reconstruction of phylogenetic relationships. Alternative uses for AFLP derived markers are summarised in Table 4.11.

With the Ma-af1 region in *Myrsine australis* and Nm-af2 in *Nothofagus menziesii* different sized products were obtained by amplification using primer pairs designed from cloned AFLP fragments. The presence of different alleles could be visualised after gel electrophoresis on polyacrylamide gels. Allele distributions are often of interest in population genetic approaches.

From studies on *Phormium tenax* (McLenachan *et al.*, 2000) primer sequences derived from cloned polymorphic AFLP bands co-amplified alleles of identical (or near identical) length. In this case it has not been possible to separate these on agarose or PAA gels and direct sequencing has produced data that are difficult to interpret (due to the presence of small deletions/insertions between co-amplified PCR products). For these fragments single base sequencing has been helpful for characterising distinct allele patterns. This approach seems also promising for investigating population samples of *Myrsine australis*.

PCR product	**one size**	**multiple products**	**single and multiple size products**	**single and multiple size products**
product type	single copy region/ multiple copy region	multiple gene regions	different copies of homologous DNA fragments	parental alleles
	⇓	⇓	⇓	⇓
approach	sequencing	screen PCR products on PAA Gel	run single base sequencing reaction	screen PCR products on PAA Gel
	⇓	⇓	⇓	⇓
	DNA sequence / DNA alignment	banding pattem	banding pattem	allele distribution
	⇓	⇓	⇓	⇓
use	reconstruct phylogeny	distinguish different groups or taxa (species/subspecies) identify hybrids	reconstruct phylogeny, distinguish taxa, identify hybrids	allele frequencies of populations

Table 4.11: Possible uses for AFLP derived molecular markers taken from McLenachan *et al.* (2000).

CHAPTER 5

Phylogenetic Analyses on Closely Related Species of *Myrsine* with Particular Reference to *M. divaricata*

5.1 INTRODUCTION

Palynological studies report the first appearance of Myrsinaceae in the New Zealand pollen record during the Oligocene (Mildenhall, 1980). Phylogenetic analyses of DNA sequence data, which were presented in Section 3.5, corroborate these finding and suggest a continuous presence of members of this family on the archipelago for at least 35 million years. Support was recovered for a monophyletic origin of an Austral-New Zealand lineage comprising the two genera *Myrsine* and *Rapanea*.

The analyses have illustrated the close phylogenetic relationships between the New Zealand endemics *M. argentea*, *M. australis*, *M. divaricata*, *M. chathamica*, *M. coxii* and *Elingamita johnsonii.* Surprisingly, close affinities between this group and the geographically distant Japanese species *M. maximowizcii* were also discovered. These results, based on sequence analysis of nuclear rDNA regions, indicate recent events of speciation and dispersal.

5.1.1 *Myrsine divaricata* A. Cunn. 1839

Figure 5.1: *Myrsine divaricata*

Myrsine divaricata (Figure 5.1) the 'weeping matipo' is typically a shrub but it can grow to be a small tree up to 4 m in height (Salmon, 1996). It has rigid, spreading branches and distinctive drooping branchlets, which are covered with fine hair when young. Leaves are gland-dotted and small, about 5-15 mm long and 5-10 mm wide. Leaf shape is broad obovate to obcordate, and usually emarginate or two lobed at the apex. Leaves of a young plant are often recurved and may have small hairs along the margins. The flowers are 5-merous and 2-3 mm in diameter. They sit on a short petiole and arise individually or in small clusters below the leaves. The fruit is a dark purple berry approximately 5 mm long.

Weeping matipo occurs throughout mainland New Zealand and on its offshore islands including the Poor Knights group, Stewart Island, Auckland and Campbell Islands. The type specimen was sampled at the 'head of the Wycaddy River' in the Bay of Islands, North Island, New Zealand by R. Cunningham in 1834. Populations found on Auckland and Campbell Islands are identical with this type specimen. Only the population of *Myrsine* cf. *divaricata,* which is present on Poor Knights Islands 15 km off shore from Whangarei in the North Island, is morphologically distinguishable from the mainland type *M. divaricata*. Instead of the typical divaricating branching pattern with drooping or 'weeping' branchlets, the Poor Knights form has a more tree-like morphology with perfectly upright branches. Moreover, leaves of this type are about twice as large as typically found on the mainland. However, plants showing this particular growth form can also be observed on the mainland, occurring side by side with the divaricating form. An individual showing both morphological types on a single plant has also been observed (P. Heenan, pers. comm.). On the Poor Knights Islands, in contrast, no plants have been reported that show characteristics of the mainland type.

5.1.2 *Myrsine argentea* Heenan and de Lange 1998

In 1998 Heenan and de Lange described a new species of *Myrsine* in New Zealand (Heenan and de Lange, 1998). The new species, called *M. argentea*, is found locally only on the dolomite soils of Mt. Burnett in the northern Nelson region. Morphologically *M. argentea* is distinguishable from *M. divaricata* by characteristics shown in Table 5.1.

	M. argentea	*M. divaricata*
growth habit:	non-divaricating and straight	divaricating and weeping
leaves	silver-green	green
petals	yellow pink	maroon
stigma	protruding from the calyx in the bud	stigma enclosed within the bud
fruit	white	purple

Table 5.1: Morphological characteristics to distinguish *M. argentea* and *M. divaricata.*

5.1.3 Aims and Expectations

The New Zealand species of *Myrsine* provide an ideal opportunity to study processes of island biogeography and radiation of species. Most published papers on these subjects deal with phylogenetic analysis of one or two molecular markers (typically the nuclear rDNA region and a chloroplast region) characterised in the taxa under study. However, with the present work, a more detailed understanding of the genomic complexity and evolution of species within the New Zealand Myrsinaceae was sought. This was to be achieved by combining information gained by AFLP profiling and DNA sequencing. In this chapter, results from phylogenetic analyses of novel AFLP derived markers which have been described in Section 4.2.1, were to be compared with results from analyses of nuclear rDNA and the chloroplast *ndh*F genome region. The aim was to test hypotheses and draw conclusions regarding (i) speciation of taxa within the New Zealand Myrsinaceae, (ii) dispersal between mainland New Zealand and off shore islands, and (iii) relationships between geographic distribution, genetic and morphological variation of *Myrsine divaricata*.

5.2 PHYLOGENETIC ANALYSIS OF MULTIPLE PLANT LOCI

Analyses of nuclear rDNA regions presented in Section 3.5 recovered close phylogenetic relationships between *M. argentea, M. australis M. divaricata, M. coxii, M. chathamica, M. maximowizcii*, and *Elingamita johnsonii*. To elucidate the genetic affiliations within this group of taxa, additional phylogenetic markers were characterised. These investigations, which are reported in this chapter, compare results from analyses of the plastid coded *ndh*F gene and nuclear rDNA with results from analyses of AFLP derived molecular markers (see Section 4.2.1). Plant accessions used in this study have been listed below in Tables 3.2 and 3.3. Except for *E. johnsonii* and *M. maximowizcii* two accessions per taxon were used for analysis.

species	sample	origin	herbarium reference
M. argentea	1	NZ, SI: Mt. Bumett	MPN 24759
M. argentea	2	NZ, SI: Mt. Bumett	MPN 24760
M. australis	1	NZ, NI: Lake Rotopounamu,	MPN 24701
M. australis	2	NZ, Stewart Island: Mason Bay,	-
M. divaricata	PK1	NZ, Poor Knights group	MPN 24756
M. divaricata	PK2	NZ, Poor Knights group: Tawhiti Rahi	-
M. divaricata	1	NZ, NI: Herepai, Ruahine Ranges	MPN 24745
M. divaricata	2	NZ, SI: Invercargill	MPN 24717
M. chathamica	1	NZ, Chatham Islands	MPN 24708
M. chathamica	2	NZ, Chatham Islands	MPN 24709
M. coxii	1	NZ, Chatham Islands	MPN 24711
M. coxii	2	NZ, Chatham Islands	MPN 24712
M. kermadecensis	1	NZ, Kermadec Islands	MPN 24761
M. kermadecensis	2	NZ, Kermadec Islands	MPN 24762
M. maximowizcii	1	Japan	MPN 24763
M. oliveri	1	NZ, Three Kings group: Great Island	MPN 24764
M. oliveri	2	NZ, Three Kings group: Great Island	MPN 24765
Elingamita johnsonii	1	NZ, Three Kings group: Great Island	MPN 24776

Table 5.2: Species of Myrsinaceae used in Section 5.2. Abbreviations: NZ: New Zealand; NI: North Island; SI: South Island.

species	sample	origin	herbarium reference
M. africana	1	Africa	MPN 24775
M. salicina	1	NZ, NI: Lake Rotopounamu	MPN 24767
M. salicina	2	NZ, SI: Buckland river	MPN 24771
Rapanea howhittiana	1	Australia	MPN 24777
Rapanea porosa	1	Australia	MPN 24778
Rapanea variabilis	1	Australia	MPN 24779

Table 5.3: Species of Myrsinaceae used only in Section 5.2.1. Abbreviations: NZ: New Zealand; NI: North Island; SI: South Island.

5.2.1 *Ndh*F Gene Sequence

The plastid gene region for *ndh*F encodes NADH-Dehydrogenase, a key enzyme in the metabolism of all organisms (Neyland & Urbetsch, 1996). The gene region has been successfully used for phylogenetic analysis at the intergeneric level (Olmstead and Reeves, 1995; Prather *et al.* 2000).

1048 bp of the *ndh*F gene were sequenced and compared between individuals of the New Zealand endemics *M. argentea*, *M. australis*, *M. divaricata*, *M. chathamica*, *M. coxii, M. kermadecensis, M. oliveri, M. salicina* and *Elingamita johnsonii.* Sequence of the same DNA region was obtained from the Japanese taxon *Myrsine maximowizcii* and from three species of the Australian genus *Rapanea*: *R. howhittiana*, *R. porosa*, and *R. variabilis.*

The DNA alignment of *ndh*F sequences of Myrsinaceae revealed very little variation between individual taxa. From a total of 1047 included sites 1038 were constant between all taxa, three out of nine variable sites were parsimony informative patterns concerning all taxa. Three variable sites were singleton base pair changes in the *ndh*F sequence of *Myrsine africana*, another two singletons were observed in the *ndh*F sequence of *Rapanea porosa*, and one parsimony informative basepair difference separated *R. porosa* and *R. howhittiana* from the remaining taxa. Interestingly all variation occurred between sites 640 and 820 of the alignment.

The parsimony patterns at sites 736, 747 and 1032 allowed the classification of all taxa into three *ndh*F-types. In the case of *M. divaricata* and *M. oliveri,* individuals displayed either one out of two sequence types (Figure 5.2).

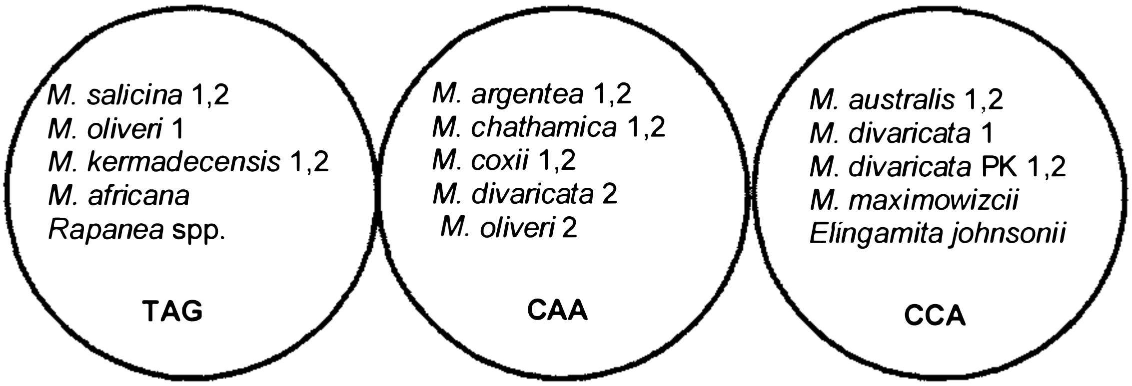


Figure 5.2: *ndh*F sequence types displayed by *Myrsine* spp. and *Rapanea* spp. Character states TAG, CAA and CCA were observed patterns at positions 736, 747 and 1032 of the multiple *ndh*F sequence alignment. Numbers refer to the accessions as specified in Tables 5.2 and 5.3.

The classification based on *ndh*F sequence types presented in Figure 5.2 shows three taxa groups. The first group comprises the Australian species of *Rapanea*, the African taxon *M. africana*, and *M. salicina*, *M. kermadecensis* and one out of two accessions of *M. oliveri* from New Zealand and its offshore islands. Outgroup analyses with *Stylogyne* sp. presented in Section 3.4 showed that these taxa are closest to the root position. The second group, which shows a derived sequence type, includes *M. argentea* and one out of two accessions of *M. divaricata* from the New Zealand mainland as well as *M. coxii* and *M. chathamica* from the Chatham Islands and accessions of *M. oliveri*. A further basepair change was observed in group three with *M. australis*, *M. maximowizcii*, accessions of *M. divaricata* and *Elingamita johnsonii*. *Ndh*F sequence is consistent with results from nuclear rDNA presented in Section 3.5 that suggest that species in group three are the most recently derived taxa.

5.2.2 Nuclear rDNA Sequence

The substitution patterns in the alignments of nuclear rDNA were more complex than that of *ndh*F. These data were visualised under split decomposition using SplitsTree (Huson, 1998).

A 464 bp alignment of ITS1 and ITS2 regions of nuclear rDNA included 18 accessions of *Myrsine* and *Elingamita* as listed in Table 5.2. Eleven informative indel positions were used as character states whereas heteroplasmic/ambiguous positions were excluded for analyses.

The unrooted splitsgraph shown in Figure 5.3 has a fit value of 100 indicating a very good representation of the data. The box-like structure at the basis of the phylogeny visualises contradictory patterns in the sequences of three taxa (*M. divaricata 1*, *M. australis 1* and *M. australis* 2).

The graph shows close relationship between the genotypes of *M. argentea*, *M. divaricata*, *M. australis*, *M. chathamica* and *Elingamita johnsonii*. More distant are *M. coxii*, *M. oliveri*, and *M. kermadecensis*, as well as *M. divaricata* samples from the Poor Knights islands and *M. maximowizcii*. Even after using indel position as informative characters, the ITS regions of both, *M. divaricata* PK1 and *M. maximowizcii* are identical. If we assume the root position, inferred by outgroup placement in Figure 3.10 in Section 3.5 (indicated by a red triangle in Figure 5.3), then the genotype of *M. chathamica* - a species from the Chatham islands – is basal to that of *M. divaricata* PK1 and *M. maximowizcii*. Both taxa are separated by over 600 km of ocean with the Chatham Islands being located south east of the Poor Knights group.

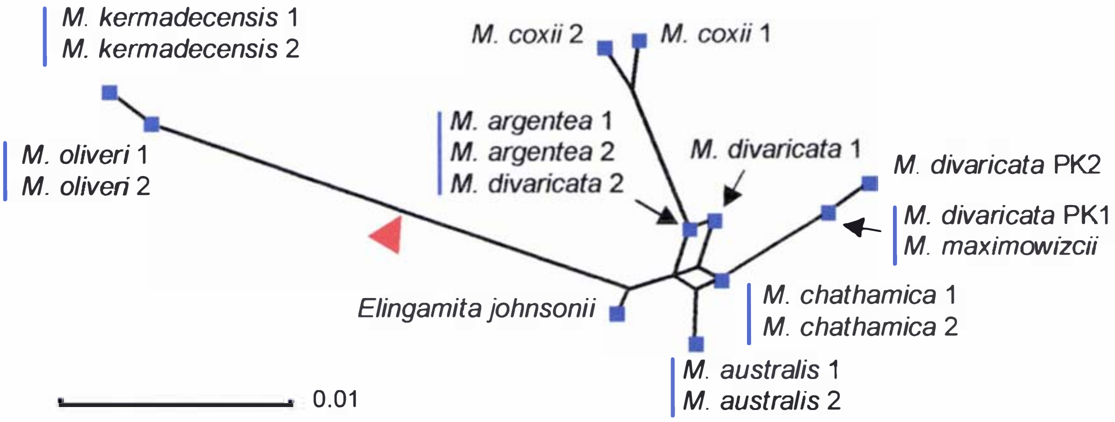


Figure 5.3: Splitsgraph for *Myrsine* spp. and *Elingamita johnsonii* based on ITS1 and ITS2 regions of nuclear rDNA. Indels are used as informative sites. The position of the root was inferred from ML analyses presented in section 3.5 and is indicated by a red triangle. Drawn to scale = 0.01 substitutions per site. Details: Fit = 100 characters = 461 (of 464) missing = 3 (off) constant = 435 nonparsimony = 439 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Samples are as listed in Table 3.2.

The ITS sequences from the two accessions of 'mainland type' *M. divaricata* differ slightly from each other as do both accessions sampled at different locations on the Poor Knights Islands. This indicates significant genetic differentiation between individual taxa of *Myrsine divaricata*.

In contrast, rDNA sequence from both accessions analysed for *M. argentea* are identical. In the splitsgraph they occupy an internal point (internal node) also occupied by one of two samples of *M. divaricata*, verifying close affiliations between the ITS regions of both species.

In the splitsgraph, *M. australis* appears on a short edge protruding from the network structure, which connects *M. argentea*, *M. divaricata* and *M. chathamica*. The low genetic distance between the rDNA genotypes of these taxa may indicate recent differentiation of these species.

M. divaricata and *M. argentea* appear basal to a relatively long edge leading to *M. coxii*. The graph suggests that the genotype of this species is directly derived from the mainland *M. divaricata/M. argentea* genotype. Both accessions analysed for *M. coxii* differ from each other, indicating variation within the genepool of this species.

The *Elingamita johnsonii* genotype shows close affiliations to those of the *M. divaricata* - *M. australis* group. However, a split also separates this species together with *M. oliveri* and *M. kermadecensis* away from the other species.

5.2.3 AFLP Derived Marker Ma-af1

Sequence data were also obtained for the AFLP derived new molecular marker Ma-af1. Analyses were performed on a 331 bp alignment, one indel of 14 bp (Section 4.2.1.1), which was observed in *M. australis*, was excluded from analysis as were sites containing heteroplasmic/ambiguous characters.

From an alignment of Ma-af1 region including 18 taxa, split decomposition recovered a very treelike unrooted graph, which is shown in Figure 5.4. A fit value of 100 indicates that distances in the data matrix are well represented by the graph.

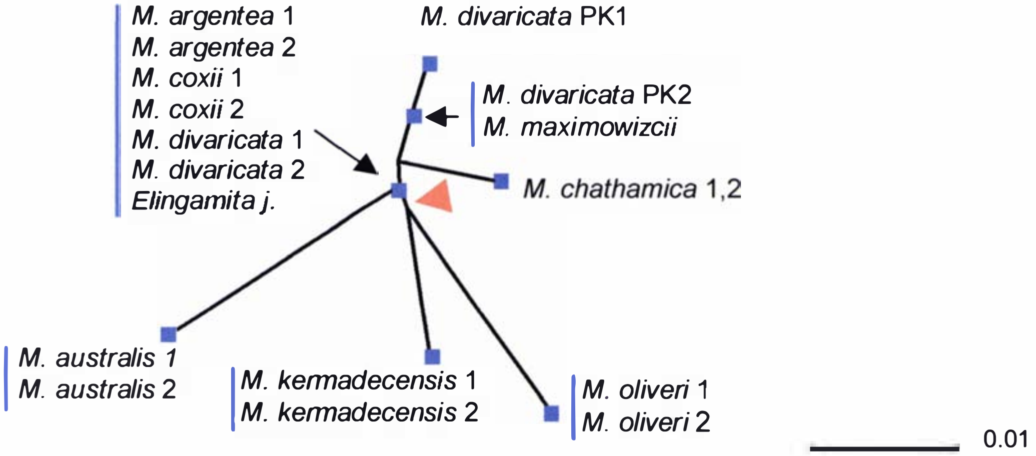


Figure 5.4: Splitsgraph for *Myrsine* spp. and *Elingamita johnsonii* based on sequence of the AFLP derived marker Md-af1. Drawn to scale = 0.01 substitutions per site. The position of the root was inferred from ML analyses presented in section 3.5 and is indicated by a red triangle. Details: Fit = 100.0, nchar = 293(of 331) gaps = 14(off), missing = 24(off), constant = 275 nonparsimony = 282 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Samples are as listed in Table 3.2.

With this genome region, no genetic differentiation was observed for the four species *M. argentea, M. coxii, M. divaricata*, and *Elingamita johnsonii*. These taxa share the same position in the splitsgraph and may well have the ancestral genotype for this gene (if we can assume the same root placement as in Figure 5.3; again the root position is indicated by a triangle in Figure 5.4). On this interpretation, all other species carry derived Ma-af1 genotypes.

Consistent with results from rDNA analysis (see Section 5.2.2.1) is that *M. divaricata* genotypes from the Poor Knights Islands form one lineage with the genotype from the Japanese species *M. maximowizcii*. However, in contrast to rDNA, the splitsgraph made with Ma-af1 indicates a closer affiliation between the genotype of *M. divaricata* PK2 and *M. maximowizcii* rather than between *M. divaricata* PK1 and *M. maximowizcii.*

In the splitsgraph, *M. australis, M. chathamica*, *M. oliveri*, and *M. kermadecensis* are situated on separate evolutionary lineages. Again, the *M. chathamica* genotype shows a close phylogenetic relationship with the Poor Knights *M. divaricata* and *M. maximowizcii*.

The Ma-af1 region of *M. australis*, *M. oliveri* and *M. kermadecensis*, on the other hand, show genetic differentiation from the genotype displayed by *M. argentea M. coxii*, *M. divaricata*, and *Elingamita johnsonii*. All three taxa are situated on relatively long extemal edges.

Consistent with the ITS results, *M. oliveri* and *M. kermadecensis* occur on adjacent edges, although significant genetic differentiation exists between the genotypes of both species. The islands of origins for these species - the Three Kings group which are home to *M. oliveri* and the Kermadec Islands, where *M. kermadecensis* originates are separated by over 1000 km. A similar amount of genetic differentiation between the two species might be expected under a biogeographical model of Tertiary dispersal from the Three Kings group to the Kermadec Islands.

5.2.4 AFLP Derived Marker Md-af9

For phylogenetic analyses of AFLP region Md-af9 in *Myrsine* and *Elingamita,* a 329 bp alignment was obtained for a group of taxa for which the region could be amplified by the use of a one-step PCR protocol (see Section 2.3.3.1). The alignment included 14 accessions from seven species: *M. australis*, *M. coxii*, *M. divaricata* PK, *M. kermadecensis*, *M. maximowizcii*, *M. oliveri* and *Elingamita johnsonii*. All accessions are listed in Table 5.2. Heteroplasmic/ambiguous positions as well as two parsimony uninformative indel positions (one 47 bp and one 1 bp indel) were excluded from analyses.

The splitsgraph made with Md-af9 sequence (Figure 5.5) visualises the very treelike properties of the data, which are supported by a high fit value of 100.

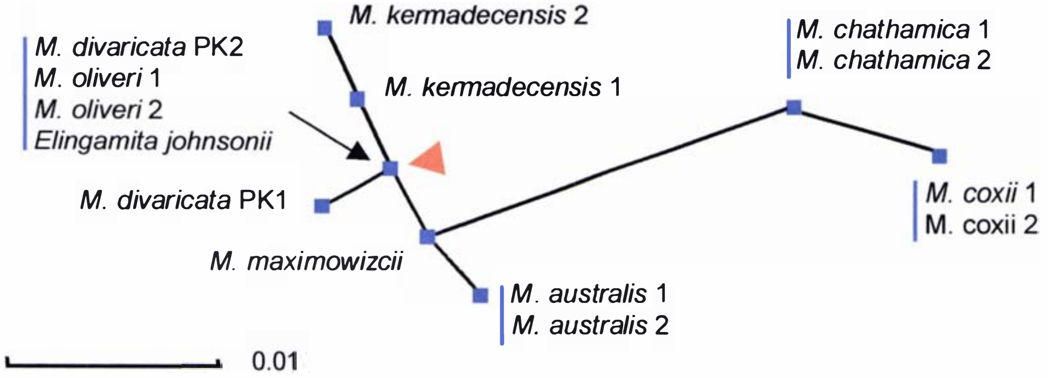


Figure 5.5: Splitsgraph for *Myrsine* spp. and *Elingamita johnsonii* based on sequence of the AFLP derived marker Md-af9. Drawn to scale = 0.01 substitutions per site. The position of the root was inferred from ML analyses presented in section 3.5 and is indicated by a red triangle. Details: Fit = 100.0, characters = 276(of 329) gaps = 48(off), missing = 5(off), constant = 264 nonparsi = 266, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Samples are as listed in Tables 3.2.

Again, if we infer a root position as indicated by a more global analysis of rDNA genotypes (Figure 3.10), we can interpret the splitsgraph as a phylogeny with a root on the internal node as indicated in Figure 5.5. The Md-af9 sequence is identical for *M. oliveri*, *Elingamita johnsonii* and *M. divaricata* PK2 and this genotype may represent the ancestral type. From this node, *M. divaricata* PK1 and *M. kermadecensis* appear as independently diverged lineages. Both individuals analysed for *M. kermadecensis* differ by 1 bp. The phylogeny separates the genotypes of two species from the Chatham Islands by a long edge from its closest relatives. The graph indicates that the *M. coxii* genotype is directly diverged from that of *M. chathamica*. *M. maximowizcii* is located on the node at the basis of the Chatham lineage connecting both groups. *M. australis* forms a separate lineage, which is also born from this node.

5.3 PHYLOGEOGRAPHIC ANALYSIS OF *MYRSINE DIVARICATA* AND *M. ARGENTEA*

The phylogenetic analyses of multiple loci described in the preceding section indicate that the taxonomic species *Myrsine divaricata* is genetically diverse. Furthermore, this genetic differentiation at the DNA level is paralleled by morphological variation between northern (the "Poor Knights type") and southern accessions. In this section of the thesis, molecular studies are described which investigate morphological and genetic variation in *M. divaricata*, the aim being to investigate regional differentiation and the possible importance/occurrence of interspecific hybridisation.

As part of this study, the taxonomic status *M. argentea*, which was treated as part of *M. divaricata* until recently, has been investigated using molecular techniques. The samples used for analyses are specified in the text along with their corresponding herbarium voucher numbers. All other herbarium vouchers are listed in Appendix II.

5.3.1 Investigating Geographical Distribution of *M. divaricata* Genotypes

5.3.1.1 *Ndh*F Sequence Type

Accessions of *M. divaricata* and *M. argentea* were typed at their *ndh*F gene region by DNA sequencing. The distribution of *ndh*F sequence types from different geographical regions has been plotted on the map of New Zealand shown in Figure 5.6.

The distribution of *ndh*F sequence types shows a clear geographic pattern, which separates accessions from northern and southern origins. All samples analysed from the North Island and northern offshore islands (Poor Knights Islands) have *ndh*F sequence type CCA. The northern *M. divaricata* samples share this sequence type with species of *M. australis*, *M. maximowizcii* and *Elingamita johnsonii* (see Section 5.2.1).

In contrast, the majority of samples from the South Island, including 2 accessions of *M. argentea*, and its offshore islands (Stewart Island and the Auckland Islands, which are not included in Figure 5.6) have sequence type CAA. This type was also found in *M. chathamica*, *M. coxii* and in one out of two accessions of *M. oliveri*. Analyses presented in Section 5.2.1 suggest that the southern sequence type may represent the more ancestral type.

M. divaricata accessions from two sites in the South Island show *ndh*F type CCA. The samples were taken from locations were *M. divaricata* and *M. australis* grow side by side. Thus the different *ndh*F sequence type may be a result of hybridisation rather than of geographical variation. This possibility was investigated further (Section 5.3.2).

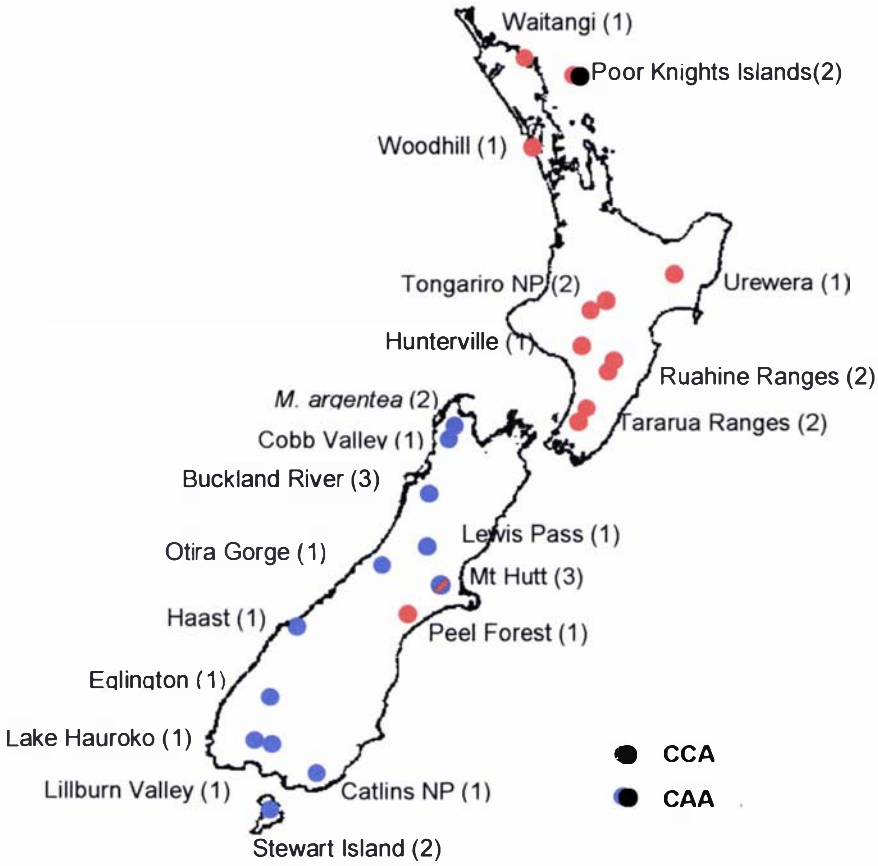


Figure 5.6: Geographical distribution of *ndh*F sequence types CCA and CAA in *Myrsine divaricata* and *M. argentea*. The numbers in brackets give the numbers of samples analysed for the location. Samples: Stewart Island: MyD_ST1, MyD_ST2; Catlins: MyD_5; Lillburn: MyD_6; Hauroko: MyD_7; Eglington: MyD_10; Haast: MyD_11; Otira Gorge: MyD_15; Peel Forest: MyD_18; Mt Hutt: MyD_Hutt1, MyD_Hutt2, MyD_Hutt3; Lewis Pass: MyD_17; Cobb Valley: MyD_21; *M. argentea*: MyArg_1, MyArg_2; Tararua Ranges: MyD_37; MyD_38; Ruahine Ranges MyD_40, MyD_41; Hunterville: MyD_42; Tongariro: MyD_44, MyD_45; Urewera; MyD_46; Woodhill: MyD_47; Waitangi: MyD_48; Poor Knights: MyD_PK1, MyD_PK2. Herbarium voucher numbers for all samples are listed in Appendix II. Abbreviations: **NP:** National Park.

5.3.1.2 Nuclear rDNA Sequence

Phylogenetic analyses were performed on sequence data of nuclear rDNA regions of *M. divaricata* from different geographical origins. The data were analysed under split decomposition.

The splitsgraph (Figure 5.7) visualises the phylogeographic information from a representative group of these taxa. A large number of accessions share a common sequence type. Among these are *M. argentea* and *M. divaricata* samples from the Tararua Ranges, Invercargill, Stewart Island, and Auckland Islands. Visualised as a starlike radiation, some taxa show individual mutations, which are derived from the most common rDNA type (the presumed ancestral type). Individuals from Northland and the Poor Knights Islands, however, form a group of northern accessions that are separated from the more southern accessions. The split between these groups is visualised as a long edge between internal nodes, and indicates that *M. divaricata* samples from Northland and the Poor Knights Islands have a distinct rDNA genotype. This is a result consistent with the possibility of northern and southern ecotypes of *M. divaricata*.

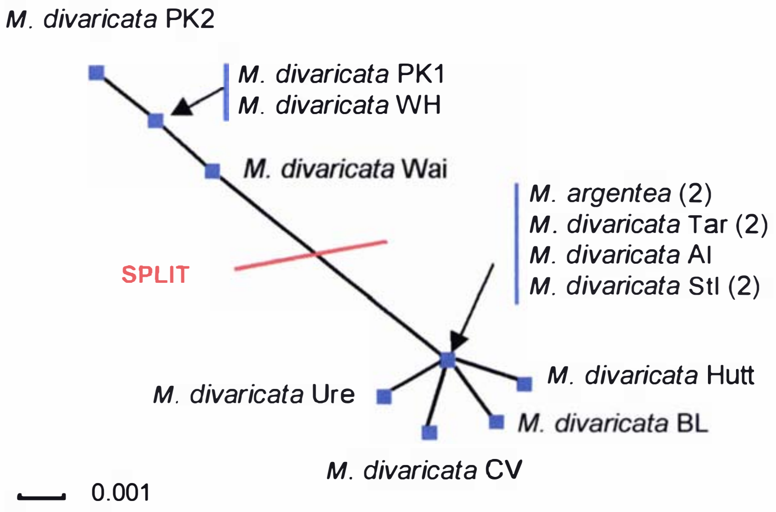


Figure 5.7: Splitsgraph for samples of *Myrsine divaricata* from different geographic locations and *M. argentea* (2 samples) based on ITS1, 5.8S and ITS2 regions of nuclear rDNA. Drawn to scale = 0.01 substitutions per site. Details: Fit = 100.0 ntax = 15 nchar = 633(of 643) gaps = 3(off) miss = 7(off) const = 623 nonparsi = 628 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Abbreviations: **AI:** Auckland Islands, **BL:** Buckland River; **CV:** Cobb Valley; **Hutt:** Mt. Hutt; **PK:** Poor Knights; **StI:** Stewart Island; **Tar:** Tararua Ranges; **Ure:** Urewera; **WH:** Woodhill; **Wai:** Waitangi. Samples: *M. divaricata*: **AI:** MyD_AI; **BL:** MyD_BL2; **CV:** MyD_21; **Hutt:** MyD_Hutt2: **PK:** MyD_PK1, MyD_PK2; **StI:** MyD_ST1, MyD_ST2; **Tar:** MyD_37; MyD_38; **Ure:** MyD_46; **WH:** MyD_47; **Wai:** MyD_48. *M. argentea*: MyArg_1; MyArg_2. Herbarium voucher numbers for all samples are listed in Appendix II.

5.3.1.3 AFLP: 'genetic fingerprints'

Northern and southern accessions of *Myrsine divaricata* were initially profiled with AT-rich and GC-rich primers to target different regions of this species' genome. The AFLP profiles from AT-rich primer combinations (AT-profiles with e.g. *Mse*-CAG and *Eco*-ATA, *Eco* -ATT etc.), showed high levels of genetic variation between individuals. These were considerably more complex and much more variable than profiles produced with more GC-rich primer combinations (GC-profiles). Although the GC-profiles contained fewer polymorphic bands they could be more easily interpreted. An effort was made to decrease the complexity of the AT-profiles by using selective primers with 4 overhanging bases. However, the number of bands could not be reduced significantly. Similar observations have also been made by Vos *et al.* (1995).

The GC-profile displayed in Figure 5.8 shows polymorphic bands, consistent with geographic differentiation. Four polymorphic markers A, B, C, and D are indicated in the profile. Bands A and D appear exclusively in samples from Northland and the Poor Knights Islands. Band B occurs in all Northland samples but also in several individuals from South Island locations. Marker C is present in all samples analysed from the New Zealand North Island and in some individuals from the northern part of the South Island. This band is less frequent in individuals analysed from more southern locations. These results suggest restricted gene flow for *M. divaricata* between different geographic regions.

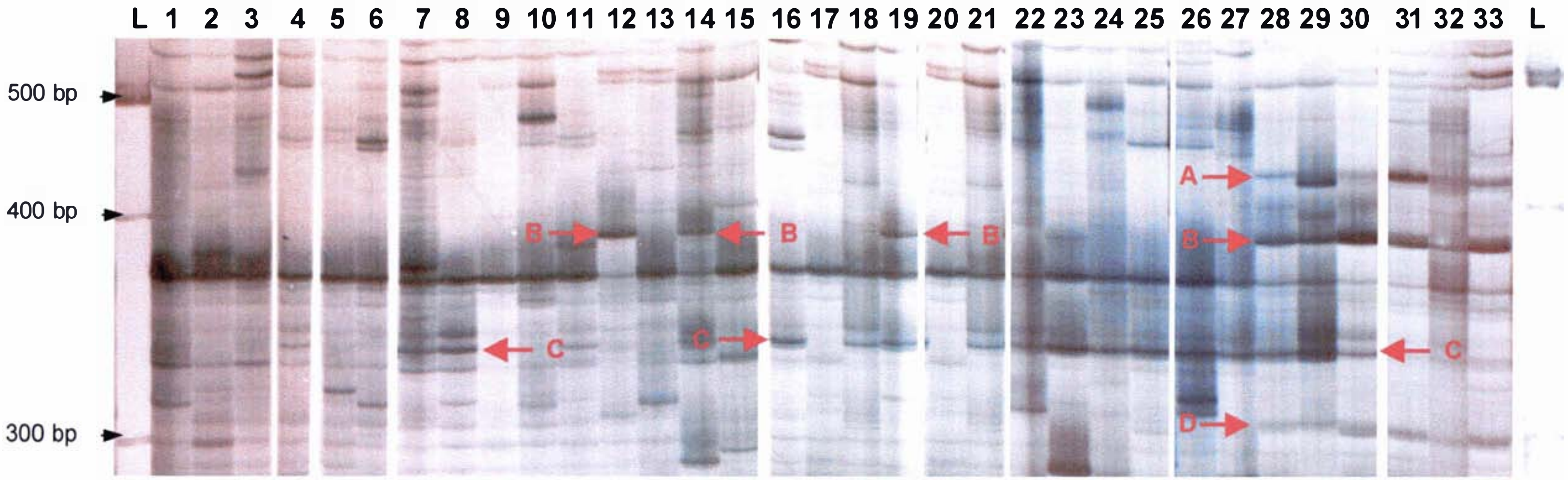


Figure 5.8: AFLP Profiles for *Myrsine divaricata* from different geographic locations in New Zealand and *M. argentea* from Mt. Burnett, Nelson made with promers *Mse*-CAG and *Eco*-AGG. Lanes: **L:** 100 bp ladder; **1:** Auckland Islands: MyD_AI ; **2:** Stewart Island: MyD_ST2; **3:** Stewart Island: Maori Beach MyD_ST1; **4:** Invercargill: MyD_4; **5:** Lill Burn Valley: MyD_6; **6:** Lake Hauroko: MyD_7; **7:** Lake Manapouri: MyD_9; **8:** Eglington Valley: MyD_10; **9:** Makaroa: MyD_10; **10:** Fox Glacier: MyD_13; **11:** Franz Joseph Glacier: MyD_13; **12:** Arthur's Pass: MyD_16; **13:** Lewis Pass: MyD_17; **14:** Peel Forest: MyD_18; **15:** Mt. Oxford: MyD_19; **16:** Maruia Valley: MyD_20; **17:** Cobb Valley: MyD_21; **18:** Cobb River: MyD_23; **19:** Cobb River: MyD_24; **20:** *M. argentea*: MyArg_1; **21:** *M. argentea*: MyArg_2; **22:** Tararua Ranges: MyD_37; **23:** Reporoa Bog: MyD_40; **24:** Ruahine Natl. Forest: MyD_41; **25:** Urewera Natl. Park: MyD_46; **26:** Mt. Ruapehu: MyD_44; **27:** Whakapapa village: MyD_45; **28:** Woodhill: MyD_47; **29:** Waitangi: MyD_48; **30:** Unahi: MyD_49; **31:** Poor Knights Islands: MyD_PK1; **32:** Poor Knights Islands: MyD_PK2; **33:** Poor Knights Islands: MyD_PK3; **L:** 100 bp ladder. The red arrows single out polymorphic bands, A,B, C and D, that may indicate geographic differentiation. Herbarium voucher numbers for all samples are listed in Appendix II.

5.3.2 Investigating Interspecific Geneflow at Sympatric Sites

To examine whether geneflow occurs between species of *Myrsine* in the same geographical regions, individual plants of *M. divaricata* and *M. australis* or *M. salicina*, were analysed from three locations where two of the species occur sympatrically. No morphological hybrids were identified at these sites. The species of all plants used in this section was determined by clear morphological features. For sequence comparisons, four plants were analysed from Lake Rotopounamu, North Island, where *M. australis* and *M. salicina* share the same habitat, five plants were sampled from the Paparoa Range (Buckland track) in the north-west of the South Island, where *M. divaricata* and *M. salicina* occur sympartrically, and five plants were taken from Mt. Hutt in the central region of the South Island, where *M. divaricata* and *M. australis* grow side by side. *Ndh*F sequence types of these plants (see Section 5.2.1) were characterised and analysed alongside sequence data of nuclear rDNA and novel AFLP derived markers.

species	sample	origin	herbarium reference
M. australis	**LR1**	NI: Lake Rotopounamu,	MPN 24701
M. australis	**LR2**	NI: Lake Rotopounamu,	MPN 24702
M. australis	**Hutt1**	SI: Mt. Hutt	-
M. australis	**Hutt2**	SI: Mt. Hutt,	MPN 24706
M. divaricata	**BL1**	SI: Buckland River	MPN 24734
M. divaricata	**BL2**	SI: Buckland River	MPN 24735
M. divaricata	**BL3**	SI: Buckland River	MPN 24736
M. divaricata	**Hutt1**	SI: Mt. Hutt	MPN 24738
M. divaricata	**Hutt2**	SI: Mt. Hutt,	MPN 24739
M. divaricata	**Hutt3**	SI: Mt. Hutt	MPN 24740
M. divaricata	**Hutt4**	SI: Mt. Hutt,	MPN 24741
M. salicina	**LR1**	NI: Lake Rotopounamu,	MPN 24767
M. salicina	**LR2**	NI: Lake Rotopounamu,	MPN 24768
M. salicina	**BL1**	SI: Buckland track	MPN 24771
M. salicina	**BL2**	SI: Buckland track	MPN 24772

Table 5.4: Accession of *Myrsine* used in Section 5.3 for analysis of geneflow at sympatric sites.

5.3.2.1 *Ndh*F Sequence Type

To investigate, whether geneflow of plastid DNA occurs between different species, the *ndh*F gene region was sequenced and compared between samples taken from three sympatric sites.

At Lake Rotopounamu, and in the Paparoa Range, *M. salicina, M. australis* and *M. divaricata* displayed distinct *ndh*F types. However, this was not the case at the Mt. Hutt site, where one out of three samples of *M. divaricata* had the *ndh*F type typical for *M. australis* (Table 5.5).

A: Lake Rotopunamu

	sample	*ndh*F type
M. salicina	BL1	TAG
	BL2	TAG
M. australis	LR1	CCA
	LR2	CCA

B: Buckland River

	sample	*ndh*F type
M. salicina	BL1	TAG
	BL2	TAG
M. divaricata	Bl1	CAA
	BL2	CAA
	BL3	CAA

B: Mt. Hutt

	sample	*ndh*F type
M. australis	Hutt1	CCA
	Hutt2	CCA
M. divaricata	**Hutt1**	**CCA**
	Hutt2	CAA
	Hutt3	CAA

Table 5.5: *ndh*F sequence types of samples taken from locations where two Myrsine species occur sympatrically. **A:** *M. salicina* and *M. australis* from Lake Rotopunamu (North Island) **B:** *M. salicina* and *M. divaricata* from Buckland track (South Island) and **B:** *M. australis* and *M. divaricata* from Mt. Hutt (South Island).

5.3.2.2 Nuclear rDNA Sequence

ITS1, ITS2 and 5.8S regions of nuclear rDNA were sequenced for 14 accessions of *M. australis*, *M. divaricata*, and *M. salicina* listed in Table 5.4. These data were aligned with ClustalX (Thompson *et al.* 1994) using the multiple alignment procedure. The alignment was analysed under SplitsTree 3.1 (Huson, 1998).

The splitsgraph which is shown in Figure 5.9 visualises genetic distances between all taxa. *M. salicina* is separated by a long edge from *M. australis* and *M. divaricata*, which show similar genotypes. While the rDNA regions of all accessions analysed for *M. australis* were identical, sequences for *M. divaricata* differed slightly from each other. Interestingly, the genotype of *M. divaricata* sample Hutt1 is more closely related to those of *M. australis* than to other *M. divaricata* samples. This finding is consistent with analyses of *ndh*F gene sequence for *M. divaricata* Hutt1 and may suggest that this taxon is a genetic hybrid. The nuclear rDNA gives no indication of gene flow between *M. salicina* and either *M. australis* or *M. divaricata*.

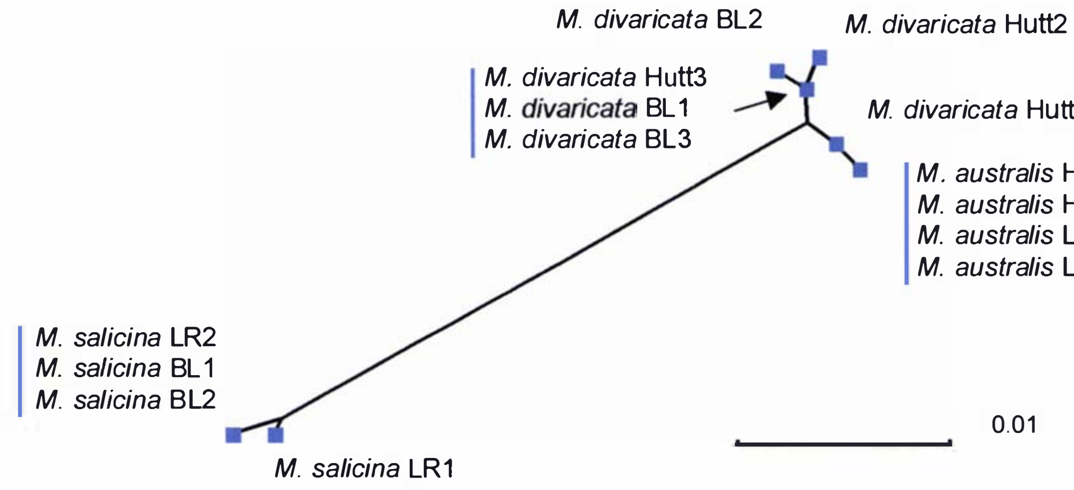


Figure 5.9: Splitsgraph for *Myrsine* spp. based on ITS1, 5.8S and ITS2 regions of nuclear rDNA. Indels used as informative sites. Drawn to scale = 0.01 substitutions per site. Details: Fit = 100 characters = 636(of 641) missing = 4 (off) constant = 612 nonparsimony = 614 -dsplits - hamming. For splitsgraph details see Section 2.4.4.4.

5.3.2.3 AFLP Derived Marker Ma-af1

For *M. australis* and *M. divaricata* from Mt. Hutt the AFLP derived molecular Ma-af1 was analysed in addition to rDNA. The data were aligned and analysed under SplitsTree 3.1 (Huson, 1998). Samples are as listed in Table 5.4.

Sequence analyses based on the AFLP derived marker Ma-af1 corroborate the hypothesis suggested from analyses of *ndh*F and nuclear rDNA that *M. divaricata* sample Hutt1 is indeed a genetic hybrid. Although, based on its morphology, this accession has been clearly identified as *M. divaricata*, its position in the splitsgraph (Figure 5.10) shows closer affiliation with the genotypes of *M. australis* than with other accessions of *M. divaricata*.

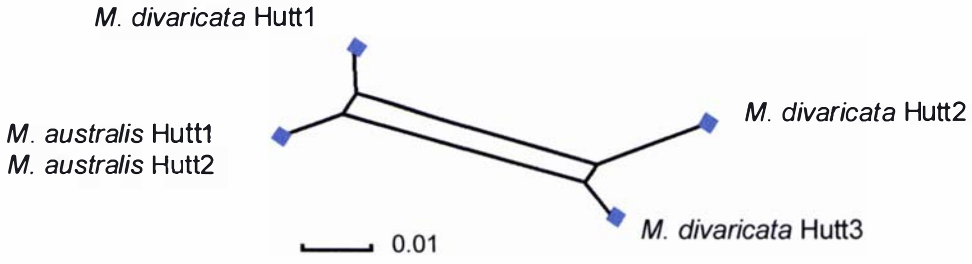


Figure 5.10: Splitsgraph for *Myrsine australis* and *M. divaricata* based on sequence data of the AFLP derived marker Ma-af1. Accessions were sampled at Mount Hutt, South Island where both species occur sympatrically. Drawn to scale = 0.01 substitutions per site. Details: Fit = 100.0 characters = 317(of 331) gaps = 14(off) constant = 295 nonparsimony = 304 -dsplits -hamming. For splitsgraph details see Section 2.4.4.4.

5.4 DISCUSSION

5.4.1 Phylogenetic Relationships within Genus *Myrsine*

Although *ndh*F sequence analyses revealed little phylogenetic resolution between New Zealand and overseas species of Myrsinaceae, the region contained phylogenetic information that allowed three groups of taxa to be distinguished. Group one was formed by the Australian species of *Rapanea*, the African taxon *M. africana*, and *M. salicina*, *M. kermadecensis* from New Zealand and its offshore islands. Inference from outgroup placement suggested that these taxa displayed the ancestral sequence type. The second group showed a derived sequence type and included *M. argentea* from the New Zealand mainland, *M. coxii* and *M. chathamica* from the Chatham Islands. A further basepair change was observed in group three with *M. australis*, *M. maximowizcii*, and *Elingamita johnsonii*. *Ndh*F sequences are consistent with results from nuclear rDNA presented in Section 3.5 that suggest that species in group three are the most recently derived taxa.

Interestingly, accessions of *M. oliveri* and *M. divaricata* showed alternative sequence types. While *M. oliveri* was found in groups one and two, accessions of *M. divaricata* were found in group two and group three. Hybridisation may be one explanation for this observation.

Nuclear rDNA and AFLP derived markers used to investigate interspecific relationships between closely related species of Myrsinaceae contained different phylogenetic information. The low correlation between nuclear rDNA region and the novel markers Ma-af1 and Md-af9 has already been demonstrated in Section 4.2.1.3. Nevertheless, there was remarkably little contradiction in the phylogenetic signals, it was rather the amount of genetic differentiation between taxa that varied between the phylogenies based on the three different direct sequencing loci.

Phylogenetic information in nuclear rDNA and AFLP derived markers is consistent in suggesting that both species from the Chatham Islands *M. chathamica* and *M. coxii* are genetically distinct from each other. This molecular result corroborates the high amount of morphological differentiation between the two taxa. While the nuclear rDNA region reveals genetic distinction of *M. coxii* (Figure 5.3), the Ma-af1 region indicates some differentiation between *M. chathamica* and other New Zealand taxa (Figure 5.4). However, both species have the same *ndh*F type and form an adjacent lineage in the Md-af9 phylogeny (Figure 5.5). This may indicate a monophyletic origin of *Myrsine* on the Chatham Islands and/or hybridisation events after the speciation of both taxa.

Close phylogenetic relationships were also recovered between the Kermadec Island *Myrsine* *M. kermadecensis* and the Three Kings species *M. oliveri*. Results from all genome loci are consistent in suggesting a common ancestor of these taxa. The two accessions investigated for *M. oliveri* display two different *ndh*F sequence types although the accessions analysed for *M. oliveri* do not originate from isolated geographical locations. This indicates that both genotypes occur sympatrically. Although the individuals used for analyses were not morphological hybrids, the different genotypes may result from geneflow, i.e. long distance hybridisation between island and mainland species of *Myrsine*. Further studies are needed to verify this hypothesis and assess the amount of possible geneflow between *M. oliveri* and other Myrsinaceae (e.g. *M. australis*)

Plastid DNA, nuclear ITS regions and Ma-af1 sequence of *Elingamita johnsonii* showed close affiliations to the *M. divaricata* - *M. australis* group. However, splitsgraphs from nuclear ITS regions and the AFLP derived marker Md-af9 also revealed signals that align this species with *M. oliveri* and *M. kermadecensis*. These results may suggest a hybrid origin of this taxon. Such a hypothesis would be consistent with the observation of possible hybridisation between *M. oliveri* and e.g. *M. australis*. Breeding experiments may help to further investigate the hypothesis of a hybrid origin of *Elingamita*. Furthermore, DNA analyses presented in this thesis clearly suggest that the generic status of this New Zealand endemic needs to be reconsidered and *Elingamita* G.T.S. Baylis should be included in *Myrsine* L.

5.4.2 Phylogenetic Relationships for *M. divaricata* and *M. argentea*

Morphological and ecological studies have suggested differentiation of *Myrsine divaricata* into several ecotypes with preference to distinct climatic and ecological environments (Heenan, pers. comm.). These ecotypes include *Myrsine* from Mt. Burnett, a very local form, which was recently described as a new species *M. argentea* (Heenan and de Lange, 1998) and the Poor Knights *Myrsine* which also shows distinct morphological features but is still included in *M. divaricata*. The occurrence of these ecotypes as well as molecular studies presented previously in Section 3.5 of the current thesis have pointed out the need for more detailed investigations to elucidate interspecific and intraspecific relationships within this plant lineage.

Interestingly, neither DNA sequencing nor AFLP revealed genetic differentiation between *M. divaricata* and the Mt. Burnett species *M. argentea*. Sequences for *M. divaricata* from a southern New Zealand origin and *M. argentea* are identical in all markers. The fact that the Md-af9 regions of both taxa could not be amplified using the AFLP derived primers and standard PCR protocols, is also consistent with a great genetic similarity. The sequence analyses presented in this chapter were not able to distinguish between *M. argentea* and South Island accessions of *M. divaricata*. The results suggest a very recent morphological and ecological differentiation and do not explicitly support the classification that *M. argentea* and *M. divaricata* are two different species.

In contrast, molecular analyses reveal genetic variation between northern and southern accessions of *M. divaricata*. Samples from the Poor Knights Islands are distinct from mainland species in all sequencing loci. This result corroborates morphological analyses and suggests that Poor Knights *M. divaricata* and mainland taxa are separate evolutionary lineages. This result is consistent with observations made by P. Heenan (pers. comm.) who suggested that the Poor Knights *Myrsine* should be excluded from *M. divaricata* as a separate species.

Furthermore, results from all gene regions suggest a close relationship between the Poor Knights *Myrsine* and the Japanese taxon *M. maximowizcii*. This relationship is most parsimoniously explained by a recent trans-oceanic dispersal event from New Zealand to Japan.

5.4.3 Geographic Structure in Populations of *Myrsine divaricata*

Molecular studies in this chapter used direct sequencing loci and AFLP fingerprinting to assess correlation between genetic information and geographic origin of *M. divaricata* samples. The studies revealed genetic differentiation between accession of *Myrsine divaricata* from different geographical locations.

Morphological and ecological analysis suggested the existence of a northern and a southern ecotype of *M. divaricata*. While the 'northern ecotype' occurs in the warmer lowland regions of the northern North Island, the 'southern type' prefers the colder climates of higher altitudes and more southerly regions of New Zealand. The northern type includes the Poor Knights *Myrsine* which is morphologically distinct from the mainland type by frequently non-divaricating branches and bigger leaves. Sequence analyses of direct sequencing loci as well as AFLP fingerprints showed that the northern ecotype on the New Zealand mainland groups has close affiliations with the Poor Knights *Myrsine*.

Phylogenetic analyses suggest a recent dispersal from the northern mainland to the Poor Knights group. The morphological differentiation between mainland and island taxa may be a result of a small founder population and the 'founder effect' (Page and Holmes, 1998). This hypothesis is supported by the observation that non-divaricating growth forms are also found on the New Zealand mainland but they are less common. AFLP fingerprints revealed that specific DNA patterns observed in the Poor Knights *Myrsine* are also present in some mainland accessions. These patterns might also indicate the occurrence of geneflow between individuals from different ecological regions and suggest that genetic differentiation between these two taxa is still an ongoing process.

Molecular studies on *M. divaricata* also revealed genetic differentiation between accessions from different geographic regions. Investigations on *ndh*F sequence types showed a split across Cook Strait - between samples from the North Island and the South Island. However, the low amount of genetic differentiation suggests a very recent origin of this genetic structure. As hybridisation has been found common between *M. divaricata* and *M. australis*, the presence of two different *ndh*F types may be the result of geneflow between these two species.

Future studies may be conducted to investigate the amount of geneflow between *M. divaricata* from different geographic regions. The presence or absence of the AFLP markers shown in Figure 5.8 will provide a useful tool for statistical analyses. A larger number of samples than are included in the present study would need to be analysed to minimise statistical error and assess whether alleles are being fixed in *M. divaricata* populations.

5.4.4 Hybridisation

Results in Section 5.3.4 clearly show that geneflow is common between *M. divaricata* and *M. australis* where both species grow side by side. No geneflow was recorded between *M. salicina* and *M. divaricata*. Interestingly, even at sympatric sites with *M. divaricata* and *M. australis* hybrids with intermediate morphological phenotypes are rare and the introgression is only observed in the genotype. Phenotypic dominance of the *M. divaricata* type may explain this observation. Although genetic hybrids cannot be identified by their morphological features, *M. divaricata* accessions from sympatric sites (e.g. Mt. Hutt in the South Island and Mt. Holdsworth in the North Island) show greater variation in leaf size than do plants from other location. In these regions, no spatial patterns of genotypic or phenotypic distribution ('hybrid zones') were identified that may indicate that hybrids have an evolutionary advantage in a particular ecological environment ('environment-dependent selection', Arnold, 1992). This observation suggests that hybridisation between *M. divaricata* and *M. australis* may be based on environment-independent selection (Barton and Hewitt, 1985).

The general consequences of hybridisation and introgression in plants have been well debated and it is now widely recognised that hybridisation is an important source of novel, stable evolutionary lineages (Abbott, 1992; Arnold, 1992). Indeed, molecular studies on hybridisation have concluded that hybrid formation may facilitate rapid speciation (Rieseberg, 1997; Ungerer *et al.*, 1998). A frequent occurrence of hybridisation has been reported in many New Zealand herbaceous alpine genera e.g. *Ranunculus* (Fisher, 1965), *Leptinella* (Lloyd, 1972), *Epilobium* (Raven, 1972), and it has been suggested that this phenomenon has been of considerable evolutionary importance in the relatively young environments of the New Zealand alpine zone (Raven, 1973; Winkworth, 2000).

Wardle (1988b) speculated on the importance of long-distance hybridisation between several New Zealand tree species via pollen flow. He noticed that in undisturbed forests adult hybrid species were rare. *Myrsine* offers an example for hybridisation within the New Zealand woody flora at sympatric sites. Future molecular studies are needed to determine the importance of hybridisation as a mechanism for speciation in New Zealand forests.

5.4.5 Evolution of Morphological Features

The hypothesis of recent speciation within the *M. divaricata* - *M. australis* group implies that some extraordinary morphological characters displayed by members of this group are also of recent origin. This includes the flower morphology of the rare species *E. johnsonii*, the evolution of different leaf types and the divaricating growth form of *M. divaricata*.

Plants with small leaves and divaricating growth forms are relatively frequent throughout New Zealand's forests (Wardle, 1963; Kuschel, 1975). While *Myrsine divaricata* and *M. coxii* display this feature during the adult stage, numerous species like e.g. *Nothofagus fusca* or *Pennantia corymbosa* show this characteristic during juvenile stages only. It has been suggested that this particular morphological feature was an advantageous protection against Moa browsing (Stevens *et al.*, 1995). In contrast, Cockayne (1911) considered that extremely arid Pleistocene climates evoked these morphological features which he regarded as xeromorphic. He suggested that the divaricating habit was impressed by past 'steppe climates' and could not have arisen in response to the environment of present forest interiors. Especially considering the juvenile divaricating, forms Wardle (1963) proposed that these morphological forms are adaptations to 'still existing, albeit dryish, forest environments'. Juvenile forms with xeromorphic small leaves are less susceptible to injury by drought and are in fact more frequent in drier forest environments. Wardle seemed sceptical about how much of this evolution could have taken place during the Pliocene and Quaternary and argued that it was unlikely that 'a primitive and isolated species like *Podocarpus spicatus* could have evolved its typically divaricating juvenile form within this time'.

As all other species of *Myrsine* that appear basal in the phylogeny carry large leaves it can be hypothesised that the common ancestor of the New Zealand Myrsinaceae was in fact a large-leafed form. Phylogenetic analyses on *Myrsine* recovered low genetic differentiation between different morphological forms which suggests that the small-leafed species like *M. divaricata* (including *M. argentea*) and *M. coxii* have evolved relatively recently - probably during the Quaternary. This result is more consistent with the hypothesis proposed by Cockayne (1911) who suggested a recent origin of divaricating growth forms in New Zealand. It will be interesting for future molecular studies whether similar results are obtained from investigations of divaricating species in other genera.

With *M. chathamica* and *M. coxii* one large-leafed and one small-leafed form is present on the Chatham Islands. The hypothesis of a monophyletic origin of *Myrsine* on the Chatham Islands would imply a recurrent evolution of either small or large leaves. While a recent dispersal event of a small-leafed (*M. divaricata*-like) ancestor would imply a subsequent evolution of large leaves, dispersal of a large leafed form would suggest an independent evolution of a small leafed species. In contrast, if a polyphyletic origin of *Myrsine* is assumed on the Chatham Islands, the phylogenetic relationship between both species may be explained by geneflow.

A close phylogenetic relationship is also observed between the large-leafed *M. maximowizcii* and the small-leafed Poor Knights form of *M. divaricata*. The phylogenetic relationships and the small genetic distance between both taxa suggest recent dispersal event of a small-leafed *M. divaricata*-like ancestor from New Zealand to Japan. This implies a subsequent evolution of large leaves in *M. maximowizcii* from a small-leafed ancestor. In this case, because of the geographic distance between both species, long-distance hybridisation between two more anciently diverged species is rather unlikely. This result supports the hypothesis of a recurrent evolution of different leaf morphologies in *Myrsine*. A recurrent evolution of leaf morphologies may not be surprising considering that conspicuous morphological changes may be caused by relatively small genetic changes. This observation has been made by Comes and Kadereit (1998), who found evidence from both natural populations and domesticated crops that in some cases, fundamental changes in plant morphology can arise from relatively small amounts of genetic change.

The low level of genetic diversity in morphologically diverse species of Myrsine is consistent with results from other New Zealand and overseas plant groups. Similar patterns of evolution have been described for e.g. *Hebe* (Wagstaff & Garnock-Jones, 1998), *Gingidia* (Mitchell *et al.*, 1998), *Myosotis* (Winkworth, 2000), and *Ranunculus* (Lockhart *et al.*, in press) from New Zealand, as well as from *Dendroseris* (Sang *et al.*, 1994) from the Juan Fernandez Islands and the Silversword alliance (Robichaux *et al.*, 1990; Baldwin, 1992) from the Hawaiian Islands. In many cases the available evidence suggests that diversification of plants particularly in insular habitats has been adaptive and is characterised by elevated speciation rates, increased morphological diversity, and marked differentiation of ecological characteristics (Givnish, 1997).

CHAPTER 6

Intraspecific DNA Studies on *Nothofagus menziesii* and *N. moorei*

6.1 INTRODUCTION

Biogeographic relationships between species of *Nothofagus* and their distribution patterns have continued to puzzle botanists, palaeobotanists, phylogeneticists, molecular systematists, and geologists alike. The genus itself exhibits a disjunct distribution across oceans (Section 3.6), but discontinuities are also found at the species level, in New Zealand as well as in Australia distributions of *Nothofagus* species exhibit large disjunctions between individual populations. Explanations for the isolated appearance of species and populations have been seen as a consequence of either dispersal or disruption of a once continuous distribution by historical events (vicariance). Ongoing debates have revolved around the question of which hypothesis may be true for *Nothofagus*.

6.1.1 Southern Beech in New Zealand

In New Zealand the four endemic *Nothofagus* species, *N. fusca* (red beech), *N. solandri* (mountain beech/black beech), *N. truncata* (hard beech) and *N. menziesii* (silver beech) share similar distribution patterns. All *Nothofagus* species are confined to the mountainous region of the country and find their northern limit at ca. 38° southern latitude. Southern beech occurs in central North Island, lower North Island, northern South Island (Nelson) and southern South Island (Southland). Populations of all species are separated by distribution gaps in the North Island (northern beech gap) and in the South Island (southern beech gap); Cook Strait disconnects populations on both islands.

In mountainous regions south of 38° southern latitude all species are absent from areas clearly within the range of ecological tolerance (Leathwick, 1998). By using projections of ecological variables to analyse the degree of equilibrium between *Nothofagus* distributions and environment, Leathwick found that disjunctions are not explicable in terms of the environmental factors he used. He concluded that distribution patterns more likely reflect the effects of historic displacement (e.g. by glaciation) coupled with poor abilities of *Nothofagus* to invade other forest types.

Burrows (1965) and Wardle (1967, 1988a) elaborated the hypothesis that glaciation and cold climates during the Pleistocene may be responsible for disjunctions of beech across the two floristic gaps in New Zealand (see Section 1.5.6). The authors proposed that during cold climates, forest species were eliminated from large areas but survived in relatively few glacial refugia in the North and South Island from which they expanded during warmer interglacial stages. Coherently, beech, which disperses slowly, has been in the process of re-occupying its potential habitat after the end of the last glacial maximum.

Haase (1992), however, argued against an incisive reduction of population size during the Pleistocene because "...the calculated rates of spread are orders of magnitudes too small to account for the distances which had to be covered in post glacial times, had the genus - together with many other forest species - become extinct on a regional scale during the maximum of the last (Otira) glaciation".

An alternative explanation was proposed by McGlone (1985) and Rogers (1989), who introduced the hypothesis that Cenozoic geological changes account for present day distribution patterns. The hypothesis enjoins a pre-glacial disruption of a formerly continuous plant cover as a result of ecological and climatic changes, but also acknowledges biological competition as a restraint for range expansion. Both authors agree to there having been reduction of *Nothofagus* over wide areas during Pleistocene ice ages but argue that small stands were likely to have survived in nearly all areas with sufficient relief to provide suitable sheltered habitats.

6.1.1.1 *Nothofagus menziesii*

Nothofagus menziesii (Hook f.) Oerst. is the only New Zealand member of subgenus *Lophozonia* (see Section 3.6). It does not hybridise with other southern beeches on the archipelago, which in addition to its poor dispersal abilities makes it ideal for studying forest dynamics. Silver beech has small, serrate, evergreen leaves and is, like other southern beeches, monoecious and wind-pollinated. The species is widespread in montane and subalpine regions, attaining its upper limits corresponding to the timber line, which is between 1430 m at East Cape (Mt. Hikurangi), 1200-1280 m in the lower North Island (Tararua Ranges) and Nelson and 910-980 m in Southland (Wardle, 1967). It descends to sea level in the humid south western districts of Fjordland and elsewhere, when nutrient poor or waterlogged soils reduce competition from other tree species (Wardle, 1984). *Nothofagus menziesii* often occurs in monotypic stands or in pure *Nothofagus* forests, typically together with one or

Figure 6.1: Distribution of *Nothofagus menziesii* in New Zealand

two of the other New Zealand beech species. Other beech forest communities include montane *Nothofagus*-conifer forests with podocarps and *Nothofagus*-conifer-*Beilschmiedia tawa* forests (Wardle, 1967).

Nothofagus menziesii, like the other New Zealand beech species, is absent from large areas clearly within its range of environmental tolerance (Figure 6.1), notably, Mt. Taranaki, central Westland, and Stewart Island (Wardle, 1967; Leathwick, 1998).

6.1.2 Southern Beech in Australia

Australia hosts three *Nothofagus* species from two subgenera: *N. moorei* (antarctic beech) and *N. cunninghamii* (myrtle beech) from subgenus *Lophozonia* and *N. gunnei* (deciduous beech) from subgenus *Fuscaspora*. Southern beech dominates the cool temperate rain forest along the eastern ranges of mainland Australia and Tasmania. *Nothofagus moorei*, a member of *Lophozonia*, is the northernmost species with populations in Queensland and New South Wales. *N. moorei* is geographically isolated by almost 1000 km from its closest Australian relative *N. cunninghamii*. This comparatively small-leafed species occurs further south in Victoria and - across Bass Strait - in Tasmania. *N. gunnii* the only extant *Fuscaspora* species is endemic to Tasmania. *N. gunnii* and *N. moorei* are generally restricted to higher altitudes above 500 m, *N. cunninghamii* being the most versatile Australian species, can also reach sea level.

Today, southern beech is a minor component of the Australian flora. It has been suggested that the present climate is unfavourable for *Nothofagus* and on the main land beech dominated rain forests are regarded as relictual (Read and Brown, 1996). Pollen studies showed that *Nothofagus* forests were more abundant during the mid-Tertiary when warmer and wetter climates dominated and distributions reached northern Queensland and Western Australia (Macphail *et al.*, 1994; Hill *et al.*, 1996). The distribution of *Nothofagus* declined at the same time that there was a general contraction of rain forest, as the Australian climate became more seasonal, cooler and drier. While warm-adapted *Brassospora* beeches disappeared from the Australian landmass, *Fuscaspora* and *Lophozonia* beeches became restricted to moister habitats. Major changes in distributions of southern beech appear to have occurred prior to the onset of the Pleistocene glacial stages, although some regional extinction continued to occur into the Pleistocene (Read and Brown, 1996).

6.1.2.1 *Nothofagus moorei*

Nothofagus moorei (antarctic beech) is an evergreen tree species, which can reach up to 40 m in height. The long ovate-elliptical leaves (4-6 X 3-4 cm) have finely toothed margins and are arranged alternately on the branches. Narrow stipules are found at the base dark green leaf.

Antarctic beech has a disjunct distribution in the Border Ranges of New South Wales (Figure 6.2). Its southernmost population is at Barrington Tops, north-west of Newcastle in New South Wales, its northernmost distribution is just beyond the McPherson Ranges (Lamington National Park) in the south of Queensland. Antarctic beech forests are found at altitudes of 500 to 1500 metres in cool temperate rainforests with 1750mm to 3000 mm of annual rain. It often forms stands with a uniform canopy of just 2 or 3 species and little undergrowth (Bale and Williams, 1993).

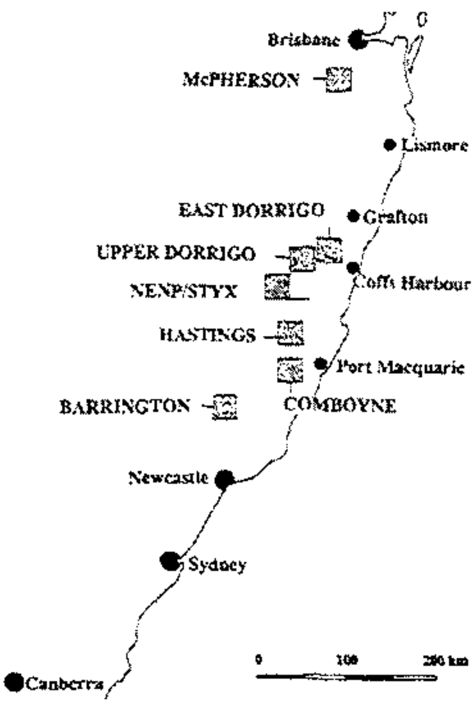


Figure 6.2: Distribution of *Nothofagus moorei* in Australia. Abbreviations: NENP: New England National Park.

6.1.3 Aims and Expectations

To date no studies have been presented that investigate molecular properties of *Nothofagus menziesii* and *N. moorei* populations. Investigations that have used molecular information to study the nature of the discontinuous distribution of *Nothofagus* and other species in New Zealand are also limited. Exceptions are isozyme analyses on New Zealand hard beech (*Nothofagus truncata*) by Haase (1992) and rimu (*Dacrydium cupressinum*) by Hawkins and Sweet (1989). These studies found low genetic variation between populations of these forest genera. Haase (1992) concluded that genetic diversity might have been lost by random genetic drift in populations of reduced size. This could have happened during Pleistocene glacial stages, which caused range contractions. Further, this loss of genetic variation was not compensated for by the occurrence of new mutations during range expansion of populations during subsequent interglacial warm periods.

In the present thesis work, DNA markers are to be used for population genetic analysis of New Zealand silver beech (*N. menziesii*) and Australian antarctic beech (*N. moorei*). The present work aims to contribute findings from intraspecific comparisons to elucidate distribution patterns of the New Zealand species *Nothofagus menziesii* and the Australian species *Nothofagus moorei*.

6.2. *NOTHOFAGUS MENZIESII*

6.2.1 Sequence Analyses of Nuclear DNA

To resolve phylogenetic relationships between populations of *N. menziesii* nuclear rDNA and Nm-af2, sequence was determined and aligned for the same ten taxa. All individuals characterised for sequence variation were sampled from geographically isolated populations of *Nothofagus menziesii* and were homozygous for allele B at the Nm-af2 locus (see Section 4.2.2.1). Protocols for DNA extraction, PCR and sequencing are described in Section 2.3. Voucher numbers for all silver beech samples analysed in this chapter are listed in Appendix III alongside with a map depicting *N. menziesii* sampling areas in New Zealand.

DNA alignments revealed sequence variation amongst individuals in ITS regions of nuclear rDNA and the Nm-af2 region. In both alignments, and at all sites showing sequence variation, heteroplasmic base signals occurred in one or more taxa. Character state patterns have been detailed in Tables 6.1 and 6.2. The phylogenetic information in these data has been visualised under split decomposition (Section 2.4.4.4). Splitsgraphs are shown in Figures 6.3 and 6.4. These indicate extensive reticulation, which would be consistent with nuclear gene exchange between now geographically isolated populations and/or retention of ancestral polymorphisms by some populations. No relationship is apparent between sequence data and geographic location of the sourced samples.

region	sample	origin	p101	p207	p428	p467	p491	p562
central North Island	N1-1	Te Aroha	T	T	C	A	G	T
	N2-1	East Cape	T+C	T	C+T	G	G	G
	N5-5	Waihaha River	T+C	T+A	C	G	G	G+T
lower North Island	N9-9	Tararua Ranges	T+C	T	C+T	A	G+A	G+T
	N10-1	Rimutaka Ranges	T+C	T	C+T	G	G	G+T
	N11-1	Aorangi Ranges	T+C	T	C+T	A	G	G+T
northern South Island	S1-1	Abel Tasman	T+C	T	C+T	G	G+A	G
	S1-9	Arthur Range	T+C	T	C+T	G+A	G	G+T
	S3-12	Buller River	T+C	T	C	G+A	G	G+T
southern South Island	S7-5	Haast	T+C	T	C	G+A	G	G+T
	S14-3	Longwood Range	T+C	T	C+T	G	G	G
	S16-12	Catlins NP	T+C	T	C+T	G	G	G

Table 6.1: Variable sites from aligned rDNA sequence of *N. menziesii* population samples. Herbarium voucher numbers for all samples are listed in Appendix III. Abbreviations: p = position.

region	sample	origin	p 17	p 115	p 154	p175	p 242
central North Island	N1-1	Te Aroha	C+G	A	T	T	C
	N3-1	East Cape	C	A	G	T	G
lower North Island	N9-9	Tararua Ranges	C	A	T	T	C
	N10-1	Rimutaka Ranges	C	A	T	T	C
	N11-1	Aorangi Ranges	C	A	G	-	C
northern South Island	S1-9	Arthur Range	C	A	T	T	C
	S3-12	Buller River	C+G	A	T	T	C
southern South Island	S7-5	Haast	C	G+A	T	T	C
	S14-3	Longwood Range	C+G	G+A	G+T	T	C
	S16-12	Catlins NP	C+G	G+A	G+T	T	C

Table 6.2: Variable sites from aligned Nm-af2 sequence of *N. menziesii* population samples. Herbarium voucher numbers for all samples are listed in Appendix III. Abbreviations: p = position.

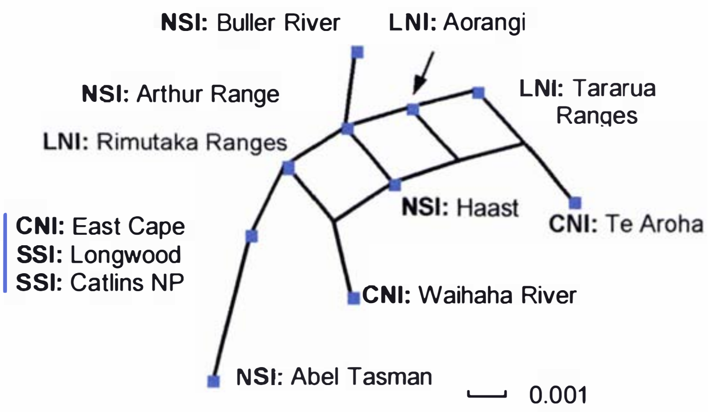


Figure 6.3: Splitsgraph for *N. menziesii* population samples based on nuclear rDNA sequence. Scale: 0.001 substitutions per site. Details: Fit = 100.0, characters = 579 (of 580), missing = 1 (off), constant = 569, non-parsimony = 574, -dsplits-hamming. For splitsgraph details see Section 2.4.4.4. Abbreviations: **CNI:** Central North Island; **LNI:** Lower North Island; **NSI:** Northern South Island; **SSI:** Southern South Island. Samples are as listed in Table 6.1.

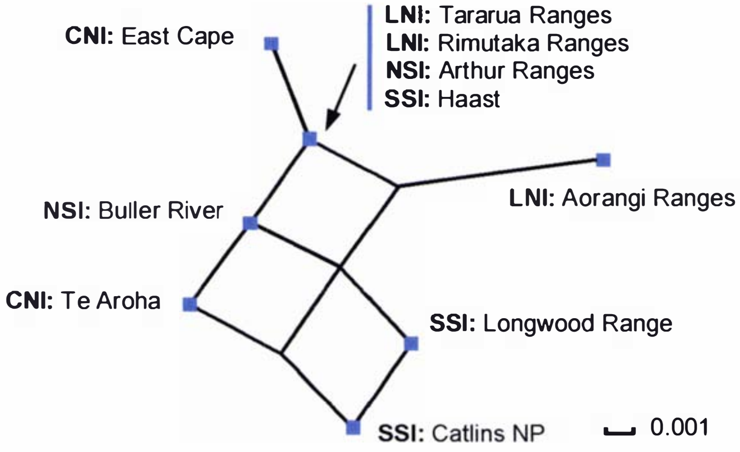


Figure 6.4: Splitsgraph for *N. menziesii* population samples based on Nm-af2 sequence. Scale: 0.001 substitutions per site. Details: Fit: 100.0, characters = 319, constant = 312, non-parsimony = 315, -dsplits-hamming. For splitsgraph details see Section 2.4.4.4. Abbreviations: **CNI:** Central North Island; **LNI:** Lower North Island; **NSI:** Northern South Island; **SSI:** Southern South Island. Samples are as listed in Table 6.2.

6.2.2 Sequence Analyses of Plastid Sequence

Plastid sequences from the *trn*L intron (443 bp) and *trn*L-*trn*F intergenic spacer (436 bp) were determined for all ten individuals of *N. menziesii* characterised for rDNA and Nm-af2. Sequences were also obtained for additional 19 taxa sampled across the distribution range. The sequence data from the two loci was analysed as a combined data set. The monospecific alignments were free of ambiguous or heteroplasmic signals, they contained no indels and revealed low genetic diversity. For *N. menziesii* one base pair difference was detected within the *trn*L-*trn*F spacer region and one within *trn*L intron sequence. The variable sites are shown in Table 6.3.

region	sample	origin	trnL intron p183	trnL-trnF spacer p128
central North Island	N1-1	Te Aroha	C	A
	N3-1	East Cape	C	A
	N2-8,	East Cape	C	A
	N2-9	East Cape	C	A
	N3-4	Urewera	C	A
	N4-2	Mamaku forest	C	A
	N5-5	Waihaha River	C	A
lower North Island	N8-1	Tararua Ranges	C	C
	N8-4	Tararua Ranges	C	C
	N8-6	Tararua Ranges	C	C
	N9-2	Tararua Ranges	C	C
	N9-10	Tararua Ranges	C	C
	N10-1	Rimutaka Ranges	C	C
	N11-1	Aorangi Ranges	C	C
northern South Island	S1-1	Abel Tasman	C	A
	S1-9	Arthur Range	C	A
	S1-10	Arthur Range	C	A
	S2-4	Bryant Range	C	A
	S3-5	Lake Rotoiti	C	A
	S3-12	Buller River	C	A
	S4-6	Maruia-Marble Hill	C	A
southern South Island	S6-1	Mt. Cook	A	A
	S7-5	Haast	A	A
	S14-3	Longwood Range	A	A
	S15-6	Catlins NP	A	A
	S15-7	Catlins NP	A	A
	S15-11	Catlins NP	A	A
	S15-12	Catlins NP	A	A
	S16-12	Catlins NP	A	A

Table 6.3: Variable sites from *N. menziesii trn*L intron and *trn*L-*trn*F intergenic spacer. Herbarium voucher numbers for all samples are listed in Appendix III. Abbreviation: p = position.

In contrast to the results from nuclear loci, base pair differences within the alignment of *N. menziesii* plastid sequences identify taxa to the geographic regions from where they were sourced. This observation is illustrated in Figure 6.6. Phylogenetic analyses using outgroup placement with other *Lophozonia* and *Fuscaspora* beeches (Section 3.6.2) reveal that samples from northern South Island populations and central North

Island populations share the ancestral DNA type (see Figure 6.5). Southland and lower North Island samples display an independently derived sequence type.

If it can be assumed that *Nothofagus* chloroplast DNA is maternally inherited as in most angiosperms, then these results suggest limited or no seed dispersal between the four disjunct provinces.

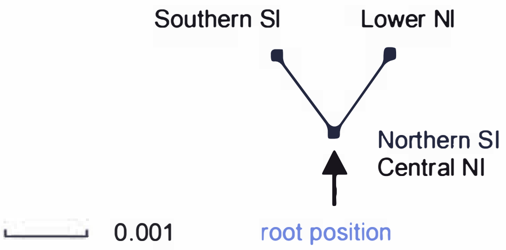


Figure 6.5: Splitsgraph for *N. menziesii* based on plastid sequence data indicating the position of the root. Scale: 0.001 substitutions per site. Details: Fit: 100.0; charcters = 845 (of 937) gaps = 92 (off); constant = 808; non-parsimony = 814; -dsplits-hamming. For splitsgraph details see Section 2.4.4.4. Abbreviations; **NI:** North Island; **SI:** South Island.

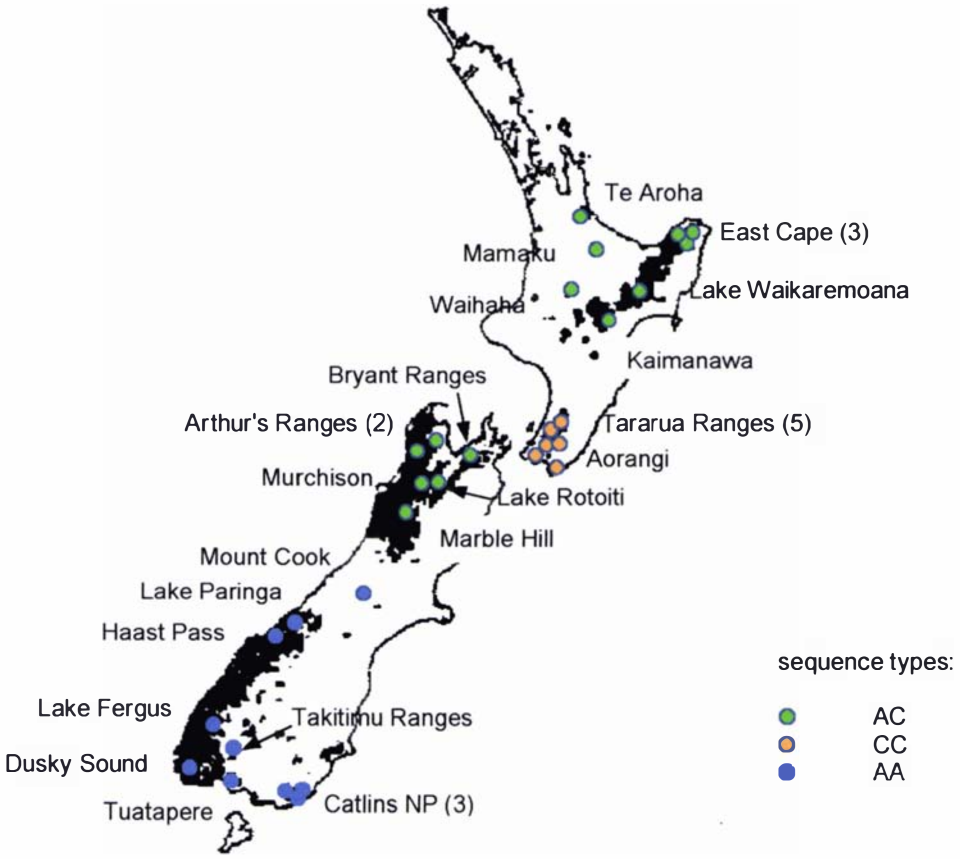


Figure 6.6: Distribution of *N. menziesii* plastid sequence types based on sequence position 183 of *trn*L intron and on position 128 of *trn*L-*trn*F intergenic spacer. Numbers in brackets indicate the number of samples displayed from one area. Samples: Te Aroha: N1-2; East Cape: N2-1, N2-8, N2-9; Urewera: N3-4; Mamaku forest: N4-2; Waihaha River: N5-5; Tararua Ranges: N8-1, N8-4, N8-6, N9-2, N9-10; Rimutaka Ranges: N10-1; Aorangi Ranges: N11-1; Abel Tasman NP: S1-1; Arthur Ranges: S1-9, S1-10; Buller River: S3-12; Mount Cook: S6-1; Haast: S7-5; Longwood Range: S14-3; Catlins NP: S15-6, S15-7, S15-11, S15-12, S16-12. Herbarium voucher numbers for all samples are listed in Appendix III.

6.2.3 Nm-af2 Allele Frequencies and Wright's Statistics

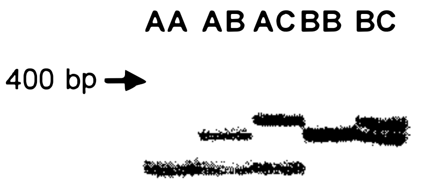


Figure 6.7: Nm-af2 allele combinations found in *N. menziesii*.

To determine allele variations and the amount of gene flow amongst populations allele distribution were investigated for the Nm-af2 nuclear DNA region (Figure 6.7) of *Nothofagus menziesii* (Section 4.2.2). The Nm-af2 DNA fragment was amplified by PCR from genomic DNA of 220 individual samples from all four provenances. The products were separated on a polyacrylamide (PAA) matrix and visualised by silver staining. Results are shown in Tables 6.4, 6.5, 6.6 and 6.7. Protocols for DNA extractions, PCR and PAA gel electrophoresis are described in Sections 2.3.1-2.3.3. Herbarium voucher numbers for all *N. menziesii* samples used in this study are listed in Appendix III alongside with a map depicting *N. menziesii* sampling areas in New Zealand.

Central North Island:

area	location	sample	genotype
N1	Te Aroha	1	BB
		2	BB
		3	BB
N2	East Cape	1	BB
		2	BB
		3	BB
		4	BB
		5	BC
		6	BC
		7	BB
		8	AC
	Hikurangi	9	AB
		10	BB
		11	BB
		12	BB
		13	BC
		14	BB
		15	BC
N3	Urewera	1	BB
		2	BC
		3	BB
		4	AA
		5	AB
		6	BC
		7	BC
		8	BB
		9	BB
		10	BB
		11	AB
N4	Mamaku Forest	2	BC
		3	BC
		4	BB
N5	Waihaha River	1	BB
		2	BB
		3	BC
		4	BC
		5	BC
		6	BC
N6	Ohakune - Turoa	1	BB
		2	BB
		3	BB
		4	BB
N7	Kaimanawa Forest East	1	AB
		2	BB
		3	BB
		4	BB
		5	BB
		6	BC
		7	BB

Total number of samples: 49

Table 6.4: Nm-af2 alleles of *N. menziesii* samples in the central North Island. Herbarium voucher numbers for all samples are listed in Appendix III.

Lower North Island:

area	location	sample	genotype
N8	Tararua Ranges West	1	AA
		2	BB
		3	BB
		4	AC
		5	BC
		6	BC
		7	AB
		8	BB
		9	BB
		10	BB
		11	BB
		12	BB
N9	Tararua Ranges East	1	BB
		2	AB
		3	BB
		4	BB
		5	AB
		6	AB
		7	BB
		8	AB
		9	BB
		10	BB
		11	BB
		12	BB
		13	BB
		14	BB
		15	BB
		16	AB
N10	Rimutaka Ranges	1	BB
		2	BB
		3	AB
		4	AB
		5	AB
		6	BB
		7	BB
N11	Aorangi Ranges	1	BB
		2	AB

Total number of samples: 35

Table 6.5: Nm-af2 alleles of *N. menziesii* samples in the lower North Island.

Northwest Nelson:

area	location	sample	genotype
S1	Abel Tasman	1	AB
		2	AB
		3	BB
		4	AB
	Takaka Hill	5	BB
		6	BB
	Arthur Range	7	AB
		8	BB
		9	BB
		10	AA
S2	Bryant Range	1	AB
		2	BB
		3	AB
		4	BB
S3	Hope Saddle	1	BB
		2	BB
		3	BB
		4	BB
	Buller River	5	BB
		6	AB
		7	BB
		8	AB
		9	BB
		10	BB
		11	BB
		12	BB
S4	Maruia River	1	BB
		2	AB
		3	AB
		4	BB
		5	AB
		6	AB
S5	Reefton	1	AB
		2	AB
		3	AB
		4	AB
		5	AB
		6	AA
	Paparoa NP	7	BB
		8	BB
	Ngahere	9	AB
		10	AB
		11	BB
	Blackball	12	BB
		13	AB
		14	AB
		15	AA

Total number of samples: 47

Table 6.6: Nm-af2 alleles of *N. menziesii* samples in the northern South Island. Herbarium voucher numbers for all samples are listed in Appendix III.

Southland:

area	location	sample	genotype
S6	Mt. Cook	1	AA
S7	Haast	1	BB
		2	AB
		3	AB
		4	BB
		5	BB
		6	AB
		7	BB
		8	AA
S8	Haast-Port Jackson	1	AB
		2	BB
		3	AB
		4	BB
		5	AA
		6	AA
		7	AA
S9	Homer Saddle	1	BB
		2	BB
		3	BB
		4	AB
		5	AB
		6	AB
		7	AB
		8	BB
S10	Te Anau	1	AA
		2	AB
		3	AA
S11	Dusky Sound	1	BB
		2	AB
		3	BB
		4	AB
		5	AB
		6	AB
		7	BB
		8	AB
		9	AB
		10	AA
		11	BB
S12	Takitumu Range West	1	AB
		2	AA
		3	AB
		4	AB
S13	South Coast	1	AB
		2	AB
		3	AB
		4	BB
		5	AB
		6	AA
		7	BB
		8	BB
S14	Longwood Range	1	AA
		2	AB
		3	BB
S15	Catlins West	1	BB
		2	AB
		3	BB
		4	AB
		5	BB
		7	AA
		8	BC
		9	AB
		10	AB
		11	AB
		12	AC
		13	AB
		14	BB
		15	AA
S16	Catlins NF East	1	AB
		2	AB
		3	AB
		4	AA
		5	BB
		6	BB
		7	AA
		8	AB
		9	BB
		10	AB
		11	AA
		12	BB
		13	AB
		14	AB
		15	AA
		16	AB
S17	Waipori	1	AB
		2	BB
		3	AB
		4	AB

Total number of samples: 84

Table 6.7: Nm-af2 alleles of *N. menziesii* samples in the southern South Island. Herbarium voucher numbers for all samples are listed in Appendix III.

Nm-af2 Allele Frequencies

To investigate the amount of genetic differentiation and measure geneflow between the four geographical regions, allele frequencies were counted and compared. The percentage of alleles A, B and C per region is visualised in Figure 6.8.

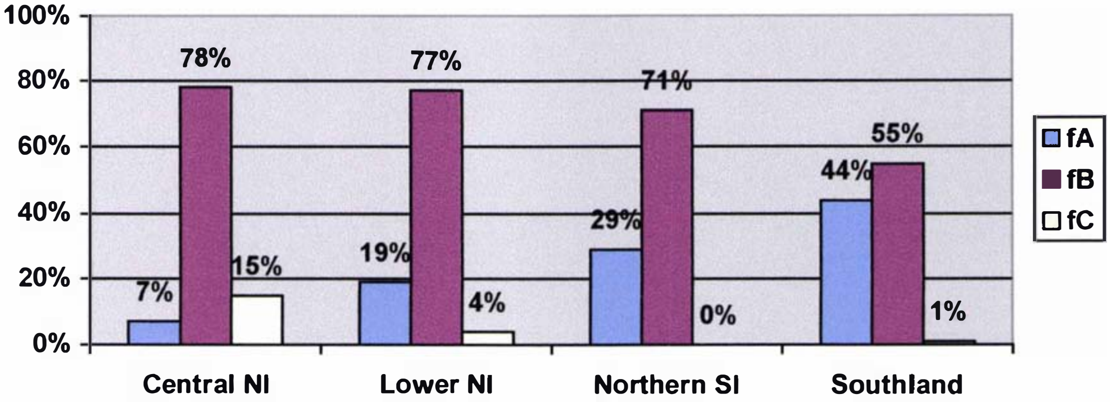


Figure 6.8. Nm-af2 allele frequencies (f) for four *Nothofagus menziesii* provenances. Abbreviations: **NI:** North Island **SI:** South Island.

Allele B is the most common in all populations. Its frequency is nearly constant between both North Island populations and the northern South Island. In the Southland population, allele B is 20% less frequent. Two allele gradients (clines) are noticeable. The frequency of allele A increases from north to south while the frequency of allele C decreases. With 15 % allele C is present in larger numbers only in the central North Island. It is not found in population samples from Nelson and occurs in less than 1 % of the Southland samples.

The clines indicate possible geneflow of nuclear DNA between geographically adjacent populations. This result is consistent with the pattern of sequence variation found in nuclear rDNA and allele B of Nm-af2 (Section 6.2.1).

Wright's Statistics

The fixation index (F_{st}) was calculated to test the assumption of random mating between individuals of all regions and measure geneflow ('migration') between the four populations. If the assumption of random mating between all individuals is fulfilled the fixation index will approximate zero. The migration rate (Nm) was derived from the F_{st} value employing the Wright equilibrium (Wright, 1951) as described in Section 2.5.2. F_{st} and Nm values for pairs of regions and the overall rates are listed in Table 6.8.

	Fixation index: F_{st}	'migration': Nm
Central NI - Lower NI	0.0168	14.64
Lower NI - Nelson	0.0099	24.96
Nelson - Southland	0.0270	8.99
Central NI - Nelson	0.0453	5.27
Lower NI - Southland	0.0622	3.77
Central NI - Southland	0.1061	2.11
overall	0.0689	3.38

Table 6.8: Fixation index (F_{st}) and migration rate (Nm) between the four *N. menziesii* populations calculated for the Nm-af2 region.

For all regions, the Wright's statistics on Nm-af2 alleles shows low F_{st} values and corresponding high migration rates (Nm). Nm values are particularly high between adjacent regions, indicating that geneflow is most frequent across Cook Strait and between the central and lower North Island populations. Geneflow decreases in correlation with the greater distances between the disjunct populations.

According to Wright's assumptions, a migration rate Nm > 1 indicates that geneflow between regions is sufficient to prevent fixation of alleles and genetic drift, respectively. Hence, per definition, the overall rate of Nm = 3.38 between all four beech provenances shows that the assumption of random mating between all individuals is fulfilled.

6.3. *NOTHOFAGUS MOOREI*

6.3.1 Sequence Analyses of Nuclear DNA

Nuclear rDNA and Nm-af2 regions were sequenced and aligned for six individuals of *Nothofagus moorei* from two disjunct populations in Australia. Samples were analysed from the northernmost population at Lamington National Park, Macpherson Ranges, Queensland, and the southernmost population at Barraga Swamp, Barrington National Park, New South Wales.

The 586 bp alignment of ITS1, ITS2 and 5.8S of nuclear rDNA contained ten sites that were variable between *N. moorei* individuals. Variability was found in the internal spacer regions only, the 5.8S sequence was identical for all samples. All variable sites showed heteroplasmic signals in more than one accession. Positions and character states of the variable sites have been listed in Table 6.9.

The 317 bp alignment of *N. moorei* Nm-af2 sequence contained one variable site at position 118 (Table 6.10) and no ambiguous/heteroplasmic patterns. The basepair difference is geographically informative and distinguishes samples from the two geographic locations.

area	sample	p83	p95	p172	p174	p366	p414	p514	p563	p566	p568
Lamington	L1	TC	TC	TC	GA	T	GA	GA	TC	TC	CA
Lamington	L2	TC	C	TC	G	TC	G	GA	TC	C	CA
Lamington	L3	C	C	TC	G	TC	G	GA	C	TC	A
Barraga Swamp	BS1	TC	TC	T	GA	T	GA	GA	TC	TC	CA
Barraga Swamp	BS2	TC	TC	T	GA	T	GA	GA	TC	TC	CA
Barraga Swamp	BS3	TC	TC	T	G	T	G	G	TC	TC	CA

Table 6.9: Variable sites from aligned rDNA sequence of *N. moorei* population samples. Herbarium voucher numbers for all samples are listed in Appendix II. Abbreviations: p = position.

area	sample	p118
Lamington	L1	A
Lamington	L2	A
Lamington	L3	A
Barraga Swamp	BS1	T
Barraga Swamp	BS2	T
Barraga Swamp	BS3	T

Table 6.10: Variable sites from aligned Nm-af2 sequence of *N. moorei*. Herbarium voucher numbers for all samples are listed in Appendix II. Abbreviations: p = position.

The phylogenetic information in both data sets has been visualised with SplitsTree 3.1 (Huson, 1998). The graphs are presented in Figures 6.9 and 6.10. Phylogeographic structure is recognisable in both nuclear data sets.

The splitsgraph shown in Figure 6.9 indicates that the patterns of variation in the nuclear rDNA region are somewhat complex, which may indicate some degree of geneflow or ancestral polymorphism between Barrington and Lamington individuals. However, polymorphic patterns provide evidence for a split that separates taxa from the northernmost and southemmost population. Because of the reticulation observed with the ITS marker it is impossible to infer the ancestral character of this locus, given the limited sampling undertaken.

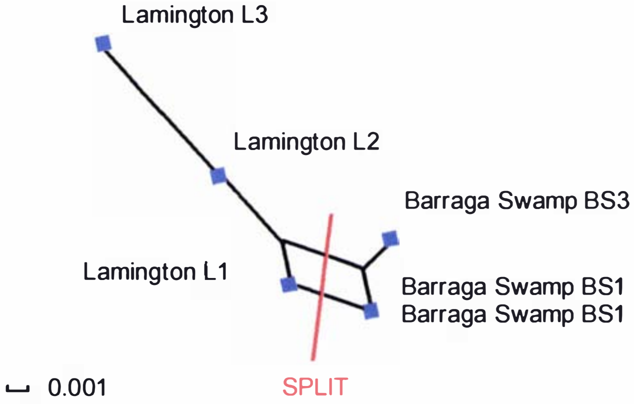


Figure 6.9: Splitsgraph based on nuclear rDNA data of *N. moorei*. Scale: 0.001 substitutions per site. Fit = 100.0, characters = 555, constant = 545, non-parsimony = 550, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Herbarium voucher numbers for all samples are listed in Appendix II.

In the Nm-af2 region of *N. moorei*, an obvious split distinguishes taxa from Barrington and Lamington National Parks. The root - determined by outgroup placement with other *Lohpozonia* and *Fuscaspora* beeches - joins at the northernmost population of Lamington National Park, indicating that the Nm-af2 allele in the southernmost population of Barrington National Park are the more derived genotype. The region differs from the ancestral type by one basepair.

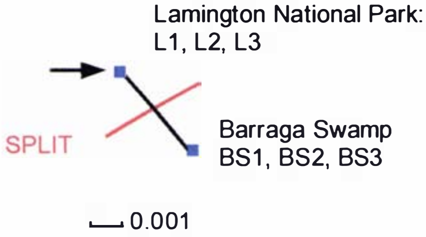


Figure 6.10: Splitsgraphs based on nuclear Nm-af2 data of *N. moorei*. The arrow indicates the position of the root. Scale: 0.001 substitutions per site. Fit = 100.0, characters = 317, constant = 316, non-parsimony = 316, -dsplits -hamming. For splitsgraph details see Section 2.4.4.4. Herbarium voucher numbers for all samples are listed in Appendix II.

6.3.2 Sequence Analyses of Plastid DNA Sequence

436 bp of the *trn*L intron and 437 bp of the *tm*L-*trn*F intergenic spacer were sequenced and analysed for the same six individuals from Barrington and Lamington National Parks. The monospecific alignments were free of ambiguous signals and identical for all individuals. In *N. moorei*, the regions were not variable enough to distinguish between individuals and populations.

6.4 DISCUSSION

Investigations in this chapter were made to study the nature of the distribution patterns of *Nothofagus menziesii* in New Zealand. DNA sequence and allele frequency data were characterised for taxa sourced from four areas. These were from: central North Island, lower North Island, northern South Island (Nelson) and southern South Island (Southland). These four provenances are separated by (i) the North Island 'beech gap', (ii) by Cook Strait and by (iii) the South Island 'beech gap'. Taxa from these four regions could be distinguished on the basis of DNA sequence of the *tm*L intron and *trn*L-*tm*F intergenic spacer. Samples from the lower North Island and Southland had independently derived genotypes while all samples investigated from central North Island and Nelson displayed the ancestral plastid sequence (Section 6.2.1.2).

A disjunct distribution of *Nothofagus moorei* in Australia was also investigated by genetic characterisation of samples from the southernmost (New South Wales) and the northernmost (Queensland) populations across the geographic range of this species. This comparison allowed some parallels to be drawn between situations in New Zealand and Australia.

6.4.1 How Much Geneflow Occurs between Disjunct Beech Populations?

6.4.1.1 *Nothofagus menziesii*

Sequence alignments of the nuclear rDNA and Nm-af2 (allele B) showed a lack of genetic distinctiveness in *N. menziesii* accessions from different provenances. Such a result would be expected when concerted evolution or gene conversion caused homogenisation between multiple copies of genes and/or alleles (Arnheim *et al.*, 1980). An alternative conclusion is the occurrence of geneflow between populations. This hypothesis was further investigated with population genetic analysis of Nm-af2 alleles of *N. menziesii* (Section 6.2.1.3).

Nm-af2 allele frequencies were distinct for *N. menziesii* populations from the four different geographic areas. Two clines, in respect of allele frequencies, were observed between northern and southern populations. Differences in allele frequencies were most apparent in non-adjacent populations, indicating that geneflow mostly occurred between neighbouring individuals or subpopulations (Briggs and Walters, 1984). Indeed, Wright's statistics on the observed allele frequencies confirmed a high degree of nuclear geneflow between adjacent populations.

Conclusions on the amount of seed dispersal were inferred from plastid DNA. Analyses of plastid coded sequences suggested genetic distinction between taxa from the four provenances. If the plastid genome is maternally inherited in silver beech, this result indicates that seed dispersal between populations plays no significant role. This result is consistent with findings from ecological studies that suggest that southern beech expands its range slowly and seeds are rarely dispersed more than a few hundred meters (Allen, 1987; Wardle, 1988a; Rogers, 1989).

As reticulation was only found in nuclear DNA regions, gene flow was interpreted to be the result of pollen mediated dispersal. In contrast to the poor seed dispersal abilities, ecological studies have shown that *Nothofagus* pollen can travel up to 500 km (Wardle, 1988b). The DNA analyses reported here confirm that pollen dispersal is effective in mediating geneflow between disjunct populations of *N. menziesii.*

6.4.1.2 *Nothofagus moorei*

DNA sequence analysis for *N. moorei* showed genetic differentiation between geographically disjunct populations of New South Wales (Barraga Swamp) and southern Queensland (Lamington National Park). A clear split was observed for Nm-af2 sequence, while nuclear ITS provided weaker evidence for genetic isolation between individuals of the two disjunct populations (Section 6.3.1.1). These latter data could be consistent with either a limited amount of gene flow or retention of ancestral polymorphisms in both the Barrington and Lamington populations. The first possibility does not seem likely considering the fact that both regions are separated by over

500 km. Although *Nothofagus* pollen has been shown to travel this distance without loosing fertility (Wardle, 1967), analyses do not suggest regular exchange of nuclear DNA between the two distant populations. In the case of *N. menziesii*, as found here, pollen mediated dispersal was relatively low between populations separated by distances greater than 200 to 300 km.

Studies on *N. menziesii* plastid geneflow presented in Section 6.2.2 and earlier ecological studies by Allen (1987) Wardle (1988a) and Rogers (1989) substantiate limited seed dispersal abilities of *Nothofagus*. These results as well as the low nuclear geneflow which was observed between geographically distant populations of *N. moorei* (Section 6.3.1), indicate that seed dispersal between the distant populations of *N. moorei* can also be largely ruled out.

6.4.2 Inferring the Age of Beech Population Disjunctions

In previous work, analyses of plastid DNAs have been useful for investigating Late Tertiary and Quaternary plant histories (e.g. Lockhart et al. 2001). However, low genetic variation has been problematic for inferring the nature of late Quaternary events (Proven *et al.* 1999). Results from the present study are consistent with these findings by revealing a low amount of genetic differentiation between individuals of disjunct populations. This low level of genetic differentiation makes estimation of divergence times difficult as the variance on such estimates is expected to be high (Bertorelle and Rannala, 1998). Further, although in the present study a higher level of variation was observed with nuclear markers (nuclear rDNA regions and Nm-af2), the extent of pollen mediated gene flow unfortunately precludes the use of these data for estimating population divergence times in *N. menziesii.*

6.4.2.1 *Nothofagus menziesii*

From fossil pollen data McGlone (1980) concluded that during the last glacial maximum (Otiran glaciation) between 25 000 and 14 000 BP forests in general had an extremely restricted distribution in New Zealand. However, nearly all pollen spectra that date to this period contain very small amounts of podocarp and beech pollen, which suggests the restricted presence of forest in most areas. Palynological evidence suggests a very local presence of silver beech in the Waikato Basin, in the Wellington area and in north Westland (McGlone, 1980 and references therein). This evidence is consistent with a disjunct distribution pattern of silver beech occurring during and since the Otiran glaciation.

Wardle (1967) and McGlone *et al.* (1996) observed that *Nothofagus menziesii* pollen has only become dominant in New Zealand in the more recent geological strata of the Pleistocene and Holocene. The greater abundance of this species has been attributed to prevailing cool and moist climates (Wardle, 1967). *Lophozonia* pollen, found in

Oligocene sediments in central Otago, has been attributed to a different extinct species with preference to warmer habitats (Mildenhall and Pocknall, 1989). Major range expansion of the present species *N. menziesii* seems to have taken place since the beginning of the Pleistocene.

Sequence analyses show a similar amount of genetic differentiation between individual populations of New Zealand silver beech as between populations of *N. moorei* in Australia. Interestingly, two out of four *N. menziesii* populations have fixed a characteristic base exchange in the plastid coded *trn*L regions (Section 6.2.1.2). Phylogenetic analyses suggest that these base exchanges are independently derived from an ancestral sequence type, which is still present in the two remaining populations.

Under a simple stochastic model of evolution and assuming a substitution rate of 0.24×10^{-9} substitutions per site and year, which was determined for the *Nothofagus* *trn*L regions (Section 3.6.2.5), one change occurs in 5 million years. With such an expectation, sequence data do not clearly reject the hypothesis that beech populations became isolated during Pliocene mountain building. However, it would seem unlikely that the present population disjunctions of silver beech predate the Late Miocene as suggested by McGlone (1985).

The time needed to fix a mutation within a population depends on the population size: in a large population fixation takes much longer than in a smaller population (Page and Holmes, 1998). Rapid fixation of the observed mutations would have been possible when beech forest was heavily reduced during the Otiran glaciation and forests were refined to small glacial refugia (McGlone, 1980; Wardle, 1988a). This hypothesis would allow a maximum of one refugium in areas with a derived sequence type such as the lower North Island and Southland. However, it has been shown that beech forests expand rather slowly especially in competitive situations (Wardle, 1963) and a fast regeneration of forests at the end of the last glacial maximum as suggested by pollen analyses (McGlone, 1980) may not have been possible from a single refugium (Wardle, 1963; 1988a).

The palynological evidence and the relatively slow rate of sequence evolution in *Nothofagus* *trn*L regions suggests that the disjunctions between silver beech populations occurred before the beginning of the most recent glacial maximum, 25 000 years ago. Disjunctions may have co-occurred with climate changes and range expansions of silver beech at the beginning of the Pleistocene.

6.4.2.2 *Nothofagus moorei*

The Australian fossil pollen record suggests that *Nothofagus* forests with *Lophozonia, Fuscaspora* and *Brassospora* beeches were much more abundant during the Early and Late Tertiary than they are today. Macphail (1997a) demonstrated an extensive regional loss of beech forest from lower altitudes of south-eastern Australia, New South Wales and southern Queensland at the beginning of Pliocene (ca. 5 million years ago). *Nothofagus* forests had disappeared from many lowland areas in New South Wales 2.5 million years ago, presumably as a result of reduced rainfall. *Nothofagus* forests have persisted only in higher altitudes in the Border Ranges of New South Wales and Southern Queensland (see Section 6.1.2.1). As pollen evidence does not support the expansion of *Nothofagus* forests during Pleistocene glacial and interglacial stages, Macphail (1997a) suggested that populations were disjunct since before the onset of the Pleistocene.

Macphail concluded that during the Late Pliocene climates became more similar to present conditions and that distributions of *Nothofagus* are controlled by the mean climatic values. The observation that *Nothofagus* rainforests are characteristically found in cool to cold climates with relatively high, evenly distributed rainfall in Australia corroborates this hypothesis.

DNA analyses of nuclear and plastid markers on *N. moorei* population samples show that individuals of disjunct populations are not highly differentiated genetically. One distinct mutation was observed between *N. moorei* populations in the nuclear Nm-af2 region. If it is assumed that a time span of 2.5 million years has not been sufficient to establish a larger amount of genetic diversity, the lack of differentiation of the plastid and nuclear DNAs indicates a slow rate of sequence evolution in *N. moorei*.

6.4.2.3 Obtaining Further Information about the Nature of the Disjunctions

Faster evolving plastid loci would be required to obtain better estimates of population divergence for *N. menziesii and N. moorei.* However, whether such loci exist is unclear. In this respect, it could be worth pursuing plant mitochondrial DNA markers. The properties of the mitochondrial genome are still unknown in beech. If they are maternally inherited and with sufficient variation, this genome might also provide a source of genetic variation for studying the phylogeographic structure of populations, as has been the case for *Hevea brasilensis* (Lou and Boutry, 1995). Plastid and mitochondrial simple sequence repeats and/or microsatellites might also provide useful data for investigating population structure as observed elsewhere (Golenberg *et al.*, 1993; Proven *et al.*, 1999; Weising and Gardner, 1999; Gugerli *et al.*, 2001).

CHAPTER 7

General Discussion

7.1 ORIGINS OF THE NEW ZEALAND FOREST FLORA

The origins of the New Zealand flora have been subject to considerable debate. Hypotheses have proposed continental shift, migration over land bridges and/or long distance dispersal across oceans to explain the observed floristic diversity (e.g. Hooker, 1853; Darlington, 1965; Raven and Axelrod, 1972; Linder and Crisp, 1995).

When New Zealand was part of the Gondwanan supercontinent, plants were presumably able to populate different regions of this large landmass without having to cross ocean gaps. Subsequently, with the break-up of Gondwana, plant genera are expected to have rafted with different continental land fragments. Biogeographic patterns that exist because populations and species have been isolated by continental drift and genetic differentiation that followed the break-up of the landmass are called patterns of vicariance. By definition, vicariant plant lineages must have had a continuous presence in New Zealand since its separation from Gondwana ca. 70 million years ago. All taxa that arrived on the New Zealand landmass thereafter must have done so by long distance dispersal.

Palynological and paleobotanical studies do not provide evidence for a continuous presence of extant genera since their first appearance in New Zealand fossil records. Extinction and re-dispersal of plant lineages cannot be ruled out (Fleming, 1979). Pole (1994) who noticed the poor match between extant plant lineages and Palaeocene fossil species developed the most extreme form of this argument. He proposed that Gondwanan plant cover may not have survived the "Oligocene drowning" of New Zealand's landmasses, and that plants and animals must have re-invaded New Zealand since that time. In support of this hypothesis, Macphail (1997b) found evidence that Gondwanan lineages have become extinct and were re-established on southern hemisphere lands during the Cenozoic. He also pointed out that, before and after the separation of New Zealand from Gondwana, many plant lineages recorded in the Australian fossil record appear much later in New Zealand geological strata.

Some biologists have tried to explain floristic relationships between New Zealand and other landmasses as being the result almost entirely of continental drift and/or by the former existence of land bridges between continents (Raven and Axelrod, 1972; Van Steenis, 1962). However, palaeobotanical and molecular studies are consistent in suggesting that floristic elements that arrived on the New Zealand landmass, while it was still part of Gondwana are rather rare.

In recent years, an increasing number of studies have been undertaken to investigate the origins and evolution of the New Zealand flora using a molecular phylogenetic approach (e.g. Mitchell and Wagstaff, 1997; Hurr, 1999; Wagstaff and Garnock-Jones, 1998; Wagstaff and Dawson, 2000). Observations from these studies have emphasised the importance of long distance dispersal in explaining the evolution of the New Zealand flora (Winkworth *et al.*, 1999, 2001). Molecular studies presented in this thesis corroborate this finding, but also provide evidence for the vicariant origin of some New Zealand plant lineages.

These results, on a number of woody plant groups native to the New Zealand forest flora, indicate that the histories and biogeographic relationships of the New Zealand flora are diverse. As forests are among the most ancient ecological environments in New Zealand (Stevens *et al.*, 1995) it has been expected that the biotic communities accommodate species that arrived on the New Zealand landmass in different geological eras. Results from molecular analyses shown here are consistent with palynological and geological evidence in suggesting that the origins of the New Zealand biota cannot be explained by one single hypothesis alone (e.g. Fleming, 1979; Winkworth *et al.*, 1999).

7.1.1 Gondwanan Origin of New Zealand Forest Plants

Once the phenomenon of plate tectonics became better understood and accepted, this provided an impetus for scientists to interpret most of the biogeographic relationships between southern hemisphere landmasses in terms of continental drift (e.g. Raven and Axelrod, 1972; Linder and Crisp, 1995).

Consistent with this interpretation, palynological and palaeobotanical research have identified a number of plant groups that were present on the New Zealand landmass before the break-up of Gondwana. Cretaceous forest floras from different locations were described by Ettinghausen (1887), Fleming (1979), Daniel *et al.* (1988), Daniel (1989), Pole (1995) Pole and Douglas (1999). Although species in the Cretaceous forests differed from extant taxa, a number of plant groups identified by these authors are also represented in the modern New Zealand flora. Among Cretaceous trees and shrubs are conifer genera like *Agathis* (Araucariaceae), *Podocarpus* and possibly *Dacrydium* (both Podocarpaceae) and angiosperms like *Nothofagus* (Nothofagaceae) *Ascarina* (Chloranthaceae), and Winteraceae (Fleming, 1979; Mildenhall, 1980; Macphail, 1997b). Most other extant plant genera first appear in the fossil record during the Tertiary and Quaternary after New Zealand had separated from Gondwana ca. 65-75 million years ago.

Phlyogenetic analyses on the New Zealand forest plants presented in this thesis support a Gondwanan origin for *Agathis australis*, *Dacrydium cupressinum*, and

Pseudowintera spp. Based on the results from molecular data, a continuous presence since the Cretaceous may also be the case for *Prumnopitys taxifolia*.

Agathis australis: Tree building analyses suggest that *Agathis australis* (New Zealand kauri) is the sole member of an ancient lineage, which has diverged early during the evolution of the genus (Section 3.2). All other extant *Agathis* species (including all three Australian taxa and species from the Malaysian region) are more closely related to each other than to *A. australis*. This result supports findings based on morphological analyses of fossil leaves that suggest that the extinct species *A. seymouricum*, which was excavated from Cretaceous sediments in New Zealand, is more closely related to *A. australis* than it is to other extant species of Araucariaceae (Daniel, 1989). Palaeobotanical and molecular results presented in this thesis are most parsimoniously explained by the hypothesis that *A. australis* evolved on the New Zealand landmass as a direct descendant of *A. seymouricum*.

Podocarpaceae: Mildenhall (1980) and Macphail (1997b) found fossil evidence that the genus *Dacrydium* had been part of a Cretaceous forest flora. In a maximum likelihood phylogeny *Dacrydium cupressinum* (rimu), the sole member of a New Zealand endemic genus, appears basal within section *Podocarpus* (see Section 3.3). Split decomposition analysis also supports this observation and suggests a unique placement of *D. cupressinum* within section *Podocarpus*. This result is consistent with an ancient and continuous presence of this genus in New Zealand. Although trans-oceanic re-dispersal (e.g. from Australia) and extinction of this genus overseas cannot be completely ruled out, rimu remains a clear candidate for being a true Gondwanan member of the New Zealand forest. Unfortunately, the small number of species presently available for studying interspecific relationships between other podocarps limits the conclusions that can be drawn about the origins of other New Zealand taxa. To better understand the evolution of the genera *Phyllocladus* and *Prumnopitys,* more overseas species need to be included in the molecular data set. It is possible that further investigations may identify a Gondwanan origin of New Zealand species of *Phyllocladus* and *Prumnopitys*.

Winteraceae: Molecular analyses on Winteraceae based on nuclear ribosomal DNA (rDNA) sequence presented in Section 3.8 of this thesis reveal great genetic differentiation between lineages from different continents. There is no evidence in the sequence data to suggest events of recent dispersal between the formerly Gondwanan landmasses Australia, New Zealand, New Caledonia and South America. Results of molecular analyses of nuclear rDNA support the vicariant evolution of *Pseudowintera* on New Zealand landmasses from Gondwanan ancestors.

Interestingly these examples, for which there is evidence of direct Gondwanan ancestry, represent isolated genera in New Zealand containing only one (*Agathis* and *Dacrydium*) or two (*Pseudowintera*) different species each. Although *Pseudowintera* species have diverged from a common New Zealand ancestor relatively recently (perhaps during the Pliocene), no intensive radiation events are observed within these lineages. Compared to recent (Pliocene-Pleistocene) adventives like *Gentiana* with 24 extant species, *Hebe* with 79 extant species, *Myosotis* with 35 extant species or *Ranunculus* with 43 extant species the 'Gondwanan relicts' are conspicuous in having a very small number of species.

Although studies have suspected a vicariant relationship between New Zealand and overseas plant lineages (Suh *et al.*, 1993; Manos, 1997), the molecular analysis of *Agathis* in particular, which is presented in this thesis, may be the first to provide clear phylogenetic evidence for the continued presence of a New Zealand plant lineage since the Cretaceous. Evidence from these studies rejects the hypothesis by Pole (1994) and Macphail (1997b) that the New Zealand flora was entirely eliminated by marine transgression of landmass during the Oligocene. Further molecular studies may identify additional examples of a Gondwanan origin within the New Zealand biota. Candidates for further study may be found particularly among fungus, moss, and fern species as well as among insects and invertebrates.

7.1.2 Dispersal Origin of New Zealand Forest Plants

Studies based on fossil and molecular data suggest that a large proportion of plants have reached the New Zealand archipelago via long distance dispersal (Mildenhall, 1980; Pole, 1994; Winkworth, 2000). The analyses presented in this thesis substantiate dispersal origins for a number of taxa native to the New Zealand forest flora.

Podocarpaceae: The maximum likelihood phylogeny of Podocarpaceae presented in Section 3.3 suggests a close relationship between the two New Zealand taxa *Dacrycarpus dacrydioides* (kahikatea) and *Podocarpus totara* (totara). Both species are shown as a monophyletic group, which also includes the Chinese species *Dacrycarpus imbricata*. The phylogeny suggests that the three species have diverged from a common ancestor relatively recently. The hypothetical ancestor may have arrived in New Zealand after the break-up of Gondwana. The close relationship between the New Zealand and the Chinese species is most parsimoniously explained by trans-oceanic dispersal of this genus.

Metrosideros: Wright *et al.* (2000) proposed dispersal of *Metrosideros* from New Zealand during, or subsequent to, the Pleistocene by means of westerly air currents. The genus has first been recognised in the New Zealand fossil pollen record during the Palaeocene, ca. 60 million years ago. On the basis of fossil evidence a possibly Gondwanan origin of *Metrosideros* cannot be ruled out. However, results from molecular analyses presented in Section 3.4 of this thesis suggest a post-Palaeocene divergence of all extant *Metrosideros* species from New Zealand and overseas. The results are consistent with an Early Tertiary dispersal origin of *Metrosideros* to New Zealand or a dispersal event from New Zealand to Africa.

Myrsinaceae: Palynological evidence suggests that *Myrsine* has been present in New Zealand since the Oligocene, ca. 35 - 30 million years ago, suggesting an origin by dispersal for this genus in New Zealand. Unfortunately, other fossil evidence is rare and little is known about the evolutionary history of this plant group. DNA analyses on New Zealand Myrsinaceae support fossil evidence for an Oligocene divergence between extant Myrsinaceae and indicate that *Myrsine* has had a continuous presence in New Zealand since the Oligocene.

Nothofagus: Mildenhall (1980) reported the first appearance of *Nothofagus* pollen in Cretaceous (subgenus *Lophozonia*) respectively in Eocene sediments (subgenus *Fuscaspora*) in New Zealand. McGlone *et al.* (1996) and Macphail (1997b), however, raised doubts as to whether these reports were valid. These authors proposed a Palaeocene arrival of *Fuscaspora* beeches in New Zealand and an Eocene arrival of *Lophozonia* beeches.

In this thesis, DNA analyses were performed on all members of the two *Nothofagus* subgenera *Lophozonia* and *Fuscaspora*. These molecular studies were based on two plastid genome regions (*trn*L intron and *trn*L-*trn*F intergenic spacer) and two nuclear genome regions (rDNA and Nm-af2 (Section 4.2.2)). Results from all DNA regions were consistent in revealing relatively low genetic differentiation between species of different continents. This result corroborates findings from other molecular phylogenetic studies on *Nothofagus* (Martin and Dowd, 1993; Manos, 1997; Setoguchi *et al.*, 1997). Genetic differentiation between taxa was greater within subgenus *Lophozonia* than within subgenus *Fuscaspora*. This result indicates that extant *Fuscaspora* beeches have diverged more recently than *Lophozonia* beeches.

Estimates for divergence times presented in Section 3.6.2.6 were calibrated assuming a Cretaceous divergence between *Nothofagus* subgenera in New Zealand and Australian. Such a calibration clearly leads to rejection of a hypothesis of vicariant evolution for *Nothofagus* species as suggested e.g. by Darlington (1965) or Linder and Crisp (1995). Results from both nuclear and plastid DNA regions suggested divergence of *Nothofagus* species during the Mid Tertiary for *Lophozonia* and during the Late Tertiary for *Fuscaspora* beeches. This result is consistent with hypotheses by Fleming (1979), Hill (1992) and Hill and Dettman (1996) who suggested a dispersal origin of extant *Nothofagus* species. The result implies that *Nothofagus* has no continuous presence in New Zealand since its first appearance in the pollen record. The distribution of extant *Lophozonia* and *Fuscaspora* beeches is most parsimoniously explained by recurrent trans-oceanic dispersal.

Proteaceae: Both extant New Zealand species of Proteaceae *Knightia excelsa* and *Toronia toru* appear in the fossil pollen record after New Zealand has separated from the super continent Gondwana. *Knightia* was first recognised in Eocene sediments, pollen of *Toronia* first appear during the lower Quaternary (Mildenhall, 1980; Macphail, 1997b).

Until recently, the endemic genus *Toronia* has been included in *Persoonia* Smith, (see Allan, 1982) which indicates the close morphological alliance of the New Zealand endemic with the speciose Australian genus *Persoonia*. Molecular analyses on Proteaceae based on *atp*B gene sequence (Section 3.7) corroborate the inference from morphological data and suggest a close phylogenetic relationship between the Australian genus *Persoonia* and the New Zealand endemic. The ML phylogeny reveals a close relationship between the *Toronia*/*Persoonia* lineage and the Australian genus *Placospermum*. This result suggests an Australian origin of this lineage and a recent dispersal event of a common ancestor of *Persoonia* sp. and *Toronia toru* from Australia to New Zealand.

In contrast to the origin by recent dispersal of *Toronia,* molecular analyses suggest that *Knightia excelsa* is an early-diverged proteaceous lineage. This result is corroborated by fossil data, which indicates that an Early Tertiary dispersal event established *Knightia* in New Zealand. Phylogenetic analysis provides evidence for a continuous presence of this lineage since its first appearance in the fossil pollen record during the Eocene. However, the *atp*B gene sequence gives poor phylogenetic resolution between all deep branching taxa in the *atp*B phylogeny, and this limits inferences that can be made concerning the basal lineages. Thus, such conclusions remain tentative, as do the intergeneric relationships between *Knightia* and other Proteaceae. The lack of resolution at the deeper phylogenetic levels results from contradictory patterns of relationship, most probably due to saturation of the *atp*B gene sequence i.e. when the actual number of base exchanges is much higher than the observed number of changes (Page and Holmes, 1998). To improve phylogenetic resolution between proteaceous genera more conserved DNA regions need to be investigated (e.g. perhaps the *rbc*L gene of the plastid genome or nuclear 18S rDNA region).

7.1.2.1 Mechanisms of Long-Distance Dispersal

Research has provided little evidence about how adventive plants were transported to New Zealand. In most cases it can only be speculated upon. As the existence of land bridges that connected New Zealand to other landmasses is based on little geological evidence, it has to be assumed that plants, which arrived after the break-up of the ancient continent, must have done so by trans-oceanic dispersal. To reach the New Zealand landmass across water, seeds must have been transported in the air or on the water.

Prevailing westerly winds have been regarded an important mechanism to mediate dispersal by air floatation of seeds across southern oceans (e.g. Raven, 1973). A number of vascular plants, ferns and moss species have airborne seeds that may be able to travel large distances. However, generally, floatation of seeds or small fruits in air currents accounts for a very small proportion of floras on islands (Carlquist, 1996). Hence, the proportion of woody plant lineages that have arrived in New Zealand via air floatation may be even smaller as seeds of trees and shrubs are generally larger than seeds of herbaceous plants.

External or internal transport of seeds by birds may be another efficient way of long distance dispersal of plant species across oceans (Carlquist, 1996). Barbed or bristly fruits may cling to feathers; other seeds may also be transported internally in intestines or crops or may be included in soil particles that are attached to the migrating bird. Carlquist suggested that bird dispersal is more efficient than other means because migratory birds tend to visit similar ecological environments, and seeds are more likely to be deposited in suitable environments. Internal transport is very likely to be the dispersal mechanism of *Myrsine* spp. and *Toronia*, as birds may eat the drupes and seeds are usually not digested. *Nothofagus* seeds as well as podocarp seeds, which have been found unlikely to be eaten by birds (Preest, 1963), are small enough to be included in soil particles which are attached to birds and transported across distances. *Knightia* may have arrived similarly on the New Zealand landmass. Carlquist (1996) thought that external transport by birds might also account for the dispersal of *Pittosporum* spp.

7.2 OVERSEAS RELATIONSHIPS OF THE NEW ZEALAND FLORA

Earlier studies have classified elements of the New Zealand biota according to their geographic relationships and possible origins. Conclusions concerning the origins of ancestral taxa have often been made from morphological and/or fossil evidence (e.g. Hill, 1987; Carlquist, 1996; Macphail, 1997b). One such classification was proposed by Fleming (1979) and is described in Section 1.5.1.

During the past decade, analyses based on molecular data have been developed as an alternative method for investigating phylogenetic relationships between species and genera. This approach has provided a useful means for reconstructing the recent and ancient biogeographic history of species (e.g. Cooper and Cooper, 1995; Ferris *et al.*, 1995). Based on molecular sequence data most progress has been made in elucidating the origins and evolution of the New Zealand alpine flora (e.g. Wagstaff and Garnock-Jones, 1998; Winkworth, 2000; Lockhart *et al.*, 2001).

7.2.1 The Trans-Tasman Connection

Comparisons between the fossil and extant biota of Australia and New Zealand reflect the close geographic link between the two landmasses (Fleming, 1979; Macphail, 1997b; McGlone, 2001). Although separated by over 1200 km many New Zealand taxa are also found in Australia. Hence, for a long time, Australia has been thought to be the major place of origin for New Zealand plant lineages (Fleming, 1979). However, it has recently been shown that plants have also dispersed from New Zealand to Australia (Winkworth *et al.*, 2000a; Lockhart *et al.*, 2001).

Investigations on elements of the New Zealand forest flora presented in this thesis also reflect the close floristic relationship between Australian and New Zealand taxa. The connection is particularly evident for *Toronia toru* (Proteaceae) for which analyses suggested a recent dispersal origin from Australia. A trans-Tasman relationship of dispersal was also recovered for *Nothofagus* species of both extant New Zealand subgenera *Lophozonia* and *Fuscaspora* as well as for species of Myrsinaceae (*Myrsine* and *Rapanea*). For both families, an Australian origin has not been established, however it remains likely.

7.2.3 The Pacific Connection

The South Pacific region has also been identified as a potential source of origin for New Zealand plant lineages (Fleming, 1979). Molecular analyses on members of the woody forest flora have substantiated these observations and identified a recent dispersal origin for *Sophora* spp. section *Sophora* in the north-west Pacific (Hurr *et al.*, 1999) and for *Corynocarpus* in New Guinea (Wagstaff and Dawson, 2000). The genus *Metrosideros* has also been considered as being of South Pacific origin (Fleming, 1979). This hypothesis has not yet been confirmed by either morphological, palaeobotanical or molecular analyses. Studies presented in this thesis rather provide most support for an African origin of *Metrosideros* in New Zealand (Section 3.4). Wright *et al.* (2000) analysed interspecific relationships within subgenus *Metrosideros* and successfully demonstrated that, rather than being dispersed to New Zealand from the South Pacific, *Metrosideros* has dispersed in the other direction. These studies showed that New Zealand has also been a centre of dispersal to other lands. To clarify the origins of *Metrosideros* in New Zealand it may be helpful to conduct further molecular studies, which include other myrtaceous genera from neighbouring landmasses in particular Australia and New Caledonia. Molecular analyses may also be helpful in elucidating the centre of dispersal for other plant groups of possibly South Pacific origin, e.g. *Beilschmiedia* (Lauraceae), *Elytrante* (Loranthaceae), and *Macropiper* (Piperaceae) or *Cordyline* (Liliaceae).

7.2.4 Have Plants Dispersed From New Zealand?

In contrast to earlier suggestions that New Zealand represents an isolated Gondwanan refuge, fossil evidence and recent genetic studies indicate that New Zealand has both received and dispersed many immigrants since the break-up of the southern supercontinent (Winkworth *et al.*, 2001). Raven (1973) considered Pleistocene westerly wind currents as an important dispersal mechanism. Indeed, phylogenetic analyses suggest that during the Late Tertiary and Quaternary eastward long distance dispersal has occurred for *Metrosideros* (Wright *et al.* 2000), Australasian apioid Umbelliferae, *Hebe, Chionohebe*, *Parahebe*, *Ranunculus* and *Gentiana* (Winkworth *et al.*, 2000b). However, Winkworth *et al.* presented compelling evidence that plants were dispersed from New Zealand in all directions and against prevailing westerly winds. Molecular analyses presented in this thesis are consistent with this observation and suggest dispersal events from New Zealand to other landmasses.

Myrsine is a particularly interesting example in this context as molecular data suggest recent dispersal events from New Zealand to several oceanic islands. The Chatham Islands to the east of the New Zealand mainland host two species of *Myrsine* which are closely related to the mainland species *M. australis* and *M. divaricata.* These phylogenetic relationships indicate recent dispersal from New Zealand to the Chatham Islands.

M. divaricata occurs on Auckland Island to the south of the New Zealand mainland, and its genotype suggests derivation from a mainland source. Higham (1991) proposed that the vegetation on Auckland and Campbell Islands was completely wiped out by Pleistocene glacial stages. Hence, the occurrence of *M. divaricata* on Auckland Island is most consistently explained by the dispersal of this species during the Holocene after the end of the Otiran glaciation.

Furthermore, results from phylogenetic analyses suggest two recent dispersal events from New Zealand to the North. Molecular data suggest that *Myrsine* on both the Kermadec Islands (*M. kermadecensis*) and in Japan (*M. maximowizcii*) are derived from New Zealand ancestors.

It will be interesting in further studies to include more species from other geographic regions to investigate worldwide dispersal and speciation patterns within *Myrsine.* Based on the findings in the present work, it may be expected that long distance dispersal events in all directions are frequent. More extensive studies are needed to identify the origins of this genus and resolve intergeneric relations within Myrsinaceae.

7.3 WHAT EVENTS MAY HAVE HAD AN INFLUENCE ON THE EVOLUTION AND BIOGEOGRAPHY OF THE NEW ZEALAND FOREST FLORA?

Fossil research suggests that the composition of plant species and genera in all ecological niches of New Zealand has been changeable (Fleming, 1979; Mildenhall, 1980). These authors find evidence that throughout the geological history new taxa appear in the fossil record, while other taxa disappear. Some changes in the composition of taxa may have been induced by alteration of geological and climatic conditions. Reconstructing the evolutionary history of plant species from fossil, morphological and/or molecular data is challenging and often a definite answer cannot be found (e.g. Hill, 1992; Pole, 1994). However, in accumulating more data, some hypotheses can be excluded while others are better supported.

7.3.1 The Cretaceous

The Jurassic and Early Cretaceous has been identified as the era of fern and moss species and gymnosperms. On all ancient landmasses the woody forest flora consisted of Podocarps and Araucariaceae (Sitte *et al.*, 1991). Although podocarp forests are widespread in New Zealand, the number of extant species is low compared to the number of angiosperm species (Salmon, 1996). Molecular analyses of extant podocarp species suggested an early divergence between some podocarp lineages. Molecular analyses suggest that the radiation between all genera of section *Phyllocladus* (*Phyllocladus*, *Prumnopitys*, *Parasitaxus* and *Lagarostrobos*) dates back to the Cretaceous. The great genetic divergence between some genera of section *Podocarpus* (e.g. *Dacrydium* and *Microstrobos*) also indicates an early diversification event within this plant group. However, analyses of molecular data also suggest recent radiation events within the Podocarpaceae (e.g. *Nageia* and Australian *Podocarpus* spp.; New Zealand *Dacrycarpus Podocarpus* and *Phyllocladus* spp.). However, more detailed studies are needed to more precisely date times of radiation within podocarp plant genera.

Molecular and fossil evidence identify the Cretaceous as a time of radiation for the three Araucarian genera *Araucaria*, *Agathis* and *Wollemia.* The genus *Araucaria* also shows significant genetic differentiation between individual species that may date back to the Late Cretaceous. Within the genus *Agathis*, an ancient Cretaceous divergence was inferred for the New Zealand lineage. In contrast, genetic differentiation between all other extant species are consistent with the Pliocene/Pleistocene being an important time for radiation within this particular plant group.

Analyses of angiosperm lineages provide less evidence for speciation events during the Cretaceous. Winteraceae are the only plant group to display infrageneric relationships dating back to the Cretaceous. It is possible that angiosperms are less conservative in that their morphological characters may evolve more rapidly. Hence, angiosperm lineages that are of Gondwanan origin may be more difficult to identify than gymnosperm lineages.

7.3.2 The 'Oligocene Drowning'

Based on the distribution of sedimentary rock facies, Suggate *et al*. (1978) documented the changing proportions of land and sea during the Cenozoic and found that the New Zealand landmass was heavily reduced during the Oligocene ('Oligocene drowning'). At times of peak flooding, New Zealand was reduced to a string of low-lying islands with a total landmass of ca. 50 000 km^2, which is equivalent to 18% of the present land area. Because of this geological evidence, the Oligocene has been regarded as an era of major environmental crisis for the New Zealand biota (Fleming, 1979; Pole, 1994; Cooper and Cooper, 1995). Based on evidence from fossil plant material, Pole proposed that the entire New Zealand flora was wiped out by the 'Oligocene drowning'. In contrast, Cooper and Cooper presented analyses on ratite populations and noticed a past reduction in genetic diversity of mitochondrial DNA. The authors interpreted the result as a reduction in population size (bottleneck) which may have been caused by the severe contraction of landmass 30 million years ago. With the slow land uplift at the end of the Oligocene, 25 million years ago, a greater variety of ecological niches was available for the biota. This situation may have opened the opportunity for a subsequent radiation of species into new habitats. Radiation may have occurred in resident species and adventive plants. Mildenhall (1980) listed a number of plant groups that first appear in the New Zealand fossil pollen record during the Oligocene. These plant species may have filled in gaps that had been caused by the reduction of species diversity during the 'drowning' of New Zealand landmass. Presumable newcomers like *Myrsine*, *Nothofagus* and possibly *Metrosideros* may have benefited from this situation. Sequence diversity within some taxa of *Myrsine* and *Metrosideros* is consistent with an Oligocene radiation.

7.3.3 The Pliocene Mountain Building

In the geological history of New Zealand, the Pliocene was the era of mountain building and the creation of alpine habitats by the uplift of the Southern Alps (Batt *et al*., 2000). The diversity of the New Zealand alpine flora has raised questions about the possible age of these lineages in New Zealand. Competing hypotheses have been extreme in suggesting a recent or ancient presence for groups now well represented in the flora. Cockayne (1928), Fleming (1979) and Wardle (1963, 1967, 1978) proposed the survival of a cool-adapted element through the warm Tertiary period either in New Zealand, where they inhabited open habitats (Cockayne, 1928; Wardle, 1978), or on cool temperate lands to the south (Wardle, 1963; Fleming, 1979). Several botanical researchers have speculated that some genera, well represented in the New Zealand alpine zone, speciated to occupy alpine and sub-alpine habitats newly formed during the Pliocene and Pleistocene (Wardle, 1968; Raven, 1973). Recent patterns of speciation within alpine plant species have been observed in *Hebe* (Wagstaff and Garnock-Jones, 1998), the Gnaphalieae (Breitwieser *et al.*, 1999), *Myosotis* and *Aciphylla* (Winkworth, 2000) and *Ranunculus* (Lockhart *et al.*, 2001). Molecular analyses on the New Zealand alpine scree weta also suggest genetic diversification between disjunct populations which may correlate with the events of mountain building (Trewick *et al.*, 2000). Radiation within New Zealand forest taxa has not yet been similarly explained. However, patterns observed in molecular data of New Zealand species of *Nothofagus* may corroborate diversification and radiation into habitats that became available with the uplift of the Southern Alps as all *Nothofagus* species populate higher mountainous regions and form the tree line in the Southern Alps. Phylogenetic analyses presented in Section 3.6 suggest a recent speciation within New Zealand Fusca-group beeches, which may reflect a Pliocene diversification. Ecological preferences of *Nothofagus* and the molecular data are consistent with the hypothesis that Pliocene Mountain building had influence on the evolution of the New Zealand forest flora.

7.3.4 The Pleistocene Ice-Ages

Numerous studies have been undertaken to investigate the effects of Pleistocene climate changes on the biota in the Northern and Southern Hemispheres (e.g. Comes and Kadereit, 1998; Wardle, 1988a, and references therein). Palynological and, more recently, plant molecular studies have focused on the localisation of glacial refugia and the reconstruction of postglacial re-colonisation. While a relatively large number of molecular studies have been concerned with situations in the Northern Hemisphere (e.g. Ferris *et al.*, 1995; Tollefrud *et al.*, 1998), little has been done to investigate effects of Pleistocene climates in the Southern Hemisphere using molecular data.

Effects of Pleistocene climates on the New Zealand flora have been discussed by Cockayne (1928), Willet (1950), Wardle (1963; 1988a) and McGlone (1980; 1985). Existing hypotheses suggest that forest plants have survived cooler glacial stages in Northland in pockets of suitable microclimates scattered throughout more southerly areas which were dominated by grass and shrub formations (Wardle, 1988a; McGlone *et al.*, 2001). The cooler temperatures and the glaciation of mountainous regions may have caused disruptions in formerly continuous plant distributions, populations may have contracted and species were reduced to smaller numbers. A geographic disjunction of populations may cause genetic differentiation, which may be followed by speciation. Changes in the environment may favour hybridisation between closely related taxa.

From the plant groups that were studied in this thesis, recent radiation events were observed within all New Zealand *Fuscaspora* Beeches and between some species of *Myrsine*. In this group, recent genetic differentiation was detected between northern and southern ecotypes of *M. divaricata* as well as between *M. divaricata* and *M. australis*. Relatively low genetic differentiation and the occurrence of interspecific hybrids (Wardle, 1988b; and shown in this thesis) indicate recent speciation events, which may have accompanied climatic changes during the Pliocene and/or Pleistocene.

Recent genetic differentiation was also observed between disjunct populations of *Nothofagus menziesii* (Chapter 6). Although phylogenetic analyses suggest that populations were disjunct by the beginning of the Pleistocene, the results are consistent with the hypothesis of disruption by Pleistocene glacial stages as suggested by Cockayne (1928), Willet (1950) and Wardle (1963; 1988a).

7.4 CONCLUSIONS

Palynological and palaeobotanical studies have suggested that biogeographic relationships of the New Zealand flora are complex (e.g. Fleming, 1979). Results from DNA analyses presented in this thesis confirm that evolutionary histories of forest plants are variable and may be unique for every single plant lineage. It could be shown that the New Zealand forest hosts tree species with both a Gondwanan origin as well as a dispersal origin. DNA analyses support the hypothesis that taxa such as *Agathis australis* Salisbury, *Dacrydium cupressinum* Lamb. and *Pseudowintera* spp. Dandy have evolved as vicariant lineages in New Zealand and on other southern hemisphere lands after the break-up of the ancient supercontinent. These molecular results refute the hypothesis that the New Zealand plant cover was eliminated by marine transgression of the New Zealand landmass during the Oligocene as suggested by Pole (1994).

Molecular studies presented in this thesis also support observations from fossil evidence that many plant lineages have arrived in New Zealand after the break-up of Gondwana by long-distance dispersal across ocean gaps (e.g. Fleming, 1979; Macphail, 1979b). Non-vicariant lineages include *Myrsine* spp. L., *Nothofagus* Blume, *Knightia excelsa* R.Br. and *Toronia toru* A. Cunn. Molecular clock tests on DNA sequence data reject a vicariant evolution for *Nothofagus* species which has been suggested by Darlington (1965), Raven and Axelrod (1972) or Linder and Crisp (1995). They support the view advocated by Hill (1992) that dispersal may account for disjunct species distribution of *Nothofagus* in the Southern Hemisphere.

Studies on New Zealand Myrsinaceae (Section 3.5 and Chapter 5) show that dispersal has also occurred from New Zealand to other landmasses. Phylogenetic results suggested that dispersal from New Zealand was multi-directional. This result corroborates findings by Winkworth (2000) who emphasised the importance of long-distance dispersal between landmasses in explaining modern plant distribution in the Southern Hemisphere.

Molecular studies on closely related taxa of *Myrsine* suggest recent speciation events and ongoing hybridisation between morphologically and ecologically distinct species. Intraspecific analyses on *Myrsine divaricata* and *Nothofagus menziesii* show that distribution patterns within New Zealand are relatively recent and may have developed during the Quaternary. Molecular results from studies on closely related taxa emphasise that evolution and distribution patterns are dynamic processes and suggest that plants respond to geological, climatic and environmental changes like land uplift and glaciation. Possible responses may be observed as genetic differentiation of ecotypes, speciation of taxa and/or hybridisation between morphologically and ecologically distinct species.

7.5 DIRECTIONS FOR FUTURE WORK

Numerous investigations have shown that molecular phylogenetic analyses are a powerful tool for testing hypotheses concerning phylogenetic relationships of plants and animals. Questions on the origins and evolution of many elements in the New Zealand flora still remain a challenging task for further studies. In many cases, molecular analyses in the present work corroborate observations on the first appearance of taxa in the pollen and/or macroflora records. This gives confidence that further DNA analyses may also be helpful to reconstruct and test evolutionary patterns and dispersal routes across southern hemisphere lands for other plant groups. The DNA studies reviewed and reported in this thesis provide evidence that in New Zealand the Oligocene regression; the Pliocene mountain building and the Pleistocene glacial-interglacial cycles have all had an impact on evolution and distribution of the flora and fauna. It will be interesting to further test these ideas with additional molecular markers and to use some of higher resolution molecular marker methods - such as the use of AFLP developed here - to investigate biotic responses of plants and animals to Late Tertiary and Quaternary climatic changes. Such molecular tools may be useful in locating glacial refugia, determining post-glacial history of the New Zealand biota and evaluating the status of New Zealand endemics for conservation purposes.

BIBLIOGRAPHY

Abbott, R. J. (1992). Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology and Evolution* **7**, 401-405.

Allan, H. H. (1947). *Transactions of the Royal Society of New Zealand* **76**, 596.

Allan, H. H. (1982). *Flora of New Zealand*, 1, P.D. Hasselberg Government Printer, Wellington, New Zealand.

Allen, R. B. (1987). Ecology of *Nothofagus menziesii* in the Catlins Ecological region, south east Otago, New Zealand. (I) Seed production, viability and dispersal. *New Zealand Journal of Botany* **25**, 5-10.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database searchprograms. *Nucleic Acids Research* **25**, 3389-3402.

Anderberg, A. A., Trift, I. & Kallersjo, M. (2000). Phylogeny of Cyclamen L. (Primulaceae). *Plant Systematics and Evolution* **220**, 147-160.

APG: The Angiosperm Phylogeny Group (1998). An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Gardens* **85**, 531-553.

Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O. & Zimmer, E. (1980). Molecular evidence for genetic exchanges among ribosomal genes on non-homologous chromosomes in man and apes. *Proceedings of the National Academy of Science USA* **77**, 7323-7327.

Arnold, M. L. (1992). Natural hybridization as an evolutionary process. *Annual Reviews in Ecology and Systematics* **23**, 237-261.

Avise, J. C. (1994). *Molecular markers, natural history and evolution*, Chapman & Hall, New York.

Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7**, 371-379.

Axsmith, B. J., Taylor, T. N. & Taylor, E. L. (1998). Anatomically preserved leaves of the conifer *Notophytum krauselii* (Podocarpaceae) from the Triassic of Antarctica. *American Journal of Botany* **85**, 704-713.

Baldwin, B. G. (1992). Phylogenetic Utility of the internal Transcribed Spacers of Nuclear Ribosomal DNA in Plants: An Example from the Compositae. *Molecular Phylogenetics and Evolution* **1**, 3-16.

Baldwin, B. G. & Sanderson, M. J. (1998). Age and rate of diversification of the Hawaiian silverworth alliance (Compositae). *Proceedings of the National Academy of Science USA* **9**, 9402-9406.

Bale, C. L. & Williams, J. B. (1993). Composition and classification of *Nothofagus moorei* communities in northern New South Wales. *Australian Systematic Botany* **6**, 429-440.

Bandelt, H.-J. & Dress, A. W. M. (1992). Split decomposition: a new and useful approach to phylogenetic analyses of distance data. *Molecular Phylogenetics and Evolution* **14**, 843-853.

Barton, N. H. & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Reviews in Ecology and Systematics* **16**, 113-148.

Bassam, B. J., Caetano-Anolles, G. & Gresshoff, P. M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* **195**, 80-83.

Batt, G. E., Braun, J., Kohn, B. P. & McDougall, I. (2000). Thermochronical analysis of the dynamics of the Southern Alps. *New Zealand Geological Society of America Bulletin* **112**, 250-266.

Beismann, H., Barker, J. H. A., Karp, A. & Speck, T. (1997). AFLP analysis sheds light on distribution of two *Salix* species and their hybrids along a natural gradient. *Molecular Ecology* **6**, 989-993.

Bennet, K. D., Tzedakis, P. C. & Willis, K. J. (1991). Quaternary Refugia of north Eurpean Trees. *Journal of Biogeography* **18**, 103-115.

Bertorelle, G. & Rannala, B. (1998). Using rarer mutations to estimate population divergence times: A maximum likelihood approach. *Proceedings of the National Academy of Science USA* **95**, 15452-15457.

Beu, A. G. & Edwards, A. R. (1984). New Zealand Pleistocene and Late Pliocene Glacio-Eustatic Cycles. *Paleogeography, Paleoclimate, Paleoecology* **46**, 119-142.

Blum, H., Beier, H. & Gross, H. J. (1987). Improved silver staining of plant-proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8**(2), 93-99.

Breitweiser, I., Glenny, D. S., Thorne, A. & Wagstaff, S. J. (1999). Phylogenetic relationships in Australasian Gnaphalieae (Compositae) inferred from ITS sequences. *New Zealand Journal of Botany* **37**, 399-412.

Briggs, B. G. & Walters, S. M. (1984). *Plant variation and evolution.* 2nd edit, Cambridge University Press, Cambridge.

Brunsfeld, S. J., Soltis, P. S., Soltis, D. E., Gadek, P. A., Quinn, C. J., Strenge, D. D. & Ranker, T. A. (1993). Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: Evidence from *rbc*L sequences. *Systematic Botany* **19**, 253-262.

Burbridge, N. T. (1960). The phytogeography of the Australian region. *Australian Journal of Botany* **8**, 75-212.

Burrows, C. J. (1965). Some Discontinuous Distributions of Plants Within New Zealand and their Ecological Significance II: Disjunctions Between Otago-Southland and Nelson-Marlborough and Related Distribution Patterns. *Tuatara* **13**, 9-29.

Burrows, C. J. & Lord, J. M. (1993). Recent colonisation by *Nothofagus fusca* at Cass, Canterbury. *New Zealand Journal of Botany* **31**, 139-146.

Cantrill, D. J. (1992). Araucarian foliage from the Lower Cretaceous of southern Victoria, Australia. *International Journal of Plant Science* **153**, 622-645.

Carlquist, S. J. (1972). Island biology: we've only just begun. *Bioscience* **22**, 221-225.

Carlquist, S. J. (1996). Plant dispersal and the origin of Pacific island floras. In *The origin and evolution of Pacific island biotas, New Guinea to Eastern Polynesia: patterns and processes* (Keast, A. & Miller, S. E., eds.). Academic Publishing, Amsterdam.

Cervera, M. T., Cabezas, J. A., Simon, B., Martinez-Zapater, J. M., Beitia, F. & Cenis, J. L. (2000). Genetic relationships among biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on AFLP analysis. *Bulletin of Entomological Research* **90**, 391-6.

Chaw, S. M., Long, H., Wang, B. S., Zharkikh, A. & Li, W. H. (1993). The phylogenetic position of Taxaceae based on 18S rRNA sequences. *Journal of Molecular Evolution* **37**, 624-630.

Chen, Z. D., Manchester, S. R. & Sun, H. Y. (1999). Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology, and paleobotany. *American Journal of Botany* **86**, 1168.

Clayton, R. N. & Stevens, G. R. (1968). Paleotemperatures of the New Zealand Jurassic and Cretaceous. *Tuatara* **16**, 3-7.

Close, R. C., Moar, N. T., Tomlinson, A. I. & Lowe, A. D. (1978). Aerial dispersal of biological material from Australia to New Zealand. *International Journal of Biometeorology* **22**, 1-19.

Cockayne, L. (1911). Observations concerning evolution derived from ecological studies in New Zealand. *Transactions of the New Zealand Institute* **44**, 1-50.

Cockayne, L. (1919). New Zealand Plants and their Story, Manual No. 1. New Zealand Board of Science and Art.

Cockayne, L. & Atkinson, E. H. (1926). On the New Zealand wild hybrids of *Nothofagus*. *Genetica* **8**, 1-43.

Cockayne, L. (1928). *The vegetation of New Zealand*, Engelmann Verlag, Leipzig.

Cohen, S. N., Chang, A. C. Y. & Hsu, L. (1972). Nonchromosomal antibiotic resistence in bacteria: transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Science USA* **69**, 2110.

Comes, H. P. & Kadereit, J. W. (1998). The effect of Quaternary climate changes on plant distribution and evolution. *Trends in Plant Science* **3**, 432-438.

Conran, J. G., Wood, G. M., Martin, P. G., Dowd, J. M., Quinn, C. J., Gadek, P. A. & Price, R. A. (2000). Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbc*L. *Australian Journal of Botany* **48**, 715-724.

Cookson, I. (1950). Fossil pollen grains of Proteaceous type from Tertiary deposits of Australia. *Australian Journal of scientific Research* **B3**, 166-177.

Cookson, I. & Duigan, S. L. (1950). Fossil *Banksiae yallourn* with notes on the morphology and anatomy of living species. *Australian Journal of scientific Research* **B2**, 133-165.

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

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

Cranwell, L. M. (1963). *Nothofagus*: living and fossil. In *Pacific basin biogeography* (Gressit, J. L., ed.), pp. 387-400. Bishop Museum Press, Honolulu.

Cronquist, A. (1998). www.inform.umd.edu/EdRes/Colleges/LFSC/life_sciences/.plant_biology/-pb250/cron1.html (Reveal, J. L., ed.).

Daniel, I. L., Lovis, J. D. & Reay, M. B. (1988). A brief introductory report on the mid-Cretaceous megaflora of the Clarence valley, New Zealand. *Proceedings of the 3IOP Conference*. 1990.

Daniel, I. L. (1989). Taxonomic investigation of elements from the Early Cretaceous megaflora from the middle Clarence Valley, New Zealand, University of Canterbury.

Darlington, P. J. (1965). *Biogeography of the southern end of the world*, Harvard University Press, Cambridge.

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

Dawson, J. W. (1976). Pacific capsular Myrtaceae XI. Redefinition of *Metrosideros* Banks ex Gaertn. and definition of infrageneric categories. *Blumea* **23**, 7-11.

Dawson, J. W. (1984). New species and combinations in New Caledonian *Metrosideros* and *Carpolepis* (Myrtaceae) with notes on other species. *Adansonia* **4**, 465-489.

de Knijff, P., Denkers, F., van Swelm, N. D. & Kuiper, M. (2001). Genetic affinities within the herring gull *Larus argentatus* assemblage revealed by AFLP genotyping. *Journal of Molecular Evolution* **52**, 85-93.

de Laubenfels, D. J. (1988). Coniferales. In *Flora Malesiana* (Dordrecht, ed.), Series I, Vol. 10., pp. 337-453. Kluwer Academic.

Dettmann, M. E., Pocknall, D., Romero, E. J. & Zamboloa, M. (1990). *Nothofagidites* Erdtmanex Potonie 1960; a catalogue of species with notes on the palaeogeographic distribution of *Nothofagus* Bl. (southern beech). *New Zealand Geological Survey Paleontological Bulletin*. **60**, 1-79.

Devereux, I. (1967). Oxygen isotope Paleotemperature measurements on New Zealand Tertiary fossils. *New Zealand Journal of Science and Technology* **10**, 988-1011.

Donoghue, M. J. & Doyle, J. A. (1989). Phylogenetic analysis of angiosperms and the relationships of Hamamelidae. In *Evolution, Systematics and Fossil History of the Hamamelidae* (Crane, P. & Blackmore, S., eds.), pp. 17-45. Clarendon Press, Oxford.

Doyle, J. J. & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**, 11-15.

Ecroyd, C. E. (1982). Biological floras of New Zealand, 8. *Agathis australis* (D. Don.) Lindl. (Araucariaceae) kauri. *New Zealand Journal of Botany* **20**, 17-36.

Efron, B. (1982). *Proceedings of CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia, PA*.

Erdogan, V. & Mehlenbacher, S. A. (2000). Phylogenetic relationships of *Corylus* species (Betulaceae) based nuclear ribosomal DNA ITS region and chloroplast *mat*K gene sequences. *Systematic Botany* **25,** 727–737.

Ettinghausen. (1887). Beiträge zur Kenntnis der fossilen Flora Neuseelands. *Denkschriften der Akademie der Wissenschaften, Wien* **53**, 143-194.

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

Ferris, C., Oliver, R. P., Davy, A. J. & Hewitt, G. M. (1995). Using chloroplast DNA to trace postglacial migration routes of oaks into Britain. *Molecular Ecology* **4**, 731-738.

Fisher, F. J. F. (1965). *The alpine Ranunculi of New Zealand*. DSIR Bulletin, 165, NZ DSIR.

Fleming, C. A. (1949). The geological history of New Zealand (with reference to the origin and history of the fauna and flora). *Tuatara* **2**, 72-90.

Fleming, C. A. (1979). *The geological history of New Zealand and its life*, Auckland University Press, New Zealand.

Fridolfsson, A.-K., Cheng, H., Copeland, N. G., Jenkins, N. A., Liu, H.-C., Raudsepp, T., Woodage, T., Chowdhary, B., Halverson, J. & Ellegren, H. (1998). Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proceedings of the National Academy of Science USA* **95**, 8147-8152.

Frisch, W. & Loeschke, J. (1993). *Plattentektonik*, Wissensch. Buchgesellschaft, Darmstadt.

Gadek, P. A. & Quinn, C. J. (1993). An analysis of relationships within the Cupressaceae sensu stricto based on *rbc*L sequences. *Annals of the Missouri Botanical Gardens* **80,** 581-586.

Gage, M. (1980). *Legends in the Rocks: An Outline of New Zealand Geology*, Whitcoulls, Christchurch.

Gancheva, A., Pot, B., Vanhonacker, K., Hoste, B. & Kersters, K. (1999). A polyphasic approach towards the identification of strains belonging to *Lactobacillus acidophilus* and related species. *Systematic and Applied Microbiology* **22**, 573-85.

Ganter, P. F. & de Barros Lopes, M. (2000). The use of anonymous DNA markers in assessing worldwide relatedness in the yeast species *Pichia kluyveri* Bedford and Kudrjavzev. *Canandian Journal of Microbiology* **46**, 967-80.

Gaudeul, M., Taberlet, P. & Till-Bottraud, I. (2000). Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. *Molecular Ecology* **9**, 1625-37.

Gerber, S., Mariette, S., Streiff, R., Bodenes, C. & Kremer, A. (2000). Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology* **9**, 1037-48.

Givnish, T. J. (1997). Adaptive radiation and molecular systematics: issues and approaches. In *Molecular evolution and adaptive radiation* (Givnish, T. J. & Sytsma, K. J., eds.), pp. 1-54. Cambridge University Press, Cambridge.

Godley, E. J. (1967). Widely distributed species, land bridges and continental drift. *Nature* **214**, 74-75.

Goldman, N., Anderson, J. P. & Rodrigo, A. G. (2000). Likelihood-based tests of topologies in phylogenetics. *Systematic Biology* **49**, 652-670.

Golenberg, E. M., Clegg, M. T., Durbin, M. L., Doebley, J. & Pow, M. D. (1993). Evolution of a non coding region of the chloroplast genome. *Molecular Phylogenetics and Evolution* **2**, 52-64.

Graham, G. C., Henry, R. J., Godwin, I. D. & Nikles, D. G. (1996). Phylogenetic position of hoop pine (*Araucaria cunninghamii*). *Australian Systematic Botany* **9**, 893-902.

Grant, P. R. (1998). *Evolution on islands*, Oxford University Press, Oxford.

Gugerli, F., Sperisen, C., Büchler, U., Magni, F., Geburek, T., Jeanderoz, S. & Senn, J. (2001). Blackwell Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious founder effect during postglacial re-colonization of the western Alps. *Molecular Ecology* **10**, 1255 –1263.

Haase, P. (1992). Isozyme variability and biogeography of *Nothofagus truncata* (Fagaceae). *New Zealand Journal of Botany* **30**, 315-328.

Hackett, C. A., Wachira, F. N., Paul, S., Powell, W. & Waugh, R. (2000). Construction of a genetic linkage map for *Camellia sinensis* (tea). *Heredity* **85**, 346-355.

Han, T. H., de Jeu, M., van Eck, H. & Jacobsen, E. (2000). Genetic diversity of Chilean and Brazilian *Alstroemeria* species assessed by AFLP analysis. *Heredity* **84**, 564-569.

Hasebe, M., Omori, T., Nakazawa, M., Sano, T., Kato, M. & Iwatzuki, K. (1994). *rbc*L gene sequences provide evidence for the evolutionary linneages of leptosporangiate ferns. *Proceedings of the National Academy of Science USA* **91**, 5730-5734.

Hasegawa, M., Kishino, H. & Yano, T.-A. (1985). Dating of the human - ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 737-739.

Hawkins, B. J. & Sweet, G. B. (1989). Genetic variation in rimu - an investigation using isozyme analysis. *New Zealand Journal of Botany* **27**, 83-90.

Hayes, D. E. & Ringis, J. (1975). The Early Opening of the Tasman Sea (Fraser, R. & McCraw, J. D., eds.), pp. 338.

Heads, M. J. (1989). Integrating earth and life sciences in New Zealand natural history: the parallel arcs model. *New Zealand Journal of Zoology* **16**, 549-585.

Heenan, P. B. & de Lange, P. J. (1998). A new and remarkably local species of *Myrsine* (Myrsinaceae from New Zealand. *New Zealand Journal of Botany* **36**, 381-387.

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

Hickson, R. E., Slack, K. E. & Lockhart, P. J. (2000). Phylogeny recapitulates geography, or why New Zealand has so many species of skinks. *Biological Journal of the Linnean Society* **70**, 415-433.

Higham, T. (1991). *New Zealand's subantarctic Islands: a guide book*, Department of Conservation, Wellington, New Zealand.

Hill, R. S. (1983). Evolution of *Nothofagus cunninghamii* and its relationship to *N. moorei* as inferred from Tasmanian macrofossils. *Australian Journal of Botany* **31**, 453-465.

Hill, R. S. (1987). Discovery of *Nothofagus* fruits corresponding to an important Tertiary pollen type. *Nature* **327**, 56-58.

Hill, R. S. & Read, J. (1987). Endemism in Tasmanian cool temperate rainforest: Alternative hypotheses. *Botanical Journal of the Linnean Society* **95**, 113-124.

Hill, R. S. & Read, J. (1991). A revised infrageneric classification of *Nothofagus* (Fagaceae). *Botanical Journal of the Linnean Society* **105**, 37-72.

Hill, R. S. (1992). *Nothofagus*: Evolution from a Southern Perspective. *Trends in Ecology and Evolution* **7**, 190-194.

Hill, R. S. & Dettmann, M. E. (1996). Origins and diversification of the genus *Nothofagus*. In *The ecology and biogeography of Nothofagus forests* (Veblen, T. T., Hill, R. S. & Read, J., eds.), pp. 11-24. Yale University Press, New Haven.

Hill, R. S., Jordan, G. J. & Macphail, M. K. (1996). History and paleoecology of Australian *Nothofagus* forests. In *The ecology and biogeography of Nothofagus forests* (Veblen, T. T., Hill, R. S. & Read, J., eds.), pp. 182-199. Yale University Press, New Haven.

Hill, R. S. & Brodribb, T. J. (1999). Turner Review No. 2 - Southern conifers in time and space. *Australian Journal of Botany* **47**, 639-696.

Hillis, D. M., Moritz, C., Porter, C. A. & Baker, R. J. (1991). Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science* **251**, 308-310.

Holmes, D. S. & Quigley, M. (1981). A rapid method for the preparation of bacterial plasmids. *Annals of Biochemistry* **114**, 193.

Hooker, J. D. (1853). Flora Novae Zelandiae. Part 1. Introductory essay. In *Botany of the Antarctic voyage of H.M. Discovery ships 'Erebus' and 'Terror' in the years 1831-1843.*, Vol. 2, pp. 209-223. Reeve, London.

Hoot, S. B. & Douglas, 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.

Hornibrook, N. & B, d. (1971). New Zealand Tertiary Climate. *New Zealand Geological Survey Report* **47**, 19.

Hosaka, E. Y. (1940). A revision of the Hawaiian species of *Myrsine* (*Suttonia, Rapanea*) (Myrsinaceae). *Occ. Pap. Bishop Mus* **16**, 25-76.

Hurr, K., Lockhart, P., Heenan, P. B. & Penny, D. (1999). Dispersal of the *Edwardsia* Section of *Sophora* (Leguminosae) around the southern oceans: Molecular Evidence. *Journal of Biogeography* **26**, 565-577.

Huson, D. H. (1998). SplitsTree: analysing and visualising evolutionary data. *Bioinformatics* **14**, 68-73.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M. & Kersters, K. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* **142**, 1881.

Johnson, L. A. S. & Briggs, B. G. (1975). On the Proteaceae - the evolution and classification of a southern family. *The Botanical Journal of the Linnean Society* **70**, 83-182.

Jones, C. J., Edwards, K. J. & Castaglione, S. (1998). Reproducibility testing of AFLPs by a network of European laboratories. In *Molecular Tools for Screening Biodiversity* (Karp, A., Isaac, G. P. & Ingram, S. D., eds.), pp. 191-192. Chapman and Hall, London.

Jones, W. (1995). Wollemi Pine - the missing link? [rediscovery of population of 23 adults and 16 juveniles in Wollemi National Park, N.S.W.]. *On the Brink* **6**, 2.

Jones, W. G., Hill, K. D. & Allen, J. M. (1996). Wollemia nobilis, a new living Australian genus and species in the Araucariaceae. *Telopea* **6**, 2-3.

Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mamalian Protein Metabolism III* (Munro, H. N., ed.), pp. 21-132. Academic Press, New York.

Kelch, D. G. (1997). The Phylogeny of the Podocarpaceae based on morphological evidence. *Systematic Botany* **22**, 1-19.

Kelch, D. G. (1998). Phylogeny of the Podocarpaceae: Comparison of evidence from morphology and 18S rDNA. *American Journal of Botany* **85**, 986-996.

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

Klein, P. E., Klein, R. R., Cartinhour, S. W., Ulanch, P. E., Dong, J., Obert, J. A., Morishige, D. T., Schlueter, S. D., Childs, K. L., Ale, M. & Mullet, J. E. (2000). A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. *Genome Research* **10**, 789-807.

Krauss, S. L. (2000a). Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology* **9**, 1241-1245.

Krauss, S. L. (2000b). The realized effect of postpollination sexual selection in a natural plant population. *Proceedings of the Royal Society of London Bulletin Biological Sciences* **267**, 1925-1929.

Kuschel, G. (1975). *Biogeography and ecology in New Zealand*. Monographiae Biologicae (Illies, J., Ed.), Dr. W. Junk b.v. Publishers, The Haague.

Lake, J. A. (1991). The order of sequence alignment can bias the selection of tree topology. *Molecular Biolology and Evololution* **8**, 378-385.

Laurent, N., Bremer, B. & Bremer, K. (1999). Phylogenetic and generic interrelationships of the Stylidiaceae (Asterales), with a possible extreme case of floral paedomorphosis. *Systematic Botany* **23**, 289-304.

Leathwick, J. R. (1998). Are New Zealand's *Nothofagus* species in equilibrium with their environment? *Journal of Vegetation Science* **9**, 719-732.

Lee, J. C., Cole, M. & Linacre, A. (2000). Identification of hallucinogenic fungi from the genera *Psilocybe* and *Panaeolus* by amplified fragment length polymorphism. *Electrophoresis* **21**, 1484-1487.

Leroy, J.-F. (1980). Nouvelles remarques sur le genre *Takhtajania* (Winteraceae - Takhtananioidies). *Adansonia* **20**, 9-20.

Lin, J.-J. & Kuo, J. (1995). AFLP: A novel PCR-based assay for plant and bacterial DNA fingerprinting. *Focus* **17**, 66-70.

Linder, H. P. & Crisp, M. D. (1995). *Nothofagus* and Pacific biogeography. *Cladistics* **11**, 5-32.

Llyod, D. G. (1972). A revision of the New Zealand, Subantarctic, and South American species of *Cotula* section *Leptinella*. *New Zealand Journal of Botany* **10**, 277-372.

Lockhart, P. J., Beanland, T. J., Howe, C. J. & Larkum, A. W. D. (1992). Substitutional bias confounds inference of cyanelle origins from sequence data. *Journal of Molecular Evolution* **34**, 153-162.

Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. (1994). Recovering evolutionary trees under more realistic model of sequence evolution. *Molecular Biology and Evolution* **11**, 606-621.

Lockhart, P. J., Larkum, A. W. D., Steel, M. A., Waddell, P. J. & Penny, D. (1996). Evolution of chorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis. *Proceedings of the National Academy of Science USA* **94**, 1930-1934.

Lockhart, P. J. & McLenachan, P. A. (1997). Isolating polymorphic plant DNA fragments identified using AFLP^{TM} technology without acrylamide gels: markers for evolutionary studies. *Focus* **19**, 70-71.

Lockhart, P. J., McLenachan, P. A., Havell, D., Glenny, D. S., Huson, D. & Jensen, U. (2001). Phylogeny, dispersal and radiation of New Zealand alpine buttercups: molecular evidence under split decomposition. *Annals of the Missouri Botanical Gardens* **in press**.

Luo, H. & Boutry, M. (1995). Phylogenetic relationships within *Hevea brasilensis* as deduced from a polymorphic mitochondrial DNA region. *Theoretical and Applied Genetics* **91**, 876-884.

Macphail, M. K., Alley, N., Truswell, E. M. & Sluiter, I. R. (1994). Early Tertiary vegetation: evidence from spores and pollen. In *History of the Australian Vegetation: Cretaceous to Recent* (Hill, R. S., ed.), pp. 189-262. Cambridge University Press, Cambridge.

Macphail, M. K. (1997a). Late Neogene climates in Australia: Fossil pollen- and spore-based estimates in retrospect and prospect. *Australian Journal of Botany* **45**, 425-464.

Macphail, M. K. (1997b). Comment on M. Pole (1994): "The New Zealand flora - entirely long-distance dispersal?". *Journal of Biogeography* **24**, 113-114.

Manos, P. (1997). Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. *American Journal of Botany* **84**, 1137.

Mark, A. F. & Adams, N. M. (1995). *New Zealand alpine plants*, Godwit Publishing Ltd, Auckland.

Martin, P. G. & Dowd, J. M. (1993). Using Sequences of *rbc*L to study phylogeny and biogeography of *Nothofagus* species. *Australian Journal of Botany* **6**, 441-447.

McCraw, J. D. (1975). Quarternary airfall deposits in New Zealand. In *Quarternary Studies* (Suggate, R. P. & Cresswell, M. M., eds.), Vol. 13, pp. 35-44. The Royal Society of the New Zealand Bulletin.

McGlone, M. S., Ed. (1980). Lowland forests in New Zealand. Edited by Thompson, K., Hodder, A. P. H. & Edmonds, A. S. University of Waikato, Hamilton, New Zealand: University of Waikato.

McGlone, M. S. (1985). Plant biogeography and the Late Cenozoic history of New Zealand. *New Zealand Journal of Botany* **23**, 723-749.

McGlone, M. S., Mildenhall, D. C. & Pole, M. S. (1996). History and paleoecology of New Zealand *Nothofagus* forests. In *The ecology and biogeography of Nothofagus forests* (Veblen, T. T., Hill, R. S. & Read, J., eds.), pp. 83-130. Yale University Press, New Haven.

McGlone, M. S., Duncan, R. P. & Heenan, P. B. (2001). Endemism, species selection and the origin and distribution of the vascular flora of New Zealand. *Journal of Biogeography* **28**, 199-216.

McLenachan, P. A., Stöckler, K., Winkworth, R. C., McBreen, K., Zauner, S. & Lockhart, P. J. (2000). Markers derived from amplified fragment length polymorphism gels for plant ecology and evolution studies. *Molecular Ecology* **9**, 1899-1903.

Melotto, M., Afanador, L. & Kelly, J. D. (1996). Development of a SCAR marker linked to the I gene in common bean. *Genome* **39**, 1216-1219.

Merril, C. R., Goldman, D., Sedman, S. A. & Ebert, M. H. (1981). Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science* **211**, 1437-1438.

Mez. (1902). *Pflanzenreich* **9**.

Mildenhall, D. C. (1980). New Zealand late Cretaceous and Cenozoic plant biogeography: a contribution. *Paleogeography, Paleoclimate, Paleoecology* **31**, 197-233.

Mildenhall, D. C. & Pocknall, D. (1989). Miocene-Pleistocene spores and pollen from Central Otago, South Island, New Zealand. *New Zealand Geological Survey Paleontological Bulletin* **59**, 62-64.

Miller, C. N. (1977). Mesozoic conifers. *Botanical review* **43**, 217-280.

Miller, C. N. (1988). The origin of modern conifer origins. In *Origin and evolution of gymnosperms* (Beck, C. B., ed.), pp. 448-486. Columbia University Press, New York.

Mitchell, A. D. & Wagstaff, S. J. (1997). Phylogenetic relationships of Pseudopanax species (Araliaceae) inferred from parsimony analysis of rDNA sequence data and morphology. *Plant Systematics and Evolution* **208**, 121-138.

Mitchell, A. D., Webb, C. J. & Wagstaff, S. J. (1998). Phylogenetic relationships of species of *Gingidia* and related genera (Apiaceae, subfamily Apioideae). *New Zealand Journal of Botany* **36**, 417-424.

Mitchell, A. D. & Heenan, P. B. (2000). Systematic relationships of New Zealand endemic Brassicaceae inferred from nrDNA ITS sequence data. *Systematic Botany* **25**, 98-105.

Moar, N. T. (1970). Contributions to the Quarternary history of the New Zealand flora: Aranuian pollen diagrams from Canterbury, Nelson and North Westland, Southland. *New Zealand Journal of Botany* **9**, 80-145.

Molloy, B. P. J. (1995). *Manoao* (Podocarpaceae), a new monotypic conifer genus endemic to New Zealand. *New Zealand Journal of Botany* **33**, 183-201.

Molloy, B. P. J. (1996). A new species name in *Phyllocladus* (Phyllocladaceae) from New Zealand. *New Zealand Journal of Botany* **34**, 287-297.

Molloy, B. P. J. & Markham, K. R. (1999). A contribution to the taxonomy of Phyllocladus (Phyllocladaceae) from the distribution of key flavonoids. *New Zealand Journal of Botany* **37**, 375-382.

Molnar, R. E. (1980). *Proceedings of the Fifth International Gondwana Symposium, Wellington, New Zealand.*

Molvray, M., Kores, P. J. & Chase, M. W. (1999). Phylogenetic relationships within *Korthalsella* (Viscaceae) based on nuclear ITS and plastid *trn*L-F sequence data. *American Journal of Botany* **86**, 249-260.

Moore, D. M. (1971). Connections within cool temperate floras with particular reference to southern South America. In *Taxonomy, phytogeography, and evolution* (Valentine, D. H., ed.), pp. 115-138. Academic Press, London.

Moore, J. G. & Clague, D. A. (1992). Volcano growth and evolution of the island of Hawaii. *Geological Society of America Bulletin* **104**, 1471-1484.

Moores, E. M. (1990). *Shaping the earth - tectonics of continents and oceans,* W.H. Freeman and Company, New York.

Morrison, D. A. & Ellis, J. T. (1997). Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa. *Molecular Biology and Evolution* **14**, 428-41.

Mueller, U. G. & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* **14**, 389-394.

Mullis, K. B., Faloona, F. A., Scharf, S. J., Saiki, R. K., Horn, G. T. & Erlich, H. A. (1986). *Cold Spring Habor Symp. Quant. Biol.*

Mullis, K. B. & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**, 335-350.

Nair, S., Schreiber, E., Thong, K. L., Pang, T. & Altwegg, M. (2000). Genotypic characterization of Salmonella typhi by amplified fragment length polymorphism fingerprinting provides increased discrimination as compared to pulsed-field gel electrophoresis and ribotyping. *Journal of Microbiological Methods* **41**, 35-43.

Neyland, R. & Urbatsch, L. E. (1996). The *ndh*F chloroplast gene detected in all vascular plant divisions. *Planta* **200**, 273-277.

Oliver, W. R. B. (1951). *Rec. Auck. Inst. Mus.* **4**, 111.

Olmstead, R. G. & 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 Gardens* **82**, 176–193.

Page, R. D. M. & Holmes, E. C. (1998). *Molecular Evolution - A Phylogenetic Approach*, Blackwell Science, Oxford.

Pasada, D. & Crandall, K. A. (2001). Selecting the best-fit model of nucleotide substitution. *Systematic Biology* **50**, 580-601.

Penny, D., Hendy, M. D. & Steel, M. A. (1992). Progress with evolutionary trees. *Trends in Ecology and Evolution*, 73-79.

Perrie, L. R., Lockhart, P. J., Brownsey, P. J. & Large, M. F. (2000). Morphological and molecular concordance for the recognition of two species in the New Zealand *Polystichum richardii* (Hook.) J. Smith complex. *Plant Systematics and Evolution* **224**, 97-107.

Pole, M. (1993). *Nothofagus* from the Dunedin volcanic group (Mid Late Miocene), New Zealand. *Alcheringa* **17**, 77-90.

Pole, M. S. (1994). The New Zealand flora - entirely long distance dispersal? *Journal of Biogeography* **22**, 625-635.

Pole, M. S. (1995). Late Cretaceous macrofloras of eastern Otago, New Zealand: Gymnosperms. *Australian Systematic Botany* **8**, 1067-1106.

Pole, M. S. (1997). Paleocene plant macrofossils from Kakahu, South Canterbury, New Zealand. *Journal of the Royal Society of New Zealand* **27**, 371-400.

Pole, M. S. (1998). Structure of a near-polar latitude forest from the New Zealand Jurassic. *Palaeogeography, Palaeoecology, Palaeoclimatology* **147**, 121-139.

Pole, M. S. & Douglas, B. (1999). Plant macrofossils of the upper Cretaceous Kaitangata coalfield, New Zealand. *Australian Systematic Botany* **12**, 331-364.

Poole, A. L. & Adams, N. M. (1994). *Trees and shrubs of New Zealand*, Manaaki Whenua Press, Lincoln, New Zealand.

Prather, L. A., Ferguson, C. J. & Jansen, R. K. (2000). Polemoniaceae phylogeny and classification: implications of sequence data from the chloroplast gene *ndh*F. *American Journal of Botany* **87**, 1300-1308.

Preest, D. S. (1963). A note on the dispersal characteristics of the seed of the New Zealand podocarps and beeches, and their biogeographical significance. In *Pacific Basin Biogeography. 10th Pacific Science Congress* (Gressit, J. L., ed.), pp. 415-424. Bishop Museum Press, Honolulu.

Qiu, Y. L., Lee, J., Bernasconi-Quadroni, F., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V. & Chase, M. W. (1999). The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* **402**, 358-359.

Raven, P. H. (1972). Evolution and endemism in the New Zealand species of *Epilobium*. In *Taxonomy, Phytogeography and Evolution* (Valentine, D. H., ed.), pp. 259-274. Academic Press.

Raven, P. H. & Axelrod, D. I. (1972). Plate tectonics and Australasian paleobiogeography. *Science* **176**, 1379-1386.

Raven, P. H. (1973). Evolution of the subalpine and alpine plant groups in New Zealand. *New Zealand Journal of Botany* **11**, 177-200.

Read, J. & Brown, M. J. (1996). Ecology of Australian *Nothofagus* forests. In *The ecology and biogeography of* Nothofagus *forests* (Veblen, T. T., Hill, R. S. & Read, J., eds.), pp. 131-181. Yale University Press, New Haven.

Reveal, J. L. (1993). Reveal System of Classification. http://lpc-linux.unil.ch/botanique/reveal/reve1.html.

Rieseberg, L. H. (1997). Hybrid origins of plant species. *Annual Reviews in Ecology and Systematics* **28**, 359-389.

Robichaux, R. H., Carr, G. D., Liebman, M. & Pearcy, R. W. (1990)b. Adaptive radiation of the Hawaiian silversword alliance (Compositae-Madiinae): Ecological, morphological, and physiological diversity. *Annals of the Missouri Botanical Gardens* **77**, 64-72.

Rodriguez, F., Oliver, J. L., Marin, A. & Medina, J. R. (1990). The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**, 485-501.

Rogers, G. M. (1989). The nature of the lower North Island floristic gap. *New Zealand Journal of Botany* **27**, 221-241.

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

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

Salmon, T. J. (1996). *The native trees of New Zealand* (Walker, G., Ed.), Reed Books, Auckland.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. 2 edit, 1-3, Cold Spring Harbor Laboratory Press, New York.

Sang, T., Crawford, D. J., Kim, S. C. & Stuessy, T. F. (1994). Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez Islands: evidence from sequences of the ITS region of the nuclear ribosomal DNA. *American Journal of Botany* **81**, 1494-1501.

Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* **94**(3), 441-448.

Savard, L., Michaud, M. & Bousquet, J. (1993). Genetic diversity and phylogenetic relationships between birches and alders using ITS, 18S rRNA and *rbc*L gene sequences. *Molecular Phylogenetics and Evolution* **2**, 112-8.

Savolainen, V., Chase, M. W., Morton, C. M., Hoot, S. B., Soltis, D. E., Bayer, C., Fay, M. F., de Bruijn, A., Sullivan, S. & Qiu, Y. L. (2000). Phylogenetics of flowering plants based upon a combined analysis of plastid *atp*B and *rbc*L gene sequences. *Systematic Botany* **49**, 306-362.

Schupp, J. M., Price, L. B., Klevytska, A. & Keim, P. (1999). Internal and flanking sequence from AFLP fragments using ligation-mediated suppression PCR. *Biotechniques* **26**, 905-912.

Seibold, E. & Berger, W. H. (2000). *The sea floor. An introduction to marine geology.* 3rd ed. edit, Springer Verlag.

Setoguchi, H., Ono, M., Doi, Y., Koyama, H. & Tsuda, M. (1997). Molecular phylogeny of *Nothofagus* (Nothofagaceae) based on the *atp*B-*rbc*L intergenic spacer of the chloroplast DNA. *Plant Research* **110**, 469-484.

Setoguchi, H., Osawa, T. A., Pintaud, J.-C., Jaffré, T. & Veillon, J.-M. (1998). Phylogenetic relationships within Araucariaceae based on *rbc*L gene sequences. *American Journal of Botany* **85**, 1507-1516.

Shan, X., Blake, T. K. & Talbert, L. E. (1999). Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat. *Theoretical and Applied Genetics* **98**, 1072-1078.

Sharbel, T. F., Haubold, B. & Mitchell-Olds, T. (2000). Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Molecular Ecology* **9**, 2109-2118.

Sharp, P. A., Sugden, B. & Sambrook, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenza* using analytical agarose. *Biochemistry* **12**, 3055.

Sitte, P., Ziegler, H., Ehrendorfer, F. & Bresinsky, A. (1991). *Strasburger Lehrbuch der Botanik.* 33 edit, Gustav Fischer Verlag, Stuttgart; Jena; New York.

Smith, J. B. (1986). Origins of the Australasian tropicalpine and alpine floras. In *Flora and Fauna of alpine Australasia* (Barlow, B. A., ed.), pp. 109-128. CSIRO, Melbourne.

Soltis, P. S., Soltis, D. E. & Chase, M. W. (1999). Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* **402**, 402-404.

Sproll, W. P. & Dietz, R. S. (1969). Morphological continental drift fit of Australia and Antarctica. *Nature* **222**, 345-348.

Steel, M. (1994). The maximum likelihood tree is not unique. *Journal of Molecular Evolution* **34**, 115-125.

Steel, M. & Penny, D. (2000). Parsimony, likelihood, and the role of models in molecular phylogenetics. *Molecular Biology and Evolution* **17**, 839-850.

Steel, M. A., Cooper, A. C. & Penny, D. (1996). Confidence intervals for the divergence time between two clades. *Systematic Biology* **45**, 127-134.

Steel, M. A., Huson, D. & Lockhart, P. J. (2000). Invariable sites models and their use in phylogeny reconstruction. *Systematic Biology* **49**, 225-232.

Stevens, G., McGlone, M. S. & McCulloch, B. (1995). *Prehistoric New Zealand* (Stevens, G. & Ridge, P., Eds.), Reed Books, Auckland.

Stewart, R. B. & Neall, V. E. (1984). Chronology of paleoclimatic change at the end of the last glaciation. *Nature* **311**, 47-48.

Stockey, R. A. (1982). The Araucariaceae: an evolutionary perspective. *Review of Palaeobotany and Palynology* **37**, 133-154.

Stockey, R. A. (1994). Mesozoic Araucariaceae: Morphology and systematic relationships. *Journal of Plant Research* **107**, 493-502.

Stockey, R. A., Nishida, M. & Nishida, H. (1994). Upper Cretaceous araucarian cones from Hokkaido and Saghalien *Araucaria nipponensis* sp. nov. *International Journal of Plant Science* **155**, 800-809.

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

Suggate, R. P., Stevens, G. R. & Te Punga, M. T. (1978). *Geology of New Zealand.* 2nd volume, Government Printer, Wellington.

Suh, Y., Thien, L. B., Reeve, H. E. & Zimmer, E. A. (1993). Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal DNA in Winteraceae. *American Journal of Botany* **80**, 1042-1055.

Sutton, D. G. (1994). *The origins of the first New Zealanders,* Auckland University Press, Auckland.

Swenson, U. & Bremer, K. (1997). Patterns of floral evolution of four Asteraceae genera (Senecioneae, Blennospermatinae) and the origin of white flowers in New Zealand. *Systematic Botany* **22**, 493-508.

Swenson, U., Backlund, A., McLoughlin, S. & Hill, R. S. (2001). *Nothofagus* biogeography revisited with special emphasis on the enigmatic distribution o subgenus *Brassospora* in New Caledonia. *Cladistics* **17**, 28-47.

Swenson, U. & Hill, R. S. (2001). Most parsimonious araeograms versus fossils: the case of *Nothofagus* (Nothofagaceae). *Australian Journal of Botany* **49**, 367-376.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996). Phylogenetic inference. In *Molecular Systematics* 2 edit. (Hillis, D. M., Moritz, C. & Mable, B. K., eds.), pp. 407-514. Sinauer Associates.

Swofford, D. L. (1998). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods) 4th edit. Sinaur Associates., Massachusetts.

Swofford, D. L. (1999a). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods) 4.65 edit. Sinaur Associates.

Swofford, D. L. (1999b). *PAUP 4.65 Manual*, Sinaur, Sunderland, Massachusetts.

Sykes, W. R. & Godley, E. J. (1968). Transoceanic dispersal in *Sophora* and other genera. *Nature* **218**, 495-496.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Tarling, D. H. (1971). Gondwanaland, paleomagnetism and continental drift. *Nature* **229**, 17-21.

Tautz, D. (1989). Hypervariabilty of simple sequences as a general source for polymorphic DNA markers. *Nucleid Acid Research* **17**, 6463-6471.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **24**, 4876-4882.

Tollefrud, M. M., Bachmann, K., Jakobsen, K. S. & Brochmann, C. (1998). Glacial survival does not matter - II: RAPD phylogeography of nordic *Saxifraga cespitosa*. *Molecular Ecology* **7**, 1217-1232.

Trewick, S. A., Wallis, G. P. & Morgan-Richards, M. (2000). Phylogeographical pattern correlates with Pliocene mountain building in the alpine scree weta (Orthoptera, Anostostomatidae). *Molecular Ecology* **9**, 657 -666.

Ungerer, M. C., Baird, S. J. E., Pan, J. & Rieseberg, L. H. (1998). Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Science USA* **95**, 11757-11762.

Van Steenis, C. G. (1962). The land bridge theory in botany with particular reference to tropical plants. *Blumea* **11**, 235-542.

Vijverberg, K., Mes, T. H. M. & Bachmann, K. (1999). Chloroplast DNA evidence for the evolution of *Microseris* (Asteraceae) in Australia and New Zealand after long-distance dispersal from western North America. *American Journal of Botany* **86**, 1448-1463.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Homes, M., Fritjers, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: A new concept for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407.

Waddell, V. G., Milinkovitch, M. C., Berube, M. & Stanhope, M. J. (2000). Molecular phylogenetic examination of the delphinoidea trichotomy: congruent evidence from three nuclear loci indicates that porpoises (Phocoenidae) share a more recent common ancestry with white whales (Monodontidae) than they do with true dolphins (Delphinidae). *Molecular Phylogenetics and Evolution* **15**, 314-8.

Wagstaff, S. J. & Garnock-Jones, P. J. (1998). Evolution and biogeography of the *Hebe* complex (Scrophulariaceae) inferred from ITS sequences. *New Zealand Journal of Botany* **36**, 425-437.

Wagstaff, S. J., Heenan, P. B. & Sanderson, M. J. (1999). Classification, origins, and patterns of diversification in New Zealand Carmichaelineae (Fabaceae). *American Journal of Botany* **86**, 1346-1356.

Wagstaff, S. J. & Dawson, M. I. (2000). Classification, origin, and patterns of diversification of *Corynocarpus* (Corynocarpaceae) inferred from DNA sequences. *Systematic Botany* **25**, 134-149.

Wakely, J. (1996). The excess of transitions among nucleotide sucstitutions: new methods of estimating transition bias underscore its significance. *Trends in Ecology and Evolution* **11**, 158-163.

Walker, J. W., Brenner, G. J. & Walker, A. G. (1983). Winteraceous pollen in the lower Cretaceous of Israel: early evidence of a Magnolialean angiosperm family. *Science* **220**, 1273-1275.

Wardle, J. (1984). The New Zealand beeches. Ecology, utilization and management. New Zealand Forest Service, Christchurch.

Wardle, P. (1963). Evolution and distribution of the New Zealand Flora, as effected by Quaternary climates. *New Zealand Journal of Botany* **1**, 3-17.

Wardle, P. (1967). Biological flora of New Zealand 2. *Nothofagus menziesii* (Hook. F) Oerst. (Fagaceae) Silver Beech. *New Zealand Journal of Botany* **5**, 276-302.

Wardle, P. (1978). Origins of the New Zealand mountain flora, with special reference to trans-Tasman relationships. *New Zealand Journal of Botany* **16**, 535-550.

Wardle, P. (1988a). Effects of glacial climates on floristic distribution in New Zealand 1. A review of evidence. *New Zealand Journal of Botany* **26**, 541-555.

Wardle, P. (1988b). Effects of glacial climates on floristic distribution in New Zealand 2. The role of long distance hybridisation in disjunct distributions. *New Zealand Journal of Botany* **26**, 557-564.

Wardle, P. (1991). *The vegetation of New Zealand*, Cambridge University Press, Cambridge.

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

Wegener, A. (1912). Die Entstehung der Kontinente. *Geologische Rundschau* **3**, 276-292.

Weising, K. & Gardner, R. C. (1999). A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* **42**, 9-19.

Welsh, J. & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleid Acid Research* **18**, 388-396.

Whitmore, T. C. (1980). A monograph of *Agathis*. *Plant Systematics and Evolution* **135**, 41-69.

Willet, R. W. (1950). The New Zealand Pleistocene snow line, climatic conditions, and suggested biological effects. *New Zealand Journal of Science and Technology* **32B**, 18-46.

Winkworth, R. C., Robertson, A. W., Ehrendorfer, F. & Lockhart, P. J. (1999). The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography* **26**, 1323-1325.

Winkworth, R. C. (2000). Evolution of the New Zealand alpine flora: origins, diversification and dispersal. PhD thesis, Massey University.

Winkworth, R. C., Wagstaff, S. J., Glenny, D. S. & Lockhart, P. J. (2001). Plant Dispersal N.E.W.S. from the South Pacific. *Trends in Ecology and Evolution* **in press**.

Wolfe, A. D., Xiang, Q.-Y. & Kephart, S. R. (1998a). Assessing hybridization in natural populations of Penstemon (Scrophulariaceae) using hypervariable inter simple sequence repeat markers. *Molecular Ecology* **7**, 1107-1125.

Wolfe, A. D., Xiang, Q.-Y. & Kephart, S. R. (1998b). Diploid hybrid speciation in *Penstemon* (Scrophulariaceae). *Proceedings of the National Academy of Science USA* **95**, 5112-5115.

Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987). Rates of nulceotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNA. *Proceedings of the National Academy of Science USA* **84**, 9054-9058.

Wright, S. D., Yong, C. G., Dawson, J. W., Whittaker, D. J. & Gardner, R. C. (2000). Riding the ice age El Nino? Pacific biogeography and evolution of *Metrosideros* subg. *Metrosideros* (Myrtaceae) inferred from nuclear ribosomal DNA. *Proceedings of the National Academy of Science USA* **97**, 4118-4123.

Yang, Z., Goldman, N. & Friday, A. (1994). Comparison of models for nucleotide substitution used in maximum likelihood phylogenetic estimation. *Molecular Biology and Evolution* **11**, 316-324.

Yang, Z. (1996). Among-site variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution* **11**, 367-371.

Yao, Z. Q., Liu, L. J., Rothwell, G. W. & Mapes, G. (2000). Szecladia new genus, a late Permian conifer with multiveined leaves from South China. *Journal of Palaeontology* **74**, 524-531.

Zabeau, M. & Vos, P. (1993). Selective restriction fragment amplification: a general method for DNA fingerprinting. In *European Patent Application no. 92402629.7*.

APPENDIX I

Geological Time Scale

Era	Period		Epoch	Duration	Years Before Present
CAINOZOIC	QUATERNARY		Holocene (Recent)	10 000 yrs	
			Pleistocene	1.5-1.8 m yrs	1.5-1.8 m yrs
	TERTIARY	Neogene	Pliocene	3-5 m yrs	5-7 m yrs
			Miocene	19 m yrs	26 m yrs
		Palaeogene	Oligocene	11-12 m yrs	37-38 m yrs
			Eocene	16 m yrs	53-54 m yrs
			Palaeocene	10 m yrs	64-65 m yrs
MESOZOIC	CRETACEOUS	Late	Senonian- Maastrichian Campanian Santonian Coniacian	35 m yrs	
			Turonian Cenomanian		100 m yrs
		Early	Albian Aptian Barremian Hauterivian Valanginian Berriasian	36 m yrs	136 m yrs
	JURASSIC			54-59 m yrs	190-195 m yrs
	TRIASSIC			33 m yrs	225 m yrs
PALAEOZOIC	PERMIAN			55 m yrs	280 m yrs
	CARBONIFEROUS			65 m yrs	345 m yrs
	DEVONIAN			50 m yrs	395 m yrs
	SILURIAN			35-45 m yrs	430-440 m yrs
	ORDOVICIAN			60-70 m yrs	500 m yrs
	CAMBRIAN			70 m yrs	570 m yrs

APPENDIX II

Herbarium Voucher Numbers of Samples Used For Molecular Analyses

Araucariaceae:

species	sample	origin	voucher
Agathis atropurpurea	1	cultivated at the DRCI Atherton, QLD, Australia	MPN 25010
	2	cultivated at the DRCI Atherton, QLD, Australia	MPN 25011
Agathis australis	1	cultivated at the Esplanade Gardens, PNth, NZ	
	2	Terzaghi Farm, Warkworth,NZ	MPN 25018
Agathis mycrostachya	1	cultivated at the DRCI Atherton, QLD, Australia	MPN 25012
	2	cultivated at the DRCI Atherton, QLD, Australia	MPN 25013
Agathis robusta	1	cultivated at the DRCI Atherton, QLD, Australia	MPN 25014

Myrsinaceae:

species	sample	origin	voucher
Myrsine africana	Maf	South Africa	MPN 24775
Myrsine argentea	My_arg1	Nelson, Mt Bumett, SI, NZ	MPN 24759
	My_arg2	Nelson, Mt Bumett, SI, NZ (Percy's)	MPN 24760
Myrsine australis	MyA_LR1	Lake Rotopunamu, NI, NZ	MPN 24701
	MyA_LR2	Lake Rotopunamu, NI, NZ	MPN 24702
	MyA_To	Pohangina Valley, Totara Reserve, NI, NZ	MPN 24703
	MyA_AT	Abel Tasman Natl. Park, SI, NZ	MPN 24704
	MyA_Kai	Kaikoura, SI, NZ	MPN 24705
	MyA_Hutt1	Mt. Hutt, SI, NZ	
	MyA_Hutt2	Mt. Hutt, SI, NZ	MPN 24706
	MyA_PF	Peel Forest, SI, NZ	MPN 24707
	MyA_StI	Freshwater Bay, Stewart Island, NZ	
Myrsine chathamica	MyC_1	Chatham Islands, NZ (Otari)	MPN 24708
	MyC_2	Chatham Islands, NZ	MPN 24709
	MyC_3	Chatham Islands, NZ	MPN 24710
Myrsine coxii	Mycox_1	Chatham Islands, cultivated at Lincoln, NZ	MPN 24711
	Mycox_2	Chatham Islands, cultivated at Lincoln, NZ	MPN 24712
	Mycox_3	Chatham Islands, cultivated at Lincoln, NZ	MPN 24713

species	sample	origin	voucher
Myrsine divaricata	MyD_AI	Auckland Islands (Otari)	MPN 24714
	MyD_StI1	Stewart Island, Maori Beach	MPN 24715
	MyD_StI2	Stewart Island, Freshwater Bay	MPN 24716
	MyD_4	Invercargill, Otatata, SI, NZ	MPN 24717
	MyD_5	Catlins, McLennan Range, Waikawa, SI, NZ	MPN 24718
	MyD_6	Lill Burn Valley, road to Hauroko, SI, NZ	MPN 24719
	MyD_7	Lake Hauroko, lake shores, SI, NZ	MPN 24720
	MyD_8	Breaksea Island, SI, NZ	MPN 24721
	MyD_9	Manapouri, Supply Cove, SI, NZ	MPN 24722
	MyD_10	Eglington Valley, SI, NZ (Otari)	MPN 24723
	MyD_11	Makaroa, Haast - Hawera Rd,SI, NZ	MPN 24724
	MyD_12	Lake Paringa, SI, NZ	MPN 24725
	MyD_13	Fox Glacier, SI, NZ	
	MyD_14	Franz Joseph Glacier, SI, NZ	MPN 24726
	MyD_15	Otira Gorge, SI, NZ	MPN 24727
	MyD_16	Castle Hill, Arthur's Pass, SI, NZ	MPN 24728
	MyD_17	Lewis Pass, SI, NZ (Otari)	MPN 24729
	MyD_18	Peel Forest Otago, SI, NZ	MPN 24730
	MyD_19	Mt. Oxford Otago, SI, NZ	MPN 24731
	MyD_20	Maruia Valley, SI, NZ	MPN 24732
	MyD_21	Cobb Valley, SI, NZ (Percy's)	MPN 24733
	MyD_23	Cobb River, SI, NZ	
	MyD_24	Cobb River, SI, NZ	
	MyD_25	Cobb River, SI, NZ	
	MyD_BL1	Buckland Track, SI, NZ	MPN 24734
	MyD_BL2	Buckland Track, SI, NZ	MPN 24735
	MyD_BL3	Buckland Track, SI, NZ	MPN 24736
	MyD_BL5	Buckland Track, SI, NZ	MPN 24737
	MyD_Hutt0	Mt. Hutt, SI, NZ	MPN 24738
	MyD_Hutt1	Mt. Hutt, SI, NZ	MPN 24739
	MyD_Hutt2	Mt. Hutt, SI, NZ	MPN 24740
	MyD_Hutt3	Mt. Hutt, SI, NZ	MPN 24741
	MyD_Hutt4	Mt. Hutt, SI, NZ	MPN 24742
	MyD_Hutt5	Mt. Hutt, SI, NZ	MPN 24743
	MyD_37	Tararua Ranges, Mt.Holdsworth, NI, NZ	MPN 24744
	MyD_He1	Tararua Ranges, Herepai, NI, NZ	MPN 24745
	MyD_He2	Tararua Ranges, Herepai, NI, NZ	MPN 24746
	MyD_40	Reporoa Bog, , NI, NZ	MPN 24747
	MyD_41	Ruahine Natl. Forest, Rangi Hut, NI, NZ	MPN 24748
	MyD_42	Hunterville, NI, NZ	MPN 24749
	MyD_43	Tongariro Natl. Park, Mt. Taranaki, NI, NZ	MPN 24750
	MyD_44	Tongariro Natl. Park, Mt. Ruapehu, NI, NZ	MPN 24751
	MyD_45	Tongariro Natl. Park, Whakapapa, NI, NZ	MPN 24752
	MyD_46	Urewera Natl. Park, NI, NZ	MPN 24753
	MyD_47	Woodhill, NI, NZ	
	MyD_48	Bay of Islands,Waitangi, NI, NZ	MPN 24754
	MyD_49	Northland, Unahi, NI, NZ	MPN 24755
	MyD_PK1	Poor Knights Islands (Otari)	MPN 24756
	MyD_PK2	Poor Knights Islands, Tawhiti Rahi	
	MyD_PK3	Poor Knights Islands (Percy's)	MPN 24757

species	sample	origin	voucher
M. kermadecensis	MyK_1	Kermadec Islands (Otari)	MPN 24700
	MyK_2	Kermadec Islands	MPN 24762
M. maximowizcii	MyMax	Japan	MPN 24763
M. oliveri	MyO_1	Three King's Islands (Otari)	MPN 24764
	MyO_2	Three King's Islands	MPN 24765
	MyO_3	Three King's Islands	MPN 24766
M. salicina	MyS_LR1	Lake Rotopounamu, NI, NZ	MPN 24767
	MyS_LR2	Lake Rotopounamu, NI, NZ	MPN 24768
	MyS_He1	Ruahine Ranges, Herepai, NI, NZ	MPN 24769
	MyS_AT	Abel Tasman Natl. Park, SI, NZ	MPN 24770
	MyS_BL1	Buckland Track, SI, NZ	MPN 24771
	MyS_BL2	Buckland Track, SI, NZ	MPN 24772
	MyS_TeA	Te Aroha	
cf. M. australis X divaricata		cultivated at Lincoln, NZ	MPN 24773
cf. M. divaricata X salicina		cultivated at Lincoln, NZ	MPN 24774
Elingamita johnsonii		Three King's Islands, NZ (Otari)	MPN 24776
Rapanea howhittiana		Australia	MPN 24777
Rapanea porosa		Australia	MPN 24778
Rapanea variabilis		Australia	MPN 24779

Myrtaceae:

species	origin	voucher
Metrosideros angustifolia	South Africa	MPN 25017

Nothofagaceae:

species	sample	origin	voucher
Nothofagus solandri	1	Urewera Natl. Park, NI, NZ	MPN 24990
	2	Kaikoura, Blue Duck Reserve, SI, NZ	MPN 24991
	3	Wainuiomata, NI, NZ	MPN 24992
Nothofagus fusca	1	Kaimanawa Ranges, NI, NZ	MPN 24993
	2	Lawerance, SI, NZ	MPN 24994
Nothofagus truncata	1	Taranaki, NI, NZ	MPN 24995
	2	Tararua Ranges, NI, NZ	MPN 24996
Nothofagus menziesii	1	Catlins Natl. Park, SI, NZ	MPN 24981
	2	Waihaha River, NI, NZ	MPN 24816
Nothofagus alpina		in cultivation, South America	MPN 24997
Nothofagus alessandri		in cultivation, South America	MPN 24998
Nothofagus glauca		in cultivation, South America	MPN 24999
Nothofagus gunnei		in cultivation, Australia	MPN 25000
Nothofagus obliqua		in cultivation, South America	MPN 25001
Nothofagus cunninghamii	1	in cultivation, Australia	MPN 25020
	2	in cultivation, Australia	MPN 25021
Nothofagus moorei	L1	Lamington Natl. Park, QLD, Australia	MPN 25022
	L2	Lamington Natl. Park, QLD, Australia	MPN 25023
	L3	Lamington Natl. Park, QLD, Australia	MPN 25024
	BS1	Barraga Swamp, NSW, Australia	MPN 25025
	BS2	Barraga Swamp, NSW, Australia	MPN 25026
	BS3	Barraga Swamp, NSW, Australia	MPN 25027

APPENDIX III

Herbarium Voucher Numbers of *Nothofagus menziesii* Population Samples

Central North Island:

area	origin	sample	
N1	Te Aroha	1	MPN 24780
		2	MPN 24781
		3	MPN 24782
N2	East Cape	1	MPN 24783
		2	MPN 24784
		3	MPN 24785
		4	MPN 24786
		5	MPN 24787
		6	MPN 24788
		7	MPN 24789
		8	MPN 24790
	Hikurangi	9	MPN 24791
		10	MPN 24792
		11	MPN 24793
		12	MPN 24794
		13	MPN 24795
		14	-
		15	-
N3	Urewera	1	MPN 24797
		2	MPN 24798
		3	MPN 24799
		4	MPN 24800
		5	MPN 24801
		6	MPN 24802
		7	MPN 24803
		8	MPN 24804
		9	MPN 24805
		10	MPN 24806
		11	MPN 24807
N4	Mamaku Forest	1	MPN 24808
		2	MPN 24809
		3	MPN 24810
		4	MPN 24811
N5	Waihaha River	1	MPN 24812
		2	MPN 24813
		3	MPN 24814
		4	MPN 24815
		5	MPN 24816
		6	MPN 24817

area	origin	sample	
N6	Ohakune - Turoa	1	MPN 24818
		2	-
		3	MPN 24819
		4	-
N7	Kaimanawa Forest East	1	MPN 24820
		2	MPN 24821
		3	MPN 24822
		4	MPN 24823
		5	MPN 24824
		6	MPN 24825
		7	MPN 24826

Lower North Island

area	origin	sample		
N8	Tararua Ranges West	1	MPN	24827
		2	MPN	24828
		3	MPN	24829
		4	MPN	24830
		5	MPN	24831
		6	MPN	24832
		7	MPN	24833
		8	MPN	24834
		9	MPN	24835
		10	MPN	24836
		11	MPN	24837
		12	-	
N9	Tararua Ranges East	1	MPN	24699
		2	MPN	24838
		3	MPN	24839
		4	MPN	24840
		5	MPN	24841
		6	MPN	24842
		7	MPN	24843
		8	MPN	24844
		9	MPN	24845
		10	MPN	24846
		11	MPN	24847
		12	MPN	24848
		13	MPN	24849
		14	MPN	24850
		15	MPN	24851
		16	MPN	24852
N10	Rimutaka Ranges	1	MPN	24853
		2	MPN	24854
		3	MPN	24855
		4	MPN	24856
		5	MPN	24857
		6	MPN	24858
		7	MPN	24859
N11	Aorangi Ranges	1	MPN	24860
		2	MPN	24861

Northern South Island:

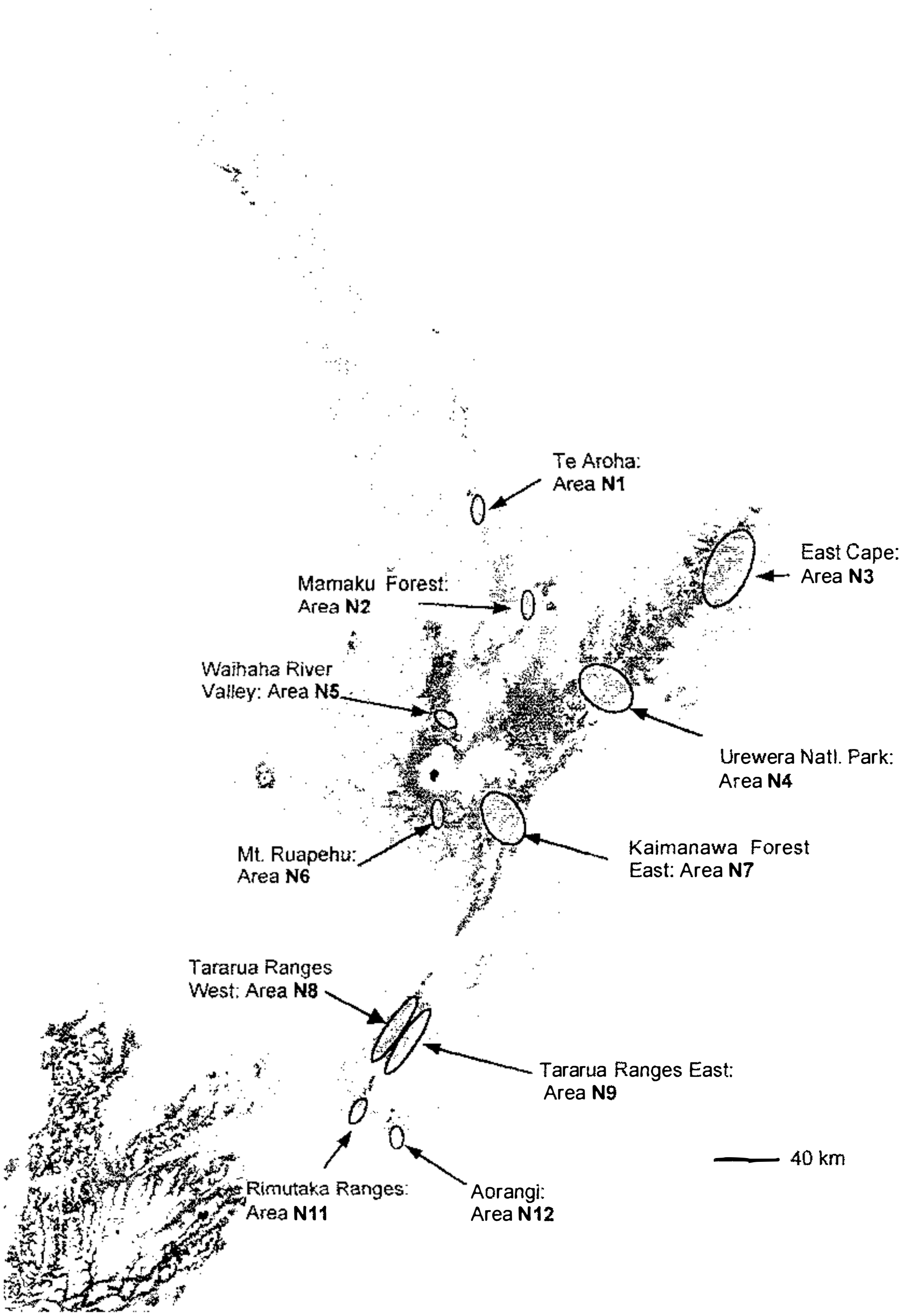
area	origin	sample		
S1	Abel Tasman	1	MPN	24862
		2	MPN	24863
		3	MPN	24864
		4	MPN	24865
	Takaka Hill	5	MPN	24866
		6	MPN	24867
	Arthur Range	7	MPN	24868
		8	MPN	24869
		9	MPN	24870
		10	MPN	24871
S2	Bryant Range	1	MPN	24872
		2	MPN	24873
		3	MPN	24874
		4	MPN	24875
S3	Hope Saddle	1	MPN	24876
		2	MPN	24877
		3	MPN	24878
		4	MPN	24879
	Buller River	5	MPN	24880
		6	MPN	24881
		7	MPN	24882
		8	MPN	24883
		9	MPN	24884
		10	MPN	24885
		11	MPN	24886
		12	MPN	24887
S4	Maruia River	1	MPN	24888
		2	MPN	24889
		3	MPN	24890
		4	MPN	24891
		5	MPN	24892
		6	MPN	24893
S5	Reefton	1	MPN	24894
		2	MPN	24895
		3	MPN	24896
		4	MPN	24897
		5	MPN	24898
		6	MPN	24899
	Paparoa NP	7	MPN	24900
		8	MPN	24901
	Ngahere	9	MPN	24902
		10	MPN	24903
		11	MPN	24904
	Blackball	12	MPN	24905
		13	MPN	24906
		14	MPN	24907
		15	MPN	24908

Southern South Island:

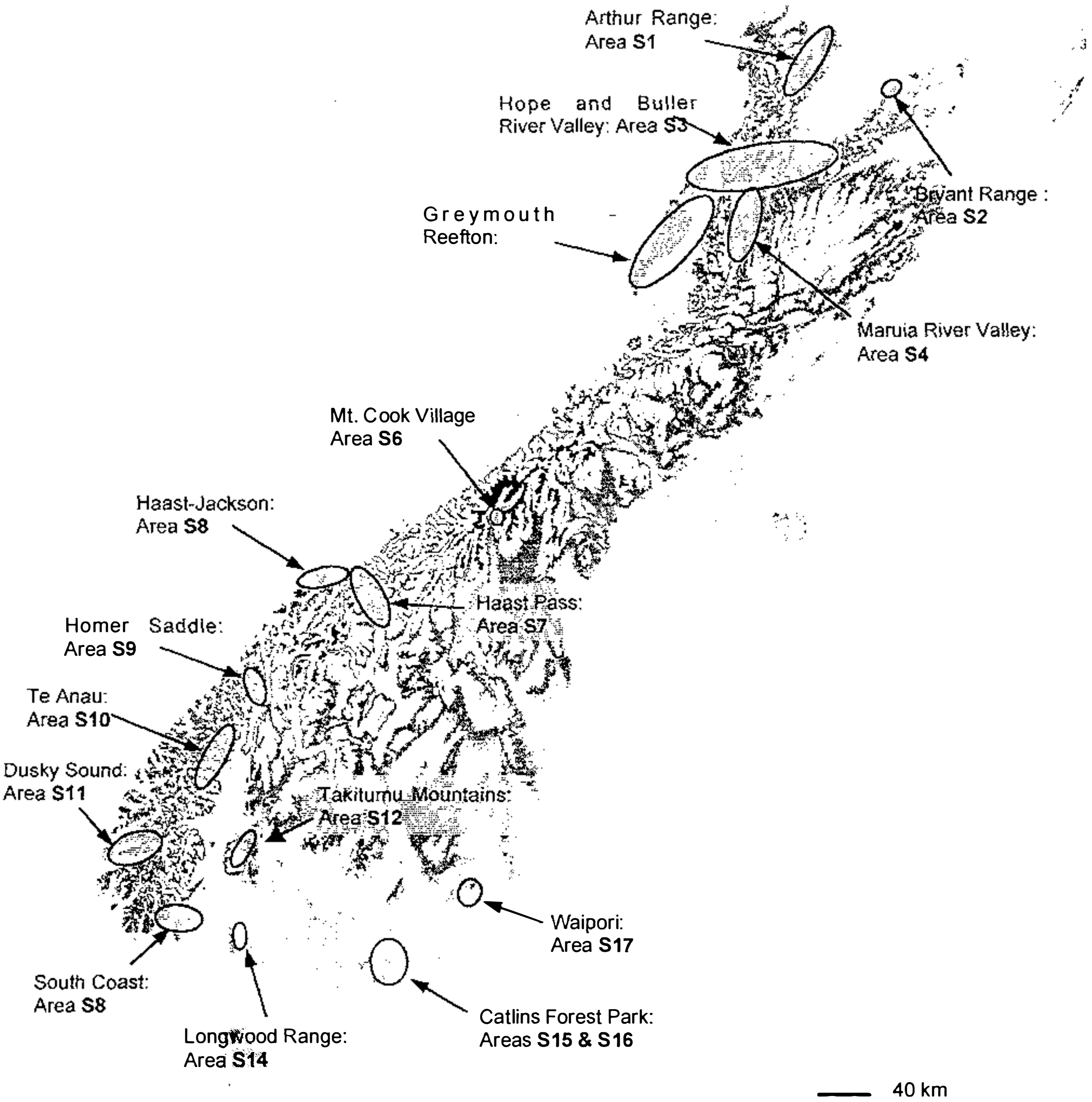
area	origin	sample		
S6	Mt. Cook	1	MPN	24909
S7	Haast	1	MPN	24910
		2	-	
		3	MPN	24911
		4	MPN	24912
		5	MPN	24913
		6	MPN	24914
		7	MPN	24915
		8	MPN	24916
S8	Haast-Port Jackson	1	MPN	24917
		2	MPN	24918
		3	MPN	24919
		4	MPN	24920
		5	-	
		6	MPN	24921
		7	MPN	24922
S9	Homer Saddle	1	MPN	24923
		2	MPN	24924
		3	MPN	24925
		4	MPN	24926
		5	MPN	24927
		6	MPN	24928
		7	MPN	24929
		8	MPN	24930
S10	Te Anau	1	MPN	24931
		2	-	
		3	MPN	24932
S11	Dusky Sound	1	MPN	24933
		2	MPN	24934
		3	MPN	24935
		4	MPN	24936
		5	MPN	24937
		6	MPN	24938
		7	-	
		8	MPN	24939
		9	MPN	24940
		10	MPN	24941
		11	MPN	24942
		12	MPN	24943
S12	Takitumu Range West	1	MPN	24944
		2	MPN	24945
		3	MPN	24946
		4	MPN	24947

area	origin	sample		
S13	South Coast	1	MPN	24948
		2	MPN	24949
		3	MPN	24950
		4	MPN	24951
		5	MPN	24952
		6	MPN	24953
		7	MPN	24954
		8	-	
S14	Longwood Range	1	MPN	24955
		2	MPN	24956
		3	MPN	24957
S15	Catlins West	1	MPN	24958
		2	MPN	24959
		3	MPN	24960
		4	MPN	24961
		5	MPN	24962
		6	MPN	24963
		7	MPN	24964
		8	MPN	24965
		9	MPN	24966
		10	MPN	24967
		11	MPN	24968
		12	MPN	24969
S16	Catlins NF East	1	MPN	24970
		2	MPN	24971
		3	MPN	24972
		4	MPN	24973
		5	MPN	24974
		6	MPN	24975
		7	MPN	24976
		8	MPN	24977
		9	MPN	24978
		10	MPN	24979
		11	MPN	24980
		12	MPN	24981
		13	MPN	24982
		14	MPN	24983
		15	MPN	24984
		16	MPN	24985
S17	Waipori	1	MPN	24986
		2	MPN	24987
		3	MPN	24988
		4	MPN	24989

NORTH ISLAND SAMPLING AREAS FOR SILVER BEECH



SOUTH ISLAND SAMPLING AREAS FOR SILVER BEECH



APPENDIX IV

DNA Sequences of Myrsine AFLP fragments

Fragment: Ma-af1

```
TGTTCAGAAATAGGTGTCATGTGTACATACCTAAAACTGAAATTGTAAATGATCGCTCCGTAAATGTGCA
GATTATTTACTTAGCAATTGCACTAAAGAACGGTGTAGTTTTATCCAGTAAGAAACTGAGTATCAATCTT
AGAGACTCATATGATTAGACTTAGATAATATAGGTCTTAGCATGCAATTCACATCTATCTTTCACTTATT
ATATARCTAATTTGTTTATATTGCTTGTCCTATTTTACTTGTCCTATTTTACTTGTTCGACTTTTTGAAC
TTTATAGRAATTTAGTGCGGAAGTGGCTAAAAAATAGCTAAAAATAATCTCGGATGATCAATCACAAACA
TCAAGAAACAATAAATACAAAAATGAGCCGAAGACATCAAAAAAAATTCCAAATGTCCAAGAATTGAAGA
ATCAGTATAGCAAAAATTGAATAATGCAAGAAATTGGCTAAGGCATTGAAGCTG
```

474 bp

Fragment: Md-af2

```
ATACTATTACGATACTTGGTATGCGATAACTTGTCACCTCCTCATGACGTGAATAAACCTTCATTGATTG
GCCACCGACTCGTGACTCAATCACTAGTCACCCACTCGTGACTCGATAACTAGCCACCTACTCGTGACTC
GATAGTCCATCACCCACTCATGTCAAGACTATGCTTGCAAAGTTCGGTCACCCATTCGTGACTCAACCAA
TGACAAACCCTCCAACTTGGTTTCTCATCTTTCGTAAAGAGATCAAATGAACTCATGGCTCTATGAATTT
ATAAGGGACTACATGCTTCGGTCCCTAAAAGACTCCCACTCCAAAAATGCATAAGTGTGATGTCTTCGAC
ACATGAATGCAATAATTAGTGTTTTCCTG
```

379 bp

Fragment: Md-af3

```
ACGCTCGTGGTATGGGAAAGACTGGTGCTGTCACGGCTGTACCTTCTTGTCAGCTGTGCGGTACAAAATA
TGAGAGTAAACGGTCTTTCGTCAGTCTAGTAAAGGAAGGGGTTCAAAGAGGAGAAATGGGGTTCGAATCC
CACTGTACTAATATCGCCTTACCTCTTTCCAAGACGGAACAAGTACTGCCACTTCTCTCCAATCAATTAG
ATGGTTGTCATCTTTCTCTGTGTTCCCCGGAAGGTTCACGGCCTGGAGCTG
```

261 bp

Fragment: Md-af5

```
ACGCTCATATTATGGTGTTGGATTACTTCTGCAACAACAATTGTTCTCAAGTGATTGTAACGTCGAAAAA
ATAAAGGTGTTCAAATCATATGAACTAGAGAACGCTACTAATCATTACAATGAGGATCGAGTGCTTGGAC
AAGGAGGTCAAGGTACTGTTTACAAAGGAATCTTGCCCAATGGGAGGCTCATTGCTG
```

197 bp

Fragment: Md-af7

```
CAGCAGATTCAGAGATATGGGTATAAGTTCTGAATTTAGGTGTATATAAGGAGCAGCTGAACCTCTAAAG
GCTGCCTACGTATCCCGATGTCGGGAATCAAGCTGCAATTGACGTAGTTCTGCAATTTCAGTCTGAACTG
AATTTATCGGTACCAGAATATAAATGTTGGAACCGAGAGATATACATCATAACCTTGAATCCGGAATCGG
AATCAGGTATACATGACTGAATAAATCGTCAGTTCCGAAATTGAACGTCGGAATAGGTCGGAAAGATGAA
CATAGCAACCAAGCGATGAACGTCTCCCCGGGGATTGAACATCTACCTAGCTACGT
```

336 bp

Fragment: Md-af8

CAGAATGGGATATAGCCCCCGATGGTTGACTCGAAGGCCGACCCGAATGAGAAGCAACACAAGCAACATT
CGACTGTTGTTTGCTCGTCAAGCTATGCTCAATCCCCATAGCCGTGTCGATGCACTCCTTGAAAGTACTA
AGGCGACCCGCCATCATGTACGTCTATATAGAAGGTCGTAACCCACCAATGAATTTATGGCAATACATCT
TCTCCGTTGACACTTTCTCGGGCGCATACTGAGACAACTCCGAAAATGAGATTTGGTAGTCTTGCACCAA
AGAACTACCTTGAACAAGGCTCGT

304 bp

Fragment: Md-af9

ACGTAATACAATATTGTCCGCTTTGGGCAAAGTTATACGTGCAGTGCTGAAACTTCTCAGGCGATGTGGG
ACGAGTGGTTGCCCCCCCTCGACACTTGTGGTAGCTCGCCCAGCATGGTAGCCTGCTAATGTGCCAGCTC
GACAGCATCTCATGCCCGCTCACCGACGTCCGCCCACACACGCCCGTCCATGCACAACCGGGGTGCTACA
ATTATTTAGCAATATTGATAGCACCCTTGTACAATACAGATGATGATGAGATATGAGACAAGCTCTGAGA
GATGTATGATGGCGCAGAAATGTCGTTACCAAGAGTGAGGAGATTTCTTCTGTGGTAATGGTAAAACATG
TGCTACTG

358 bp

Fragment: Md-af11

CAGGATCCTGCCTGATGGAGGCTCTTTAGTGAACATCTTGCCCTTCCATACTGCAAGGGAGCTTGGAATC
ACATCCGACGAAATGTCTCATAGTCGCTTGATGATCCAAGGTCTCAACCAAGATGGGCAAAAAACTATAG
GGATAATCCGCTTGGACATTCACATAGACGAAATGAGATCAAGCGTCTTGTTCCATGTCATCGACGCGAA
GACCACGTACGATGTGCTCATTGGTCGGCCGT

242 bp

Fragment: Md-af12

CGTCCTACGCGCCAAAGCTGTTCAGGCATGTTGCAAGGTTGTCAGGAGAGCCCACAACTCGACAGGCAGA
GCTATTGGTACAAGCATGGTACTTCGGGTTCCAGGTTGCTCAAGTCTTTCTGGTAACAACATTCTCTTCT
GGAGCTACAGCAGTTGCATCAAGCATCATGAAGAATCCACAATCAGCACCAAATTTGCTTGCTGAGAGTC
TGCCAAAGGCTGCTAATTTCTATCTTACCTACATCATACTGCAAGGCACTGCTGGCGCTGCTGACAATCT
G

281 bp

Fragment: Md-af13

CAGTTTAGGATTCGATCTTATCTTTGGCATTGATTGGTTATCAAACCATCAAGCCATAATTGATTGCGGA
AAACGTCAAGTCACTCTAAATACATCCTCTAGAGAACGAATCACCTATCAAGAGGGATTCAATAAACACT
TTCCCAAACTGTCCCACATATTGACCCTTTCTGAGATGATTTCTTATCTTGGAGAGATTTCCATATCAAA
TTCATCACTTCCTCCAATAGTTCG

234 bp

Fragment: Md-af15

ACGTAATACAATATTGTCCGCTTTGGGCAAAGTTATACGTGCAGTGCTGAAACTTCTCAGGCGATGTGGG
ATGAGTGGTTGCCCCCCCTCGACACTTGTGGTAGCTCGCCCAGCATGGTAGCCTGCTAATGTGCCAGCTC
GACAGCATCTCATGCCCGCTCGCCGACGTCCGCCCACACACGCCCGTCCATGCACAACCGGGGTGCTACA
ATTATTTAGCAATATTGATAGCACCCTTGTACAATACAGATGATGATGAGATATGAGACAAGCTCTGAGA
GATGTATGATGGCGCAGAAATGTCGTTACCAAGAGTGAGGAGATTTCTTCTGTGGTAATGGTAAAACATG
TGCTACTG

358 bp

Fragment: Md-af17

CAGGATCCTGCTTGATGGAGGCTCTTCAGTGAACATTTTTCCCCTCTGTACTGCAAGGGAGCTTGGAATC
ACAGCCGATGAAATGTCTCCTAGTCTCTTGATGATCCAAGGTTTCAACCAATATGGGCAAAGAGCTATAG
GGATAATTTGCTTGGACCTTCACATAGACGAAATGAGATCAAGCGTCTCATTCCATGTCACCGACACGAA
GACCATGTACAATGTGCTCCTTGGTCGGCCGT

242 bp

DNA Sequences of Nothofagus AFLP fragments

Fragment: Nm-af2

CTATCAGAACTGGTATTCTGTAAAGAAAATGCAAATTTCAATGGAAGACATTTATCTGGCCAGAGAAAAC
TAGATTCCACTACATAATGGAAAGTGTTTACCTAGTGAGGAGTGACACTGATTCCCCTGAAGGAACTTGA
TCAAGCTTCAATGGGTGTCTCTAGAACAAAAGAATGATAATCCCACGAACCTTTTCTTCTCTCCTCAACT
TCTGTTGACACAAAATTTTCCTTTTCCATTTTAGCTCGTGCTTGAACTACTTGAAGGTTGTAGTCTTGAT
TGTTCTTCACCTGCCATTCTTTTACAACAGGACAAGACAAGACAACCTTTTTTTACCTCATCTACTTCTG
TCGTTGAAGTCATGGTACATCGAACATGTATATATATGACACTCATTG

398 bp

Fragment: Nm-af4

CTATCAGANCTGGTATTCTGTAAAGAAAATNCAATTTTCAATGAAAGNCATTTATCTGGCCAGAGAAAAC
TAAATTCCACTACATAATGGAAAGTGTTTACCTAGTGAGGAGTGACACTGATTCCCCTGAAGGAACTTGA
TCAAGCTTCAATGGGTGTCTCTAGAACAAAAGAATAGTAATCCCACGAACCTTTTCTTCTCTCCTCAACT
TCTGTTGACACAAAATTTTCCTTTTCCATTTCCATTTCCATTTTAGCTCGTGCTTGAACTACTTGAAGGT
TGTAGTCTTGATTGTTCTTCACCTGCCATTCTTTTACAACAAGACAAGACAAGACAACCTTTTTTTACCT
CATCTACTTCTGTCGTTGAAGTCATGGTACATCGAACATGTATATATATGACACTCATTG

410 bp

Fragment: Nm-af7

AATAAGCCACAAAGGCAGAAACAAAGAGAAGGAAAGATACAAGACTTTCGTGTTGAAGACCTCATTATCT
GATTCAAAGACAGCAACCAGACCTTCCAAACTCTGACACATGAAGACTGCTTGCCTCAAAGAGAGAGCTT
CTGCCATATGGGCTGATTTTGCCTCCAGAGAAAAAGCACTACCATACGAACTTTTTATTGCTGCAAGATT
CAGACCATTATGCAGGCTTCTCAAGCATGGACTCTGCTTGGCTTGGAGTTTGGACCAAGAAAATAGTAAA
AAATAAACAAAAAAATTGCAGTATATAAACCACGCAACTAATACACTGAAACAGAATAG

339bp

Fragment: Nm-af8

AATTGCATCATCAATATCATCCATATCAACTGAACCGCTTTGTTTTGAAGAATTATCACCACAAGCAAAG
CTGTATGCCAAAAAATAAAGTGGTAGCTGTAAACATAACAGATCCTACCAACTCAGCAAGCTCCTCCACA
TCAGCAAAAAATCCAGCAGCATTGGGTCCCTCCACCTCAATCACTGTCCTGTCTTGAACATCACCAAAGT
CTTGTACAGCTGCATTGAGCCTTTCCCCATGTGTAGCCGTTGAGAAGAGTATGGCAGCCACTGTTGTGAC
AGTCCCTGCATCAGTATGCAGCCCCTGCAATCAAGGAAAATAGTTGAGGACCCTTGTAACTAG

343 bp